Australian Telecommunications Research Institute

Performance Analysis Of Adaptive Arrays With Projected Perturbation Sequences

Steven Antony Ivandich

"This thesis is presented as part of the requirements for
the award of the Degree of Doctor of Philosophy
of the
Curtin University of Technology"

October 1999

Acknowledgements

I take this opportunity to thank the people who have contributed in the production of this thesis.

Importantly I like to thank my supervisors Professor Antonio Cantoni and Dr Yee Hong Leung.

Professor Cantoni for initially suggesting that I enrol in the PhD degree, for his patience over the years to get to this point and for his calmness in going through the many revisions. It has been a long long road but finally it is over. Part-time again? I don't think so.

Thanks Yee Hong for the extra help in refining those alarmingly difficult Chapters.

I especially like to thank my Family and Friends who make all things worthwhile.

Abstract

Perturbation techniques are useful in the design of low complexity adaptive antenna arrays for estimating the gradient required in stochastic descent algorithms. Implementing projected perturbation sequences in an adaptive array allows the simultaneous reception of signals and the adaptation of the array weights while preserving the constraints imposed on the array weights.

This thesis quantifies the performance of narrowband adaptive array processors that employ projected perturbation techniques. For different perturbation receiver structures the performance is determined under idealised conditions and importantly also when practical implementation issues are taken into account.

The arrays performance is characterised by analysing the transient performance of the weight covariance matrix and by determining the misadjustment. By drawing similarities between two established analysis techniques a new misadjustment analysis technique is introduced.

Practical implementation can impact on the arrays performance such that the benefit of the projected perturbation approach is lost. By characterising the array's sensitivity to perturbation noise additional projections which counteract some implementation effects are identified. The level of loss of performance due to weight quantisation and the limited dynamic range of the array weights is determined.

Contents

Notation

Symbols	*Description*
a*	conjugate of the complex scalar a
$\mathbf{A}^H$	hermitian or conjugate transpose of matrix **A**
$\mathbf{A}^T$	transpose of matrix **A**
Re(a)	the real part of complex variable a
Im(a)	the imaginary part of complex variable a
$\mathbf{I}_{LL}$	the $L \times L$ identity matrix
$E[x]$	the expected value of x
$Tr[\mathbf{A}]$	the trace of the matrix **A**
$\forall$	for all
∇	gradient operator
$\frac{\partial^m}{\partial\theta^m}\rho(\theta)$	the m^{th} partial derivative of $\rho(\theta)$ with respect to θ
$\left.\frac{\partial^m}{\partial\theta^m}\rho(\theta)\right\|_{\theta = \theta_0}$	$\frac{\partial^m}{\partial\theta^m}\rho(\theta)$ evaluated at $\theta = \theta_0$
H^+	the set of $L \times L$ positive semidefinite matrices.
H	the set of $L \times L$ complex Hermitian matrices.
$\tilde{q}$	the Hilbert transform of q
i.i.d	independant and identically distributed
w.r.t	with respect to

Author's Publications

"Performance Analysis of Narrowband Adaptive Arrays Using Projected Perturbation Sequences", ISSPA 92, Signal Processing and its Applications. Gold Coast, Australia, pp 581-585, August 16-21, 1992. (with A. Cantoni)

"Performance Analysis of Narrowband Adaptive Arrays Using Projected Perturbation Sequences", IEEE Trans. Antennas Propagat., Vol. 41, no. 5,pp 625-634, May 1993. (with A. Cantoni)

"Quantisation Error Modelling of Narrowband Adaptive Arrays Using Projected Perturbation Sequences". International Conference On Acoustics, Speech and Signal Processing, Adelaide Australia, Volume 2 pp 309-312., April 1994

"Bandwidth Performance of a Narrowband Adaptive Array", Masters Preliminary Report, Department of Electrical and Electronic Engineering, The University of Western Australia 1989.

Chapter 1

Introduction

1.1 Background and Motivation

An adaptive array consists of a set of spatially distributed sensors, the outputs of which can be combined and processed in an adaptive manner so as to optimise the performance of the system subject to some criterion [9], [13], [33], [53], [59], [60], [61]. By having spatially distributed sensors the adaptive array is able to exploit the spatial characteristics of the incident signals as well as temporal characteristics to enhance system performance.

Adaptive antenna arrays were proposed, as early as the 1960s, for use in military applications. Adaptive arrays can be used to estimate the signal scenario by determining the number of signal sources that are present, the signals strengths and their direction of arrivals. And an adaptive array can be used to improve the desired received signal quality through spatial filtering of interference signals and by maximising the signal to interference and noise ratio on the antenna output. The use of adaptive arrays finds wide applications in fields such as acoustics, seismology and communications [64], [65], [66], [67], [75]. Recently there has been a surge of interest in applying adaptive arrays in cellular telecommunications systems [77], [78], [79], [80], [81], [82]. In cellular applications the use of adaptive arrays has the potential to increase the capacity of a system by improving the interference and noise suppression capabilities and reducing co-channel interference and multipath propagation effects.

The application of perturbation sequences to adaptive beamforming enables the design of low complexity adaptive antenna arrays. In a perturbation based adaptive antenna array individual array element signals are inaccessible and perturbation techniques are required to estimate the gradient used in the weight update algorithm. In a conventional array where all the array element signals are accessible the required gradient vector can usually be determined by correlating the array output with the array element signals. Compared to a conventional array an adaptive array using perturbation techniques obviates the need to have a coherent measurement of the signals on all the array elements that have adaptively controlled weights, this leads to a considerable reduction in the array hardware and the complexity and cost of the system. Perturbation based arrays also provide advantages in dealing with phase and offset errors and dynamic range issues [9] [68].

Perturbation based adaptive arrays have been studied by Widrow et. al. [1] who described a gradient based algorithm which obtained a gradient estimate by measurement of changes in smoothed output power resulting from individual weight perturbations. Cantoni [2] modified the technique by correlating the instantaneous output power of a single receiver system (or the instantaneous power difference sequence of a dual receiver systems), with an orthogonal perturbation sequence. Davis et. al. [3] introduced the coherent perturbation method in which closely spaced samples of the output power were taken to estimate the baseband voltage of the array elements, this method required the adjacent time samples to be correlated. A number of authors have contributed to the analysis and theory of perturbation techniques in adaptive arrays [3], [4], [5], [6], [46], [53].

The use of constraint preserving perturbation sequences, referred to here as projected perturbation sequences was introduced in [4]. Implementing projected perturbation sequences in an adaptive array allowed the simultaneous reception of signals and the adaptation of the array weights while maintaining the imposed constraints on the array weights. The use of projected perturbation sequences was demonstrated on an experimental narrowband adaptive array developed at the University of Newcastle [36], [37], [38]. A project was initiated within the Electrical and Electronic Engineering department at the University of Western Australia and then at the Australian Telecommunication Research Institute at Curtin University to extend the efforts conducted in Newcastle.

Within the context of low complexity adaptive arrays our work is focused on quantifying the performance of narrowband adaptive arrays that employ the projected perturbation techniques. We determine the performance of adaptive arrays under idealised conditions and importantly also when practical implementation issues are taken into account. Specifically, we develop expressions for the performance of an adaptive array with different array perturbation receiver structures, determine the impact that practical implementation effects have on the performance, and examine which projections are required on the perturbation sequence to achieve a desired noise response and to maintain the original constraint preserving response when directional mismatch and practical implementation effects occur.

A modified version of the LMS algorithm is used to update the array weights in the perturbation based arrays we investigate. The LMS algorithm has been considered quite extensively for the constrained and unconstrained case in adaptive beamforming [1], [2], [5], [6], [7], [8], [9], [10], [62], [63], [69], [70], [73]. The analysis that exists covers the transient behaviour of the weights, the convergence, the transient behaviour of the weight covariance matrix and the misadjustment.

For conventional adaptive arrays when orthogonal perturbation sequences are used to estimate the required gradient for the constrained LMS algorithm the study of the convergence and transient behaviour of the weights [5] and the transient behaviour of the weight covariance matrix and the misadjustment [6] has been carried out. Webster Evans and Cantoni proposed the use of projected perturbation sequences in adaptive beamforming [4]. Importantly these sequences allowed simultaneous adaptation and reception by use of weight perturbations that do not obstruct the look direction constraint. These sequences can also be shorter in length and offer computational savings. In Webster's et. al. [4] performance analysis of an adaptive array using these sequences, the analysis was limited to considering computational aspects and deriving expressions for the gradient estimate and the excess mean output power due to perturbations, all for a single perturbation receiver structure. We extend their performance characterisation by determining the gradient covariance, the transient analysis of the weight covariance matrix and the misadjustment for three adaptive perturbation receiver structures which use projected orthogonal perturbation sequences.

Finite precision implementation of an adaptive algorithm can have significant impact on the performance of the array. The study and characterisation of finite precision effects when implementing the LMS algorithm has been studied [13], [14], [20], [21], [22], [33], [45], [72], [76]. However, little work has appeared on the effects of finite precision on adaptive arrays using perturbation techniques. It is expected that the weight quantisation effects of a perturbation based adaptive array will be similar to that experienced on a conventional adaptive array, however when the quantisation effects due to implementing the perturbation sequence and in particular the projected perturbation sequence can not be minimised by design [5], the performance between the two arrays will differ considerably. For a perturbation based array Hudson [14] has considered characterising the interference rejection capability of the processor and Webster [36] has examined the use of dynamic scaling on the weights to minimise the quantisation error. We extend the idealised performance characterisation and develop new expressions for the gradient covariance and the misadjustment in the presence of weight quantisation for different perturbation receiver structures.

The application of derivative constraints to adaptive antenna arrays has been used to achieve controlled main beam responses and to improve the reception of desired signals when directional mismatch occurs [9], [13], [18], [28], [29], [35], [42], [71], [73]. Directional mismatch occurs when the desired signal's direction of arrival is offset from the expected direction of arrival. A flat main beam response can be achieved by setting the derivative of the beamformer response or the magnitude response in a specified direction to be zero [51]. The body of work that exists on the application of derivative constraints to adaptive arrays is mainly concerned with defining these constraints, how to implement the constraints simply and effectively and what are the limitations of these implementations. No work has appeared on the application of derivative constraints to projected perturbation sequences used in adaptive arrays. When directional mismatch occurs the inherent properties of using projected perturbation sequences, such as satisfying the system constraints and reducing the perturbation noise can be degraded. In this instance, applying a spatial derivative constraint in the generation of the projected perturbation sequence can preserve the original system response and reduce the perturbation noise that is introduced into the system. We examine the benefit of applying spatial derivative constraints in the generation of the projected perturbation sequences by developing new expressions for an array's sensitivity to perturbation noise, and identify conditions on the array under which spatial derivative constraints applied to perturbation sequences are effective in suppressing perturbation noise.

1.2 Contribution of the Thesis

The major contributions of this thesis are:

- The characterisation of the performance of the projected perturbation approach by analysing the transient performance of the weight covariance matrix and determining the misadjustment for three narrowband array receiver structures under idealised conditions. We also derive expressions for the gradient covariance.

- The introduction of a new misadjustment analysis technique.

- By considering practical implementation issues, we extend the system performance characterisation to include weight quantisation effects. In particular, we develop new expressions for the gradient covariance and the misadjustment in the presence of weight quantisation for three narrowband array receiver structures. We determine the level of loss of performance due to weight quantisation and the limited dynamic range of the array weights for an adaptive array that uses the projected perturbation approach.

- The development of new expressions for an array's sensitivity to perturbation noise to determine the benefit of projecting the perturbation sequence onto the spatial derivative constraint planes. We identify conditions on the array under which spatial derivative constraints applied to perturbation sequences are effective in suppressing perturbation noise.

1.3 Organisation of the Thesis

The thesis is organised as follows:

In Chapter 2 we place in perspective the area of research on the use of projected perturbation sequences. We introduce the notation and model assumptions used throughout the thesis, briefly review narrowband time domain beamforming techniques using the perturbation approach, define the measures that we use to characterise an arrays performance and present results on the performance of an adaptive array using orthogonal and projected orthogonal perturbation sequences.

In Chapter 3 we present new results that characterise the performance of the projected perturbation approach by determining the misadjustment for three perturbation receiver structures. We determine the misadjustment using two established approaches the *Direct* and *Bounds* approach developed in [5] and [6] and we present a new misadjustment analysis technique. We also determine the optimum perturbation step size and analyse the transient performance of the weight covariance matrix. The misadjustment is determined under idealised conditions and we present the results of simulation studies to verify the accuracy of the misadjustment expressions. For the sake of conciseness, only the basic approach and an outline of major intermediate results are presented in the chapter. A detailed derivation of the intermediate results can be found in Appendix B, C and D.

Appendix A contains the lemmas and theorems that are required in the derivation of the results contained in the thesis.

Chapter 4 extends the performance characterisation of the projected perturbation approach to include digital implementation effects. Specifically, for three perturbation receiver structures we develop new expressions for the gradient covariance and the misadjustment which allow for weight quantisation and we determine the level of loss of performance due to weight quantisation and the limited dynamic range of the array weights. Simulation results are presented to determine the accuracy of the derived expressions. Again for the sake of conciseness, only the major intermediate results are presented in the chapter. A detailed derivation of the intermediate results can be found in Appendix E.

In Chapter 5 we examine the use of spatial derivative constraints in projected perturbation sequences. The application of spatial derivative constraints is useful in countering the effects of directional mismatch. By projecting the perturbation sequence onto spatial derivative constraint planes unwanted perturbation noise (generated by directional mismatch condition) in the output signal can be reduced. When directional mismatch is not a concern, projection of the perturbation sequence onto the spatial derivative constraint planes can be used to reduce the perturbation noise contributed by signals within an angular region of the look direction. We determine the benefit of using the spatial derivative constraints in the projection operation by developing new expressions for an array's sensitivity to perturbation noise. We identify conditions on the array under which spatial derivative constraints applied to perturbation sequences are effective in suppressing perturbation noise.

Chapter 2

Overview of Narrowband Array Processing Using Perturbation Sequences

2.1 Introduction

In this chapter we give a brief introduction to narrowband time domain beamforming techniques using the perturbation approach. The aim is to place in perspective the area of research on the use of projected perturbation sequences.

We outline the optimum array processing problem and present the signal model, the narrowband array, the perturbation receiver structures and the basic assumptions and definitions used throughout the work. Measures that are used to characterise an array's performance are introduced and we present some existing results, established in [2],[4],[36] and [37], of an array's performance with orthogonal and projected orthogonal perturbation sequences. However these existing results give only a limited analysis of the performance and in later chapters we will extend these results to obtain a more comprehensive characterisation of an array's performance using the projected perturbation approach.

The chapter is organised as follows. In Section 2.2 the propagation and received signal model are introduced. In Section 2.3 we present the narrowband beamformer and the signal representation. We identify the *Narrowband Condition* used for defining narrowband signals in the context of narrowband array processing in Section 2.4. In Section 2.5 we present the inphase and quadrature signal representation. We introduce the vector notation used throughout the thesis in Section 2.6. In Section 2.7 we define the adaptive beamforming problem that is considered throughout the thesis. In Section 2.8 we review a class of narrowband beamformers using the perturbation approach. Here three perturbation receiver structures are presented and the properties of an orthogonal perturbation sequence are defined. In Section 2.9 we introduce measures used to characterise an array's performance and finally in Section 2.10 we review a projected perturbation sequence and summarise the work that has been done to characterise an array's performance using the projected perturbation approach.

2.2 Propagation and Received Signal Model

In this thesis, we assume that the array consists of L sensor elements. The elements are located in the far field of a number of point sources, $s_i(t)$ $i = 1, 2...K$, such that the wavefront impinging on the array, due to any source, can be considered to be a plane wave. As shown in Figure 1, the normal to a plane wave gives the direction of arrival of the source.

The position of the array sensor elements is described in terms of a Cartesian co-ordinate system as shown in Figure 1. The origin of the co-ordinate system can be chosen arbitrarily, for convenience we take the origin of the co-ordinate system as the time reference point relative to which all propagation times are calculated.

The output signal waveform, $x_l(t)$, of the l^{th} element of the array due to a far field point source $s(t)$, before beamforming, is assumed to be related to the scalar signal $s(t)$ by [11]

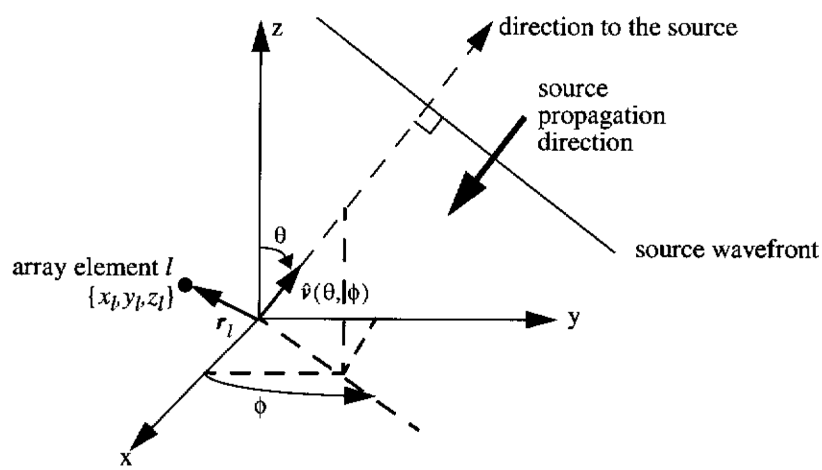


Figure 1.Definition of the Co-ordinate System

$$x_l(t) = \int_{-\infty}^{t} o_l(t-\tau)s(\tau)d\tau + n_l(t) \tag{2.1}$$

where $o_l(t)$ represents the effects of propagation from the signal source to the l^{th} element of the array and the response of the l^{th} element and $n_l(t)$ is the total noise at the output of the l^{th} element due to both internal electronic noise and external background noise.

We assume the following throughout the thesis:

Assumption 2.1: $o_l(t)$ corresponds to a pure time delay for the ideal case of non-dispersive propagation and array elements which are distortionless and omni-directional.

Using *Assumption 2.1*, for a single point source, $s(t)$, incident on the array, the expression for the output waveform of the l^{th} element before beamforming is given by

$$x_l(t) = s(t + \tau_l(\theta, \phi)) + n_l(t) \quad l=1,2,...L \tag{2.2}$$

where (θ, ϕ) is the direction of arrival of the plane wave, $\tau_l(\theta, \phi)$ is the propagation delay of the plane wave relative to the origin of the co-ordinate system defined by

$$\tau_l(\theta, \phi) = \frac{\boldsymbol{r}_l \cdot \hat{\nu}(\theta, \phi)}{\upsilon} = \frac{1}{\upsilon}[(x_l \cos\phi + y_l \sin\phi)\sin\theta + z_l \cos\theta] \tag{2.3}$$

and where $\boldsymbol{r}_l$ is the position vector of the l^{th} element of the array, $\hat{\nu}(\theta, \phi)$ is a unit vector in direction (θ, ϕ) as shown in Figure 1, υ is the speed of propagation of the plane wave and (x_l, y_l, z_l) is the position of the l^{th} element.

2.3 Signal Representation

The general structure of the narrowband beamformer considered in this thesis is shown in Figure 2. Functionally, the narrowband beamformer can be separated into four modules, an input filtering module, the inphase and quadrature signal generation module, the weighting module and the adaptive processor module.

The input filtering module is used to pass the frequency band of interest. It is possible that this function is unity or trivial. The weighting module is where the signals are combined to produce an output, and the adaptive processor module is used to perform the weight update algorithm.

In the inphase and quadrature signal generation module, the inphase and quadrature components of the array element signals are generated. This stage can be performed at the element stage and providing certain conditions hold at an intermediate frequency stage or at baseband, [11], [57]. At the element stage, the inphase and quadrature signal generation occurs with no frequency shifting. This can be accomplished with the use of quadrature filters. At an intermediate frequency stage or at baseband, the inphase and quadrature signal generation requires some frequency shifting. There are numerous ways that the inphase and quadrature signals can be generated such as with quadrature filters, quadrature frequency shifting and by the use of Fast Fourier Transforms [11],[33],[54],[56].

Two possible quadrature signal generators are shown in Figure 3 and Figure 4. In Figure 3 the inphase and quadrature signals at baseband are generated using quadrature frequency shifting. In Figure 4 the inphase and quadrature signals at an intermediate frequency stage are generated by frequency shifting and using a quadrature filter. In Figure 4, f_{IF} is the intermediate frequency and in the thesis we sometimes refer to this frequency as the mean frequency of the signal at that stage.

We assume throughout the thesis that the filtering stages at the input or those required in the inphase and quadrature signal generation can be realised.

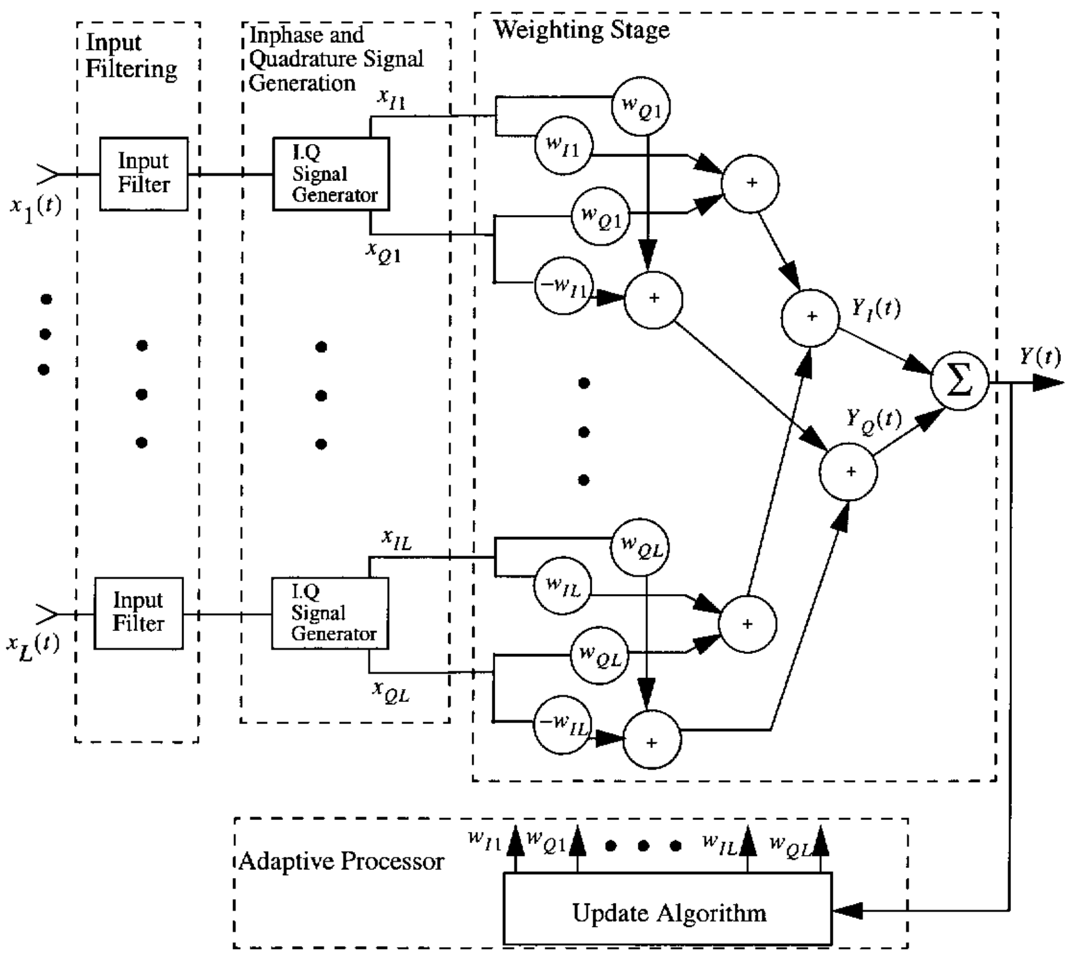


Figure 2. Narrowband Antenna Array Structure

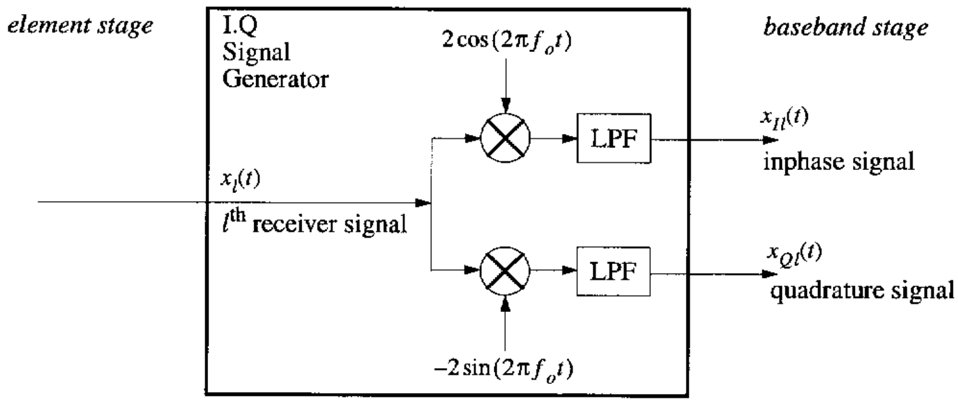


Figure 3. Inphase and Quadrature Signal Generator at Baseband

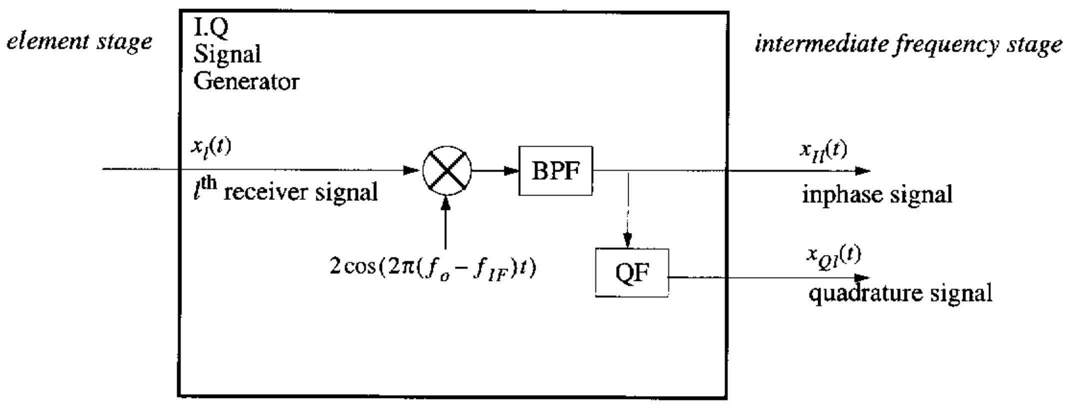


Figure 4. Inphase and Quadrature Signal Generator at an Intermediate Frequency

When using the inphase and quadrature components of the array element signals, it is convenient to represent the signals in the form of a complex envelope. The complex envelope of $s(t)$ at frequency f_o is given by [15]

$$m(t) = m_I(t) + jm_Q(t) = (s(t) + j\tilde{s}(t))e^{-j2\pi f_o t} \tag{2.4}$$

where $m_I(t)$ and $m_Q(t)$ represent the real and imaginary parts of the complex function $m(t)$ and $\tilde{s}(t)$ is the Hilbert transform of $s(t)$. By re-arranging (2.4) the signal $s(t)$ can be expressed in terms of its complex envelope, i.e.

$$s(t) = m_I(t)\cos(2\pi f_o t) - m_Q(t)\sin(2\pi f_o t). \tag{2.5}$$

Note from (2.4) that $m_I(t)$ and $m_Q(t)$ depend on the choice of f_o.

By substituting (2.5) into (2.2), for a single point source, $s(t)$, incident on the array, the output of the l^{th} receiver element before beamforming is given by

$$x_l(t) = m_I(t+\tau_l)\cos(2\pi f_o(t+\tau_l)) - m_Q(t+\tau_l)\sin(2\pi f_o(t+\tau_l)) + n_l(t) \qquad (2.6)$$

2.4 Narrowband Condition

In the case of purely temporal processing, a signal is considered narrowband if the bandwidth of the signal is small compared to the mean frequency of the signal. However, narrowband array processing also involves spatial sampling of the signal. Therefore, in an array processing context, for a signal to be considered as narrowband, it must satisfy an additional condition that allows the separation of the signal's time and space dependent quantities [9], [11], [16], [33]. When this condition, *Narrowband Condition*, is satisfied, the resulting analysis for narrowband arrays is more tractable.

Narrowband Condition: The complex envelope of the signal is constant across the array at any given instant of time, i.e. $m(t+\delta_t) \cong m(t)$. (2.7)

Note that the above condition is a sufficient condition. It is commonly referred to as the array aperture bandwidth condition. In [2], [33] and [34] the array's performance is determined when (2.7) is not satisfied.

In (2.7), δ_t is the order of time difference between the signal arriving at two elements situated closest and furthest from the point source, defined by $\delta_t = \delta_a/\upsilon$ where δ_a is the maximum aperture size of the array. It is established in [9],[11],[16],[33] that a sufficient condition for (2.7) to be true is that the bandwidth, δ_f, of the complex function $m(t)$ satisfies

$$\delta_f \ll \frac{\upsilon}{\delta_a}. \qquad (2.8)$$

Assuming *Narrowband Condition* holds, (2.6) can be re-written as

$$x_l(t) = m_I(t)\cos(2\pi f_o(t+\tau_l)) - m_Q(t)\sin(2\pi f_o(t+\tau_l)) + n_l(t). \qquad (2.9)$$

2.5 Inphase and Quadrature Signal Representation

As mentioned in Section 2.3, the inphase and quadrature signal generation can be performed at the element stage and providing certain conditions hold at an intermediate frequency stage or at baseband. In [11] it is shown that these conditions are:

Inphase and Quadrature Condition 1: When quadrature signal generation is performed at a frequency shifted stage, the mean frequency of the signal at that stage, f_{IF}, should be greater than the highest frequency component of $m_I(t)$ and $m_Q(t)$.

Inphase and Quadrature Condition 2: If the quadrature signals are generated at baseband using quadrature filters then, $m_Q(t) = \tilde{m}_I(t)$

Note that the above conditions are sufficient conditions [11].

Using the *Inphase and Quadrature Conditions,* the inphase and quadrature signals of the l^{th} array element at the beamforming stage when more than one signal is incident on the array, are given by

$$x_{Il}(t) = \sum_{i=1}^{K} [m_{Ii}(t)\cos(2\pi(f_{IF}t + f_o\tau_{li})) - m_{Qi}(t)\sin(2\pi(f_{IF}t + f_o\tau_{li}))] + n_{Il}(t) \tag{2.10}$$

$$x_{Ql}(t) = \sum_{i=1}^{K} [m_{Qi}(t)\cos(2\pi(f_{IF}t + f_o\tau_{li})) + m_{Ii}(t)\sin(2\pi(f_{IF}t + f_o\tau_{li}))] + n_{Ql}(t) \tag{2.11}$$

where $n_{Il}(t)$ and $n_{Ql}(t)$ are the inphase and quadrature noise components, $m_{Ii}(t)$ and $m_{Qi}(t)$ are the inphase and quadrature components of the complex modulating function for the i^{th} source and τ_{li} is the propagation time from the i^{th} source. Note that in (2.10) and (2.11) when beamforming is performed at the element stage $f_{IF} = f_o$, and when beamforming is performed at baseband $f_{IF} = 0$.

Without loss of generality and for simplicity of notation, we assume throughout the rest of the thesis that the inphase and quadrature signal generation is performed at baseband as shown in Figure 3.

2.6 Vector Notation for Narrowband Beamforming

In this section we introduce the vector notation adopted throughout the thesis.

We let the inphase and quadrature components of the array element signals be represented by an $L \times 1$ dimensional complex vector $X(t)$, its l^{th} component being defined by

$$X_l(t) = x_{Il}(t) + jx_{Ql}(t) \quad (2.12)$$

where $x_{Il}(t)$ and $x_{Ql}(t)$ are defined by (2.10) and (2.11) respectively.

Throughout the thesis, for convenience, we also adopt the use of the following steering vector notation, [9], to represent the array signals. The components of the steering vectors are the phase shifts of the sinusoidal signals produced in each of the array elements relative to a reference when a monochromatic plane wave is incident on the array from a specified direction. The $L \times 1$ dimensional complex steering vector in direction (θ,ϕ) is given by

$$S^T(\theta, \phi) = [e^{2\pi f_o \tau_1(\theta, \phi)}, e^{2\pi f_o \tau_2(\theta, \phi)}, \ldots\ldots\ldots\ldots, e^{2\pi f_o \tau_L(\theta, \phi)}] \quad (2.13)$$

Using the above steering vector notation, (2.10), (2.11) and assuming there is no noise present in this instance, the array signal at the beamforming stage is given by

$$X(t) = \sum_{i=1}^{K} m_i(t) S(\theta_i, \phi_i) \quad (2.14)$$

Let the array weights be represented by an $L \times 1$ dimensional complex weight vector W, with its l^{th} component being defined by

$$W_l = w_{Il} + jw_{Ql} \quad (2.15)$$

where w_{Il} and w_{Ql} are defined in Figure 2. The output of the array processor is given by the product of the signal and weight vectors, i.e

$$Y(t) = X^H(t) W \quad (2.16)$$

The output power of the narrowband processor is defined by

$$P(W(n)) = E[Y^2(t)] = W^H(n) \mathbf{R} W(n) \quad (2.17)$$

where $\mathbf{R}$ is the $L \times L$ dimensional complex array correlation matrix defined by

$$\mathbf{R} = E[X(t) X^H(t)] \quad (2.18)$$

We shall make the following assumption throughout the thesis that will impact on **R**.

Assumption 2.2a: The narrowband signals, $s_i(t)$ $i = 1, 2...K$, are mutually uncorrelated and statistically independent from the noise at any receiver.

Assumption 2.2b: The total noise at the output of the l^{th} array element can be modelled as uncorrelated white element noise

Using *Assumption 2.2a*, the array correlation matrix can be decomposed as

$$\mathbf{R} = \sum_{i=1}^{K} \mathbf{R}_i + \mathbf{R}_\sigma \tag{2.19}$$

where $\mathbf{R}_i$ is the correlation matrix due to the i^{th} source and $\mathbf{R}_\sigma$ is the correlation matrix due to noise on the array elements.

Further under *Assumption 2.2b* the noise contribution to the array correlation matrix is given by $\mathbf{R}_\sigma = \sigma\mathbf{I}_{LL}$ (2.20)

where σ is the white noise power on each array element.

2.7 Adaptive Beamforming

The adaptive beamforming problem considered in the thesis is that of computing in real time a weight vector that converges in some sense to the optimum weight vector, to be explored later, and is defined by the following optimisation problem [2], [4], [6], [9].

$$\min_{W} \quad W^H \mathbf{R} W \tag{2.21}$$

subject to $W^H\mathbf{C} = F$ (2.22)

In the thesis we assume that all the constraints placed on the system are linearly independent and consistent. When there are N constraints placed on the system, $\mathbf{C}$ is a $L \times N$ full rank matrix given by $\mathbf{C} = \left[C_1\ C_2\ ....\ C_N\right]$ and F is an $1 \times N$ row vector.

The constraints will consist of non-zero fixed responses in specified directions known as look directions and constraints that are used to control the array's response over specified regions of direction and or frequency [28], [29].

Let $\mathbf{P}$ be an $L \times L$ projection operator matrix defined by

$$\mathbf{P} = \mathbf{I}_{LL} - \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H \tag{2.23}$$

It can be shown that W the solution of the optimization problem defined by (2.21) and (2.22) satisfies the equation

$$\mathbf{P}\mathbf{R}\hat{W} = 0 \qquad (2.24)$$

and if **R** has full rank,

$$\hat{W} = \mathbf{R}^{-1}\mathbf{C}(\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C})^{-1}F^H \qquad (2.25)$$

The total output power for the optimum weight setting is obtained by substituting (2.25) into (2.17), i.e

$$P_{opt} = F(\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C})^{-1}F^H \qquad (2.26)$$

In practice **R** is unknown and must be estimated from the data. This leads to the development of adaptive algorithms. A real time algorithm [1],[8],[53] for obtaining the optimum weight vector $\hat{W}$ is given by

$$W(n+1) = \mathbf{P}[W(n) - \mu G(W(n))] + \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}F^H \qquad (2.27)$$

where $W(n+1)$ denotes the new weight vector computed at the n^{th}+1 iteration, μ is a positive scalar defining the gradient step size and $G(W(n))$ is an unbiased estimate of the true gradient of $P(W)$ with respect to $W(n)$.

As shown in [7] and [8] the gradient of $P(W)$ with respect to $W(n)$ is given by

$$\nabla_W P(W)|_{W = W(n)} = 2\mathbf{R}W(n) \qquad (2.28)$$

The conventional Least Mean Square (LMS) adaptive algorithm, [9], [17] for adjusting the weights uses the following gradient estimate

$$G(W(n)) = 2X(n+1)X^H(n+1)W(n) = 2X(n+1)Y(n+1) \qquad (2.29)$$

which clearly requires access to the signals $X(n+1)$.

Note that for a given $W(n)$ the estimate given by (2.29) is unbiased, i.e

$$E[G(W(n))|W(n)] = 2\mathbf{R}W(n) \qquad (2.30)$$

Under appropriate conditions on μ, $\lim_{n \to \infty} E[W(n)] = \hat{W}$ [2], [5], [8].

2.8 Adaptive Narrowband Beamforming Using Perturbation Techniques

In the design of low complexity adaptive arrays, perturbation techniques can be used for estimating the gradient required in stochastic descent algorithms [2], [3], [4], [5], [6], [46], [53]. For narrowband arrays, the conventional Least Mean Square algorithm requires the coherent measurement of the signals of $X(t)$ that have adaptively controlled weights [9],[17]. Perturbation techniques obviate the need to measure all these signals and this can lead to considerable reduction in the complexity and cost of the system. To apply perturbation techniques the weights of the array need to be individually adjustable. An evaluation of the degradation in performance by not adapting all the weights can be found in [2], [3], [4], [5], [6], [46], [53].

The basic idea in perturbation techniques is to apply a perturbation sequence to the adjustable weights, and then to estimate the required gradient on the basis of the measurement of the corresponding instantaneous output power sequence or instantaneous power difference sequence obtained with one or two beamformers. The perturbation based processor only requires access to the processor's output signal and the weights to estimate the gradient required in the weight update algorithm.

When using perturbation techniques to estimate the gradient, the adaptive beamforming algorithm consists of two phases. In the first phase an unbiased estimate of the true gradient $2\mathbf{RW}$ is obtained using one of the schemes which are described in Section 2.8.2. The perturbation phase occupies m time instants to yield the gradient estimate, m being the number of perturbation vectors used in the gradient estimation. In the second phase the weights are updated using the gradient projection algorithm (2.27). The algorithm is performed according to the following.

Adaptive Beamforming Algorithm using Perturbations:

Phase 1.(perturbation) Estimate the gradient $G(W(n))$ at $W(n)$ by using one of the three schemes described in Section 2.8.2

Phase 2.(weight update) $W(n+1) = \mathbf{P}[W(n) - \mu G(W(n))] + \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}F^H$ where $G(W(n))$ is the estimated gradient.

In the following we review a class of perturbation techniques that can be used to estimate the gradient required in the weight update algorithm. Firstly a number of definitions regarding an orthogonal perturbation sequence are presented. Three perturbation receiver structures are then presented. The three methods that are of interest are the single receiver system, the dual receiver dual perturbation system and the dual receiver reference receiver system. These three methods offer different trade-offs between complexity and performance.

2.8.1 Orthogonal Perturbation Sequences

The perturbation sequences used in the perturbation receiver structures are required to have certain properties for the gradient estimate to be unbiased. In this section we briefly review these properties and present an orthogonal perturbation sequence, the Time Multiplex sequence, that is used in later analysis. Methods to generate orthogonal perturbation sequences can be found in [2].

Definition 1: Orthogonality

Let $\delta(.) = \{\delta(1), \delta(2), ..., \delta(m)\}$ be a sequence of $L \times 1$ complex column vectors. The sequence $\delta(.)$ is said to be a normalised complex orthogonal vector sequence if

i) $$\frac{1}{m}\sum_{i=1}^{m} Re[\delta(i)]Re[\delta^H(i)] = \mathbf{I}_{LL} \tag{2.31}$$

ii) $$\frac{1}{m}\sum_{i=1}^{m} Im[\delta(i)]Im[\delta^H(i)] = \mathbf{I}_{LL} \tag{2.32}$$

iii) $$\frac{1}{m}\sum_{i=1}^{m} Re[\delta(i)]Im[\delta^H(i)] = 0 \tag{2.33}$$

iv) $$\frac{1}{m}\sum_{i=1}^{m} Im[\delta(i)]Re[\delta^H(i)] = 0 \tag{2.34}$$

The above equations state that the $2L$ scalar sequences comprising of the L scalar sequences forming the components of the real part of $\delta(i)$ and L scalar sequences forming the components of the imaginary part of $\delta(i)$ are mutually orthogonal over a perturbation sequence length m. Each scalar sequence has been normalized to have a unit average power.

Definition 2: Zero Mean

A sequence $\delta(.) = \{\delta(1), \delta(2), ..., \delta(m)\}$ is said to have zero mean if

$$\frac{1}{m}\sum_{i=1}^{m}\delta(i) = 0 \tag{2.35}$$

Definition 3: Odd Symmetry

A sequence $\delta(.) = \{\delta(1), \delta(2), ..., \delta(m)\}$ is said to have odd symmetry if for every i, $1 \le i \le m$ there exists a j, $1 \le j \le m$ such that $\delta(i) = -\delta(j)$

From this definition it follows that an odd symmetry sequence has zero mean and there are only $m/2$ distinct vectors in an odd symmetry sequence, the other $m/2$ vectors are the negatives of these. An odd length sequence can also have odd symmetry if it trivially includes an odd number of zero vectors.

For a Time Multiplex sequence [2], it is possible to evaluate certain expressions in closed form and consequently these sequences are often considered in detail. Two Time Multiplex sequences are examined in the thesis: a minimum length sequence and an odd symmetry sequence as defined below. The minimum and odd symmetry time multiplex sequence are of length $2L$ and $4L$ respectively.

Minimum Length Time Multiplex Sequence

$$\{\delta(1), ..., \delta(2L)\} = S = \sqrt{2L}[\mathbf{I}_{LL}, j\mathbf{I}_{LL}] \tag{2.36}$$

Odd Symmetry Time Multiplex Sequence

$$\{\delta(1), ..., \delta(4L)\} = S = \sqrt{2L}[\mathbf{I}_{LL}, j\mathbf{I}_{LL}, -\mathbf{I}_{LL}, -j\mathbf{I}_{LL}] \tag{2.37}$$

In the above two equations the sequence of complex column vectors in S correspond to the perturbation vectors.

2.8.2 Gradient Estimation using Orthogonal Perturbation Sequences

In this section we summarise the perturbation techniques, developed in [2], that are considered in the thesis. We review some basic results derived in [2] that state the conditions under which an unbiased estimate of the gradient can be obtained, these results are required in later analysis. We also compare the trade-offs between complexity and performance of each perturbation receiver structure.

(a) Dual Receiver Dual Perturbation System

Figure 5 illustrates two dual receiver structures in which perturbation sequences can be used to estimate the gradient. With the switch, *SW*, in position A, the structure corresponds to the dual receiver dual perturbation system. In this system, an estimate of the gradient is obtained by applying a perturbation sequence $\delta(.)$ in antiphase to two sets of weights and correlating the difference in power with $\delta(.)$.

At the i^{th} instant within the perturbation cycle, $1 \le i \le m$, Receiver 1 has its weights perturbed according to

$$W_p(W(n), i) = W(n) + \gamma\delta(i) \quad i=1,2...m \tag{2.38}$$

and Receiver 2 has its weights perturbed according to

$$W_m(W(n), i) = W(n) - \gamma\delta(i) \quad i=1,2...m \tag{2.39}$$

where γ is the perturbation step size that reflects the amplitude of the perturbation applied to the weight vector $W(n)$. To understand the effect of the amplitude of the perturbations in the three systems it is necessary to introduce γ since the perturbation sequence effectively has a normalised amplitude determined by the orthogonality definitions.

An estimate of the gradient is given by

$$G_1(W(n)) = \frac{1}{2\gamma m}\sum_{i=1}^{m}[f_1(W_p, i) - f_2(W_m, i)]\delta(i) \tag{2.40}$$

where $f_1(W_p, i)$ and $f_2(W_m, i)$ denote the instantaneous output powers from Receivers 1 and 2 respectively which are given by

$$f_1(W_p, i) = W_p^H(W(n), i)X(l+i)X^H(l+i)W_p(W(n), i) \tag{2.41}$$

$$f_2(W_m, i) = W_m^H(W(n), i)X(l+i)X^H(l+i)W_m(W(n), i) \tag{2.42}$$

and l is the time instant at which the perturbation cycle is initiated.

We note that in (2.38) to (2.42) n corresponds to the n^{th} iteration of the weight update algorithm, l corresponds to the time at which the perturbation cycle is started for the gradient estimation at the $(n+1)^{th}$ weight update and i corresponds to the i^{th} instant of time within the perturbation cycle.

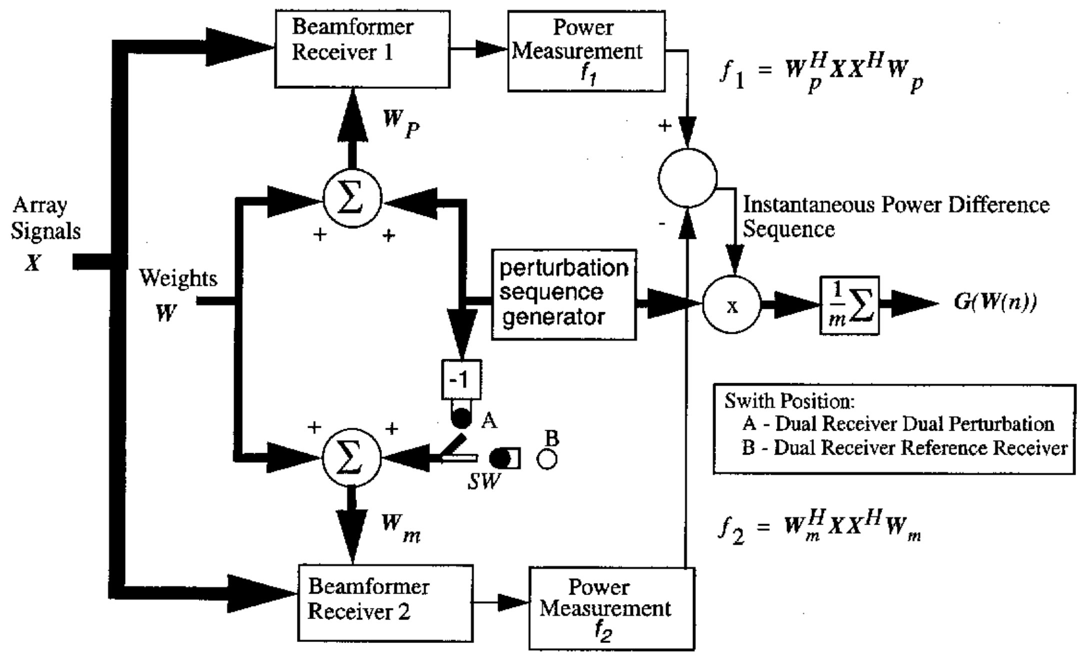


Figure 5. Dual Receiver Perturbation Based Structures

For a given $W(n)$ the gradient estimate given by (2.40) satisfies the following result

Result 2.8.a. Unbiased Gradient Estimate for a Dual Receiver Dual Perturbation System

If the perturbation sequence $\delta(.)$ is orthogonal then for any $\gamma > 0$ the conditional mean of the gradient estimate defined by (2.40) satisfies

$$E[G_1(W(n))|W(n)] = 2\mathbf{R}W(n)$$

❑❑❑

(b) Dual Receiver Reference Receiver System

In Figure 5 with the switch, *SW*, in position B, the structure corresponds to the dual receiver reference receiver system. In this system, an estimate of the gradient is obtained by applying a perturbation sequence $\delta(.)$ to one set of weights, Receiver 1, while the second set of weights remains at a nominal value W.

At the i^{th} instant within the perturbation cycle, $1 \le i \le m$, Receiver 1 has its weights perturbed according to

$$W_p(W(n), i) = W(n) + \gamma\delta(i) \quad i=1,2......m \tag{2.43}$$

An estimate of the gradient is given by

$$G_2(W(n)) = \frac{1}{\gamma m}\sum_{i=1}^{m} [f_1(W_p, i) - f_2(W(n), i)]\delta(i) \tag{2.44}$$

where $f_1(W_p, i)$ and $f_2(W(n), i)$ denote the instantaneous output powers from Receivers 1 and 2 respectively. $f_1(W_p, i)$ is given by (2.41) and $f_2(W(n), i)$ is given by

$$f_2(W(n), i) = W^H(n)X(l+i)X^H(l+i)W(n) \tag{2.45}$$

For a given $W(n)$ the gradient estimate given by (2.44) satisfies the following result

Result 2.8.b. Gradient Estimate for a Dual Receiver Reference Receiver System

If the perturbation sequence $\delta(.)$ is orthogonal then for a given W and for any $\gamma > 0$ the conditional mean of the gradient estimate defined by (2.44) satisfies

$$E[G_2(W(n))|W(n)] = 2\mathbf{R}W(n) + b_1(W(n)) \tag{2.46}$$

where the gradient bias is given by

$$b_1(W(n)) = \frac{\gamma}{m}\sum_{i=1}^{m} \delta^H(i)\mathbf{R}\delta(i)\delta(i) \tag{2.47}$$

Additionally, if the sequence $\delta(.)$ has odd symmetry then

$$E[G_2(W(n))|W(n)] = 2\mathbf{R}W(n)$$

❑❑❑

(c) Single Receiver System

Figure 6 shows a single receiver system. In this system, an estimate of the required gradient is obtained by perturbing the array weights about their nominal value W and correlating the instantaneous output power sequence with the perturbation sequence. At the i^{th} instant within the perturbation cycle, $1 \le i \le m$, the weight vector is given by

$$W_p(W(n), i) = W(n) + \gamma\delta(i) \quad i=1,2......m \tag{2.48}$$

An estimate of the gradient is given by

$$G_3(W(n)) = \frac{1}{\gamma m}\sum_{i=1}^{m} [f_1(W_p, i)]\delta(i) \tag{2.49}$$

where $f_1(W_p, i)$ is the instantaneous output power given by

$$f_1(W_p, i) = W_p^H(W(n), i)X(l+i)X^H(l+i)W_p(W(n), i) \tag{2.50}$$

For a given $W(n)$ the gradient estimate given by (2.49) satisfies the following result.

Result 2.8.c. Gradient Estimate for a Single Receiver System

If the sequence $\delta(.)$ satisfies the conditions of orthogonality and has zero mean, then for any $\gamma > 0$ and given W, the conditional mean of the gradient estimate defined by (2.49) satisfies

$$E[G_3(W)|W(n)] = 2\mathbf{R}W(n) + b_2(W(n)) \tag{2.51}$$

where

$$b_2(W(n)) = \frac{\gamma}{m}\sum_{i=1}^{m}\delta^H(i)\mathbf{R}\delta(i)\delta(i) \tag{2.52}$$

Additionally, if the sequence $\delta(.)$ has odd symmetry, then

$$E[G_3(W(n))|W(n)] = 2\mathbf{R}W(n)$$

❑❑❑

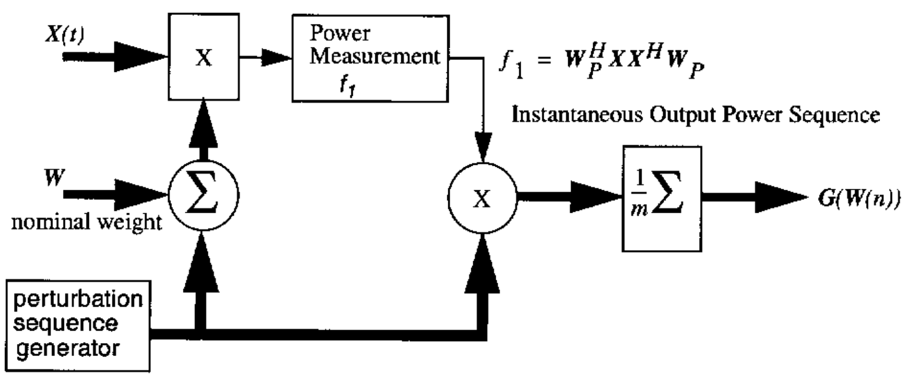


Figure 6. Single Receiver Perturbation Based Structure

(d) Comparison of Receiver Structures

We note from Results 2.8a, 2.8b and 2.8c that for the dual receiver dual perturbation system it is only necessary for the perturbation sequence to satisfy the conditions for orthogonality for the gradient estimate to be unbiased, while for the dual receiver reference receiver and single receiver structures we additionally require the perturbation sequence to have odd symmetry. Thus to obtain an unbiased gradient estimate, the dual receiver dual perturbation system can operate either with a $2L$ length Time Multiplex sequence or the $4L$ length Time Multiplex sequence. For the dual receiver reference receiver and single receiver structures it is necessary to operate with a $4L$ length Time Multiplex sequence.

The length of a perturbation sequence determines the real time rate of convergence of the weights. That is, if fewer perturbations can be used to obtain the same estimate of the gradient at each weight update then the real time required for the adaptive beamforming algorrithm convergence will be reduced. When the dual receiver dual perturbation system operates with the minimum length perturbation sequence it can have an increased real time rate of convergence as compared to the other structures. This faster convergence time is obtained at the expense of increased hardware complexity.

All the perturbation structures result in a noise component in the output signal due to the perturbation of the weights about their nominal values. This additional noise component is referred to as perturbation noise. The dual receiver reference receiver system has an advantage over the dual receiver dual perturbation system in that only one of its receivers is being perturbed which allows a signal free of perturbation noise to be available always. The same hardware requirements as the dual receiver dual perturbation system are still required.

The single receiver system has the simplest hardware requirements of the perturbation structures, requiring only one set of weights and one power measurement stage. While its hardware requirements are less than the dual receiver systems, it is observed that it does suffer from longer real time convergence as compared to the dual receiver dual perturbation system and there is no signal free of perturbation noise during adaptation.

2.9 Perturbation Technique Performance Characterisation

In the thesis we provide new results that characterise the array's performance using the different adaptive perturbation systems. In particular we examine the misadjustment and the excess output power, ξ, due to perturbation noise. Here we present the definitions for the misadjustment and ξ. We also detail the importance of the misadjustment measure and present properties of the perturbation noise for an orthogonal perturbation sequence.

2.9.1 Misadjustment

The misadjustment is a dimensionless measure of how closely the adaptive algorithm approaches the optimum power. The misadjustment is defined by [25]

$$M = \lim_{n \to \infty} \frac{E[P(W(n))] - P_{opt}}{P_{opt}} \tag{2.53}$$

where P_{opt} is the optimum power given by (2.26) and $P(W(n))$ is given by (2.17). The first term in the numerator of (2.53) is the steady state mean output power produced by the adaptive weight adjustment algorithm. The second term is the mean output power that would result if the weights were fixed at the optimum value which is the solution of the constrained beamforming problem. Thus the misadjustment is a normalized measure of the penalty in output power due to weight adaptation.

For a simple signal scenario the misadjustment can also be related more directly to the signal to noise ratio performance of an adaptive array. Consider a scenario of signals incident on the array that consists of a signal of power ρ_0 that arrives from the look direction and interferences. Then, the array signal vector can be expressed as

$$\mathbf{X}(t) = \sqrt{\rho_o} s(t) S(\theta_0, \phi_0) + \mathbf{I}(t) \tag{2.54}$$

where $s(t)$ is a unity variance, zero mean random variable, $S(\theta_0, \phi_0)$ is the steering vector in the look direction (θ_0, ϕ_0) and $\mathbf{I}(t)$ is an $L \times 1$ complex vector representing the interferences. Note we assume there is no white noise present.

It is well known [33] that when the number of elements is greater than the number of directional interferences, a very high attenuation of the interferences can be achieved. Hence, the optimum array power will be nearly equal to the look direction signal power ρ_0, that is,

$$P(\hat{\mathbf{W}}) = \hat{\mathbf{W}}^H \mathbf{R} \hat{\mathbf{W}} \cong \rho_o \tag{2.55}$$

Furthermore, for a single look direction constraint, since the adaptive algorithm ensures that

$$\mathbf{W}^H(n) S(\theta_0, \phi_0) = 1 \tag{2.56}$$

it follows that

$$\lim_{n \to \infty} E[\mathbf{W}^H(n)\mathbf{X}(n+1)\mathbf{X}^H(n+1)\mathbf{W}(n)] = \rho_o + \rho_n \tag{2.57}$$

where ρ_n is the power of the unwanted signal components at the array output.

Substitution of (2.55) and (2.57) in (2.53) yields $M \approx \frac{\rho_n}{\rho_o}$ (2.58)

Thus the output signal to noise ratio, is given approximately by $SNR \approx \frac{1}{M}$ (2.59)

2.9.2 Excess Output Power due to Perturbation Noise

We refer to perturbation noise as the noise that results from the application of perturbation sequences to the array weights. The perturbation noise has been briefly discussed in Section 2.8.

An indication of the effect of the perturbation noise can be obtained by determining the mean excess output power, ξ, due to perturbations of the weights about a nominal value $\boldsymbol{W}(n)$. It is defined as, [2],

$$\xi = \frac{1}{m}E[P_p(\boldsymbol{W}(n), i) - P(\boldsymbol{W}(n))] \tag{2.60}$$

where $P_p(\boldsymbol{W}(n), i)$ is the output power of the receiver at the i^{th} instant of time within the perturbation cycle and is given by (2.41), (2.42), (2.43) or (2.48).

It is shown in [2] that on substitution for any orthogonal perturbation sequence ξ is given by

For the Single Receiver System

$$\xi = 2\gamma^2 Tr(\mathbf{R}) + \frac{\gamma}{m}\sum_{i=1}^{m} 2Re[\delta^H(i)\mathbf{R}\boldsymbol{W}(n)] \tag{2.61}$$

For a Dual Receiver System

$$\xi_+ = 2\gamma^2 Tr(\mathbf{R}) + \frac{\gamma}{m}\sum_{i=1}^{m} 2Re[\delta^H(i)\mathbf{R}\boldsymbol{W}(n)] \tag{2.62}$$

$$\xi_- = 2\gamma^2 Tr(\mathbf{R}) - \frac{\gamma}{m}\sum_{i=1}^{m} 2Re[\delta^H(i)\mathbf{R}\boldsymbol{W}(n)] \tag{2.63}$$

In (2.61), ξ represents the perturbation noise that results when the weights are perturbed according to (2.48). For the dual receiver systems, ξ_+ represents the perturbation noise that results on a receiver that has its weights perturbed positively as given by (2.41) and (2.43). And ξ_- represents the perturbation noise that results when negative weight perturbation are applied to the receiver as given by (2.42).

In the thesis we define the *Look Direction Perturbation Noise* as the mean excess output power due to perturbation that results when only a look direction signal, of unity power, is incident on the array.

Using (2.61) the following results have been established in [2], [4]

Result 2.9.a. Mean Excess Output Power due to Perturbation Noise for a Zero Mean Perturbation Sequence

For any perturbation sequence $\delta(.)$ satisfying the conditions of orthogonality and having zero mean the mean excess power due to perturbations about a nominal weight $\boldsymbol{W}$ is given by

$$\xi = 2\gamma^2 Tr(\mathbf{R}) \quad (2.64)$$

❑❑❑

Result 2.9.b. Look Direction Perturbation Noise for a Zero Mean Perturbation Sequence

For any zero mean orthogonal perturbation sequence, the mean excess power from the constrained look direction due to perturbations about the nominal weight $\boldsymbol{W}$ is given by,

$$\zeta = 2\gamma^2 L \quad (2.65)$$

❑❑❑

2.10 Projected Perturbation Sequences

In [4], a class of perturbation sequences referred to as the projected perturbation sequences were introduced. These sequences permit simultaneous adaptation of weights and signal reception by the use of weight perturbations that do not violate the look direction constraint. They can also be shorter in length than the previously described sequences and offer computational savings, and can be used with all the receiver structures identified earlier. In this section we describe the generation of projected perturbation sequences and include a brief example. Also, to place in perspective the work that has been done by others on the performance analysis of narrowband arrays using these sequences, we review the properties of the projected perturbation sequences and the results established in [4], [36], [37].

2.10.1 Generation of a Reduced set of Perturbation Vectors

The motivation behind the use of projected perturbation sequences can be observed by examining a simple two dimensional problem defined by

$$\min_{W} W^T \mathbf{R} W \text{ subject to } W^T C = 1$$

where C is defined by $C^T = \begin{bmatrix} 1 & 1 \end{bmatrix}$ and $\mathbf{R} = \mathbf{I}_{22}$ (2.66)

The weight vector W is a 2x1 real vector defined by $W^T = \begin{bmatrix} w_1 & w_2 \end{bmatrix}$

and the constraint surface and the constraint subspace are defined by the lines given by $\Gamma = \{W: W^T C = 1\}$ and $\Lambda = \{W: W^T C = 0\}$.

In Figure 7, a weight update is illustrated when a projected and non projected perturbation sequence are used. Figure 7 (a) illustrates the use of the non projected sequence. The initial weight vector, $W(k)$, lies on the constraint surface. The perturbation vectors, defined by S in the figure, are illustrated as weight displacements centred at $W(k)$. When the perturbations are applied it can be observed that the weights are removed from the constraint surface. However by restricting the perturbation vectors to be in the constraint subspace it is clear that this would not occur. These constrained perturbation vectors can be obtained by projecting the perturbation sequence onto the constraint subspace.

Figure 7 (b) illustrates the use of a projected perturbation sequence, defined by T in the figure. The perturbation sequence T does not move the weights off the constraint surface. Hence the desired look direction response is satisfied at all times.

The projected perturbation vectors are derived by projecting the perturbation sequence S onto the constraint plane. Let this new sequence be T', then

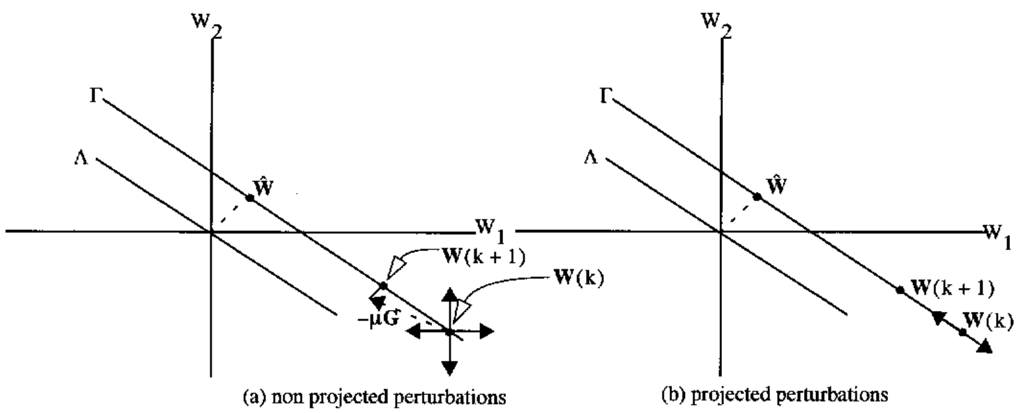
$$T' = [\mathbf{P}\delta(1), \mathbf{P}\delta(2), \ldots, \mathbf{P}\delta(m)] = \mathbf{P}S \quad (2.67)$$

where $\mathbf{P}$ is defined by (2.23). Note that after projection the projected perturbation sequence does not span the entire weight space but is constrained to the solution space of the problem.

Since the projection matrix does not have full rank (in general it has rank L-N) the sequence T' contains a number of redundant vectors. Hence by using a reduction method a reduced set of vectors T can be obtained which is a basis of T'.

$$T = basis(T') \tag{2.68}$$

Note that the redundant vectors in T' will correspond to vectors in S which are orthogonal to the constraint surface and do not belong to the solution space. In general, as the number of orthogonal constraints on the system increases the dimension of the solution space is reduced, leading to a reduced set of perturbation vectors.



non projected perturbation sequence, $S = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$

projection operator matrix, $\mathbf{P} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$

projected perturbation vectors, $T' = \mathbf{P}S = \left\{ \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \right\}$

reduced perturbation vectors, $T = \left\{ \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right\}$

Figure 7. Weight Update

2.10.2 Approaches to Gradient Estimation Using Projected and Non Projected Perturbation Sequences

In this section we identify the different approaches possible for extracting the required gradient when using projected and non-projected perturbation sequences. The gradient estimation approaches differ in the type of perturbation sequences used to perturb the array weights and the type of perturbation sequences used to correlate the instantaneous output power of the single receiver system (or the instantaneous power difference sequence of the dual receiver systems).

We refer to the different gradient estimation approaches as follows:

Non Projected Perturbation Approach- the array weights are perturbed with a non projected perturbation sequence $\delta(.)$, and the instantaneous output power of the single receiver system (or the instantaneous power difference sequence of the dual receiver systems), is correlated with the same non-projected perturbation sequence $\delta(.)$.

Projected Perturbation Approach- the array weights are perturbed with a projected perturbation sequence $\mathbf{P}\delta(.)$, and the instantaneous output power of the single receiver system (or the instantaneous power difference sequence of the dual receiver systems), is correlated with the same projected perturbation sequence $\mathbf{P}\delta(.)$.

Hybrid Perturbation Approach- the array weights are perturbed with a projected perturbation sequence $\mathbf{P}\delta(.)$, and the instantaneous output power of the single receiver system (or the instantaneous power difference sequence of the dual receiver systems), is correlated with the non-projected perturbation sequence $\delta(.)$.

The hybrid perturbation approach can provide computational advantages by reducing the complexity and nature of the multiplications that take place in correlation. In the case of correlation with the Time Multiplex perturbation sequence the number of multiplications is reduced as there is only one non-zero element in each perturbation vector. In comparison, for the projected perturbation approach in the case of correlation with the projected Time Multiplex Sequence all the components of the perturbation vectors may be significant.

In later Chapters we note in the relevant places what the effect is of using the hybrid perturbation approach as compared to the projected perturbation approach.

A second hybrid perturbation approach exists whereby the array weights are perturbed with a non-projected perturbation sequence $\delta(.)$, and the instantaneous output power of the single receiver system (or the instantaneous power difference sequence of the dual receiver systems), is correlated with the projected perturbation sequence $\mathbf{P}\delta(.)$. This second hybrid approach is of little value as it does not result in computational advantages or perturbation noise reduction. We do not explore it in detail in later chapters and have only identified it in this section for completeness.

2.10.3 Properties of Projected Perturbation Sequences

In this section we briefly review the major results, established in [4], [36] and [37], on the use of projected perturbation sequences in narrowband arrays. We review these results to demonstrate that they are limited in characterising an array's performance. Some of these results are also required in later sections of the thesis. Most of the following results are extensions of the results presented in Sections 2.8 and 2.9.

Result 2.10.a. Gradient Estimate using the Projected Perturbation Approach

The conditional mean of the gradient estimate, obtained with any of the perturbation receiver structures identified earlier, using an appropriate projected perturbation sequence T' is unbiased for any γ>0, and satisfies

$$E[\boldsymbol{G}(\boldsymbol{W})|\boldsymbol{W}(n)] = 2\mathbf{PR}\boldsymbol{W}(n) \qquad (2.69)$$

❑ ❑ ❑

Result 2.10.b. Gradient Estimate using the Hybrid Perturbation Approach

The conditional mean of the gradient estimate, obtained with any of the perturbation receiver structures identified earlier, using an appropriate projected perturbation sequence $\mathbf{P}\delta(.)$ to perturb the array weights and the corresponding non-projected perturbation sequence $\delta(.)$ in correlation, is unbiased for any γ>0, and satisfies

$$E[\boldsymbol{G}(\boldsymbol{W})|\boldsymbol{W}(n)] = 2\mathbf{PR}\boldsymbol{W}(n) \qquad (2.70)$$

❑ ❑ ❑

The proof of *Results 2.10.b* follows simply from the proof of *Results 2.10.a* which is contained in [4].

Result 2.10.c. Perturbation Noise Mean Excess Output Power for a Zero Mean Projected Perturbation Sequence

For a projected zero mean perturbation sequence T', the mean excess output power due to perturbations about a nominal weight $\boldsymbol{W}$ is given by

$$\xi = 2\gamma^2 Tr(\mathbf{PRP}) \le 2\gamma^2 Tr(\mathbf{R}) \qquad (2.71)$$

❑❑❑

Result 2.10.d. Look Direction Perturbation Noise Mean Excess Output Power for a Zero Mean Projected Perturbation Sequence

For a zero mean projected perturbation sequence T', the mean excess power from the constrained look direction due to perturbation about a nominal weight $\boldsymbol{W}$ is zero.

$$\zeta = 0 \qquad (2.72)$$

❑❑❑

From *Results 2.10.c* and *2.10.d* it may seem desirable, in any perturbation receiver structure, to use the projected perturbation approach in preference to the non projected approach since less perturbation noise will result. However, from these results no qualitative indication as to how well the beamformer performs with either approach can be made. The conditions under which the projected sequences provide the most benefit in terms of a performance gain are not yet clearly understood. In Chapter 3 we develop qualitative measures to characterise an array's performance.

Result 2.10.e. Necessary Weight Update Projections

If a projected perturbation sequence is used, the projection operation on the weight adjustment term $[\boldsymbol{W}(n) - \mu \boldsymbol{G}(\boldsymbol{W}(n))]$ is not required at each weight update.

❑❑❑

To obtain *Result 2.10.e* it is assumed that the weight generated at the previous weight update is on the constraint surface. This result suggests that compared to the non-projected perturbation approach a significant computational effort can be obtained by eliminating the number of weight projections that take place. When reducing the number of weight projections that take place caution must be exercised since errors may still be propagated. These errors are due to finite precision arithmetic in the processor and quantisation of the perturbed weights. Consequently it would still be necessary to introduce a suitable number of weight projections to eliminate these errors.

When implementing the projected sequences additional implementation issues will result. For example, in the case of quantisation, for the projected time multiplex sequence there may be more than one non zero element in each perturbation vector as compared to the time multiplex sequence. These additional components can make the quantisation process more difficult when the additional components are small but not insignificant. In Chapter 4 we examine these quantisation issues and the level of loss of performance due to quantisation effects.

2.10.4 Excess Noise Output Power in Adaptive Narrowband Beamforming Using Projected Perturbation Sequences

In this section, we examine the components that contribute to the excess noise output power of an adaptive array that uses the projected perturbation approach or the hybrid perturbation approach.

For a single receiver we denote ξ_{Total} as the maximum total excess output power due to noise. It is given by

$$\xi_{Total} = M_{\xi} + \xi \qquad (2.73)$$

where M_{ξ} represents the excess output power due to the misadjustment in the weight update algorithm and ξ is excess output power due to the perturbation of the weights. ξ is given by (2.60). We are only considering an infinite precision system here, the noise effects due to digital implementation such as weight quantisation are considered in Chapter 4.

Now considering a system that has one desired signal and a number of interference signals incident on the array. Using (2.19) and (2.71) in (2.73), ξ_{Total} can be expressed as

$$\xi_{Total} = M_{\xi} + 2\gamma^2 Tr(\mathbf{PR}_{ds}) + 2\gamma^2 Tr(\mathbf{PR}_I) \qquad (2.74)$$

where $\mathbf{R}_{ds}$ is the correlation matrix due to the desired signal and $\mathbf{R}_I$ is the correlation matrix due to all the interference sources and the noise on the array elements.

Examining (2.74) one can observe that:

- Depending on the offset of a signal's direction of arrival from the look direction the interference signals' contribution of perturbation excess noise can be as significant as the desired signal's contribution.
- We can modify ξ_{Total} through the perturbation step size γ and the projection operator $\mathbf{P}$.

The selection of an optimum γ is examined in Chapter 3.

When the direction of arrival of the desired signal corresponds to the look direction, using Result 2.10.d, it can be observed that the excess noise output power contributed by the desired signal will be minimised. That is, the projection operation nulls the desired signal's contribution and effectively removes the desired signal's contribution from the total array correlation matrix. ξ_{Total} is now given by

$$\xi_{Total} = M_{\xi} + 2\gamma^2 Tr(\mathbf{P}\mathbf{R}_I)$$

However, when there is directional mismatch such that the direction of arrival of the desired signal and the look direction are offset, the perturbation noise contributed by the desired signal is not necessarily minimised. An array's sensitivity to perturbation noise contributed by the desired signal can be improved through the use of spatial derivative constraints [9],[13], [18], [28], [29], [35], [42]. Note that by modifying $\mathbf{P}$ to change the perturbation noise contributed by the desired signal will also impact on the perturbation noise contributed by the interference signals. The application of spatial derivative constraints and an arrays sensitivity to perturbation noise is examined in Chapter 5.

Chapter 3

Performance Analysis using Projected Perturbation Sequences

3.1 Introduction

In [2], [5] and [6] Cantoni and Godara have evaluated the performance of an adaptive antenna array which uses orthogonal perturbation sequences. For an array using the constrained LMS algorithm they derived expressions for the gradient estimate, the gradient covariance, and the misadjustment. They considered three different perturbation receiver schemes.

Webster, Evans and Cantoni proposed the use of projected perturbation sequences with adaptive antenna arrays in [4]. The projected perturbation sequences allow simultaneous adaptation and reception by using weight perturbations that always satisfy the look direction constraint. In [4], the performance analysis of an array using these projected sequences was limited to considering computational aspects of the projected perturbation scheme and deriving expressions for the gradient estimate and the excess mean output power due to perturbations. In their analysis only the single receiver perturbation scheme was considered. Their major results are reviewed in the previous Chapter.

A major contribution of this thesis is that we extend the characterisation of the projected perturbation approach by analysing the transient performance and determining the misadjustment for the three receiver structures described earlier. We derive expressions for the gradient covariance. The misadjustment analysis is based on the *Direct* and *Bounds* approach developed in [5] and [6]. In [6] the Direct approach is referred to as the *Exact* approach. The results presented here have been published by the author in [47] and [48].

Also, a new, more straightforward, misadjustment analysis technique is introduced. The new technique is based on the *Bounds* approach and solves the weight covariance matrix.

This Chapter is organised as follows. In Section 3.2 we derive the gradient covariance for the different projected perturbation schemes. The gradient covariance is required to determine the weight covariance matrix and the excess mean square output power.

In Section 3.3 we investigate the perturbation step size selection for the single receiver system. The selection of a suitable perturbation step size allows the simplification of expressions in later analysis.

We present the *Direct* analysis in Section 3.4. The transient behaviour of the weight covariance matrix is analysed first, under the condition that the weight covariance matrix and the **PRP** matrix can be diagonalised by the same unitary transformation. The *Direct* misadjustment analysis is then performed. The misadjustment is expressed in terms of the weight covariance matrix and the analysis is performed for all receiver structures.

In Section 3.5 we consider an alternative bounding technique, which we refer to as the *Bounds* approach, for obtaining the misadjustment. This misadjustment analysis examines the excess mean square output power and it requires the norm of the weight error vector to converge. Bounds on the excess mean square output power and hence the misadjustment are established and the asymptotic misadjustment bounds are evaluated.

A new bounding technique for establishing the misadjustment is presented in Section 3.6. The new bounds technique establishes bounds on the weight covariance matrix.

In Section 3.7 we compare the misadjustment expressions obtained with the different misadjustment techniques and finally present results of simulation studies in Section 3.8.

In this Chapter, for the sake of conciseness, only the basic approach and an outline of major intermediate results are presented. A detailed derivation of the intermediate results can be found in the appropriate Appendices as referenced in the following text.

3.2 Gradient Covariance Results

The covariance of the gradient estimate plays a key role in the analysis of the misadjustment and transient behaviour of the array weights. It can be used to determine the excess mean square power and the weight covariance matrix. In this section expressions for the covariance of the gradient estimate for the single and dual receiver systems using the projected perturbation approach and the hybrid perturbation approach are developed.

In the first instance results are stated for the general classes of sequences, then results are derived for the projected Time Multiplex sequence. The covariance expressions for the projected perturbation approach, the non projected perturbation approach and the hybrid perturbation approach will then be compared.

3.2.1 Gradient Estimation Generic Results for the Projected Perturbation Approach

To determine the conditional covariance of the gradient estimate, we make the following assumptions:

Assumption 3.1: **X**(.) is an independent and identically distributed, zero mean, complex gaussian process

Assumption 3.2: For each perturbation receiver structure, the appropriate projected orthogonal perturbation sequences is used such that an unbiased estimate of the gradient is obtained

Assumption 3.1 is a commonly used assumption, [9], [5], [6], that is required to make the analysis tractable.

In Appendix *B*, using the approach described in [2], [5] the following ***Generic Expressions*** for the gradient covariance were developed. We refer to the expressions as being generic since they are applicable to any projected orthogonal perturbation sequence.

Result 3.2.a. Gradient Covariance of Dual Receiver Dual Perturbation System

$$\mathbf{V}_{G1}(W(n)) = \frac{1}{m^2}\sum_{i=1}^{m}[(W^H\mathbf{R}\delta_p(i))^2 + (\delta_p^H(i)\mathbf{R}W)^2 + 2(W^H\mathbf{R}W)(\delta_p^H(i)\mathbf{R}\delta_p(i))]\delta_p(i)\delta_p^H(i) \quad (3.1)$$

❑❑❑

Result 3.2.b. Gradient Covariance of Dual Receiver Reference Receiver System

$$\mathbf{V}_{G2}(W(n)) = \gamma^2\mathbf{D} + \mathbf{E} \quad (3.2)$$

where

$$\mathbf{D} = \frac{1}{m^2}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))^2\delta_p(i)\delta_p^H(i) \quad (3.3)$$

and

$$\mathbf{E} = \frac{2W^H(n)\mathbf{R}W(n)}{m^2}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))\delta_p(i)\delta_p^H(i) + \frac{1}{m^2}\sum_{i=1}^{m}[(W^H(n)\mathbf{R}\delta_p(i))^2 + (\delta_p^H(i)\mathbf{R}W(n))^2]\delta_p(i)\delta_p^H(i) \quad (3.4)$$

❑❑❑

Result 3.2.c. Gradient Covariance of Single Receiver System

$$\mathbf{V}_{G3}(W(n)) = \gamma^2\mathbf{A} + \frac{1}{\gamma^2}\mathbf{B} + \mathbf{C} \quad (3.5)$$

where

$$\mathbf{A} = \frac{1}{m^2}\sum_{i=1}^{m}[\delta_p^H(i)\mathbf{R}\delta_p(i)]^2\delta_p(i)\delta_p^H(i)$$

$$\mathbf{B} = \frac{(W^H(n)\mathbf{R}W(n))^2}{m^2}\sum_{i=1}^{m}\delta_p(i)\delta_p^H(i)$$

$$\mathbf{C} = \frac{1}{m^2}\sum_{i=1}^{m}[W^H(n)\mathbf{R}\delta_p(i) + \delta_p^H(i)\mathbf{R}W(n)]^2\delta_p(i)\delta_p^H(i) + \frac{2W^H(n)\mathbf{R}W(n)}{m^2}\sum_{i=1}^{m}\delta_p^H(i)\mathbf{R}\delta_p(i)\delta_p(i)\delta_p^H(i) \quad (3.6)$$

❑❑❑

In the above equations $\delta_p(i)$ represents the i^{th} vector in the projected perturbation sequence, it is given by

$$\delta_p(i) = \mathbf{P}\delta(i) \quad i = 1, 2, \ldots m \quad (3.7)$$

3.2.2 Gradient Estimation using the Projected Perturbation Approach with the Projected Time Multiplex Sequence

For the special case of the projected Time Multiplex perturbation sequence, in Appendix *B*, (3.1), (3.2) and (3.5) have been derived and are shown below. For simplicity the analysis assumes that the projected sequence defined by (2.67) is used. For the dual receiver dual perturbation system the $2L$ length sequence corresponds to the minimum length Time Multiplex sequence and the $4L$ length sequence corresponds to an odd symmetry Time Multiplex sequence. Expressions for the $4L$ length sequence have been derived to enable comparisons with the other receiver structures.

Result 3.2.d. Gradient Covariance of Dual Receiver Dual Perturbation System with Projected Time Multiplex Sequence

$$\mathbf{V}_{G1}(W(n)) = 2W^H(n)\mathbf{R}W(n)\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} \text{ for a } 4L \text{ length sequence.} \quad (3.8)$$

$$\mathbf{V}_{G1}(W(n)) = 4W^H(n)\mathbf{R}W(n)\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} \text{ for a } 2L \text{ length sequence.} \quad (3.9)$$

❑❑❑

Result 3.2.e. Gradient Covariance of Dual Receiver Reference Receiver System with Projected Time Multiplex Sequence

$$\mathbf{V}_{G2}(W(n)) = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P} + 2W^H(n)\mathbf{R}W(n)\mathbf{P}(Diag(\mathbf{PRP}))\mathbf{P}$$

for a $4L$ length sequence (3.10)

❑❑❑

Result 3.2.f. Gradient Covariance of Single Receiver System with Projected Time Multiplex Sequence

$$\mathbf{V}_{G3}(W(n)) = \mathbf{P}\Big(\gamma^2 2L(Diag(\mathbf{PRP}))^2 + \frac{(W^H(n)\mathbf{R}W(n))^2}{\gamma^2 2L}$$
$$+ 2Diag(\mathbf{PR}W(n)W^H(n)\mathbf{RP}) + 2W^H(n)\mathbf{R}W(n)Diag(\mathbf{PRP})\Big)\mathbf{P}$$

for a $4L$ length sequence (3.11)

❑❑❑

The equivalent gradient covariance expressions for the non projected perturbation approach for the three receiver structures can be found in Appendix *F*.

Examining (3.8), (3.9), (3.10) and (3.11) the following observations can be made:

- For all perturbation receiver structures the covariance of the gradient estimates are positive definite. Comparing these expressions with the equivalent expressions in *Appendix F* for non projected perturbation approach one can observe that in general the variance of the gradient estimate for the projected cases is smaller due to the terms involving **PRP** removing the contribution of the desired signal. This difference in the variance of the gradient estimates depends on the look direction signal power.

- The covariance is directly proportional to the output power, $W^H(n)\mathbf{R}W(n)$. This dependence decreases in time since the adaptive algorithm is attempting to minimise this quantity. Thus as the gradient estimate improves as the weight vector approaches its optimum value, the covariance becomes independent of the desired signal.

- The covariance expression for the dual receiver dual perturbation system is independent of the perturbation step size. For the dual receiver reference receiver system the covariance is a monotonically increasing function of the perturbation step size and the single receiver system is a convex function of γ^2. Similar relations also apply to the non-projected perturbation approach [2],[5].

- The terms in the covariance expression for the dual receiver dual perturbation system with perturbation sequence length $4L$ are common to the dual receiver reference receiver and the single receiver covariance expressions. Also the terms in the covariance expression for the dual receiver reference receiver system are common to the single receiver covariance expression.

- From (3.8) and (3.9) the covariance of the dual receiver dual perturbation system with perturbation sequence length $2L$ is double that of the dual receiver dual perturbation system with perturbation sequence length $4L$. The latter system has a smaller covariance since it has a longer time to make a better gradient estimate.

3.2.3 Gradient Estimation Using the Hybrid Perturbation Approach

When estimating the gradient covariance for the hybrid perturbation approach *Assumption 3.1* and *Assumption 3.2* are still required. It is shown in Appendix *B* that the ***Generic*** expression for the gradient covariance for the dual receiver dual perturbation system is given by

Result 3.2.g. Gradient Covariance of Dual Receiver Dual Perturbation System using the Hybrid Perturbation Approach

$$\mathbf{V}_{G1}(W(n)) = \frac{1}{m^2}\sum_{i=1}^{m}[(\mathbf{W}^H\mathbf{R}\delta_p(i))^2 + (\delta_p^H(i)\mathbf{R}\mathbf{W})^2 + 2(\mathbf{W}^H\mathbf{R}\mathbf{W})(\delta_p^H(i)\mathbf{R}\delta_p(i))]\delta(i)\delta^H(i) \qquad (3.12)$$

❑❑❑

Comparing (3.1) with (3.12) the following observation can be made:

- The two expressions are similar except for the last matrix term defined by the product of the perturbation vectors. By substituting $\delta_p(i) = \mathbf{P}\delta(i)$ in the perturbation vector product in (3.1), we observe that the covariance expression defined by (3.1) is equivalent to pre- and post- multiplying the gradient covariance expression defined by (3.12) by the projection matrix **P**. A similar relationship exists for the gradient covariance expressions for the other receiver structures. This is discussed in Appendix B.

Generic expressions for the covariance of the gradient estimate of the single receiver and the dual receiver reference receiver system using the hybrid perturbation approach can be found in Appendix *B*.

3.3 Optimal Perturbation Step Size for the Projected Perturbation Approach

In this section we determine the optimal perturbation step size $\hat{\gamma}$, which minimises the gradient covariance for the single receiver system.

For the single receiver case, allowing the perturbation step size to be a scaled multiple of $\hat{\gamma}$ simplifies the gradient covariance expression. This is useful in later analysis.

While we use a similar approach found in [5], [39] to determine $\hat{\gamma}$, an additional assumption is required in the estimation of $\hat{\gamma}$, and a new method is needed to implement the terms in $\hat{\gamma}$.

Note that in this section and in the rest of the Chapter we consider the single linear constraint adaptive beamforming problem.

3.3.1 Estimation of Optimum γ

As shown in [5], [39] it can be observed that for the single receiver system the covariance is a convex function of γ^2 and an optimal value of γ^2 exists for which the covariance is a minimum. This is also true with the projected perturbation approach and the hybrid perturbation approach. For the dual receiver systems no optimum perturbation step size exists to minimise the covariance of the gradient estimate as discussed previously. For the single receiver system, as with the dual receiver systems, the choice of perturbation step size is also constrained by the maximum perturbation noise which is acceptable in the system and by the wordlength used in the implementation.

The following assumption has been found necessary to estimate the optimum perturbation step size.

Assumption 3.3: $Diag(\mathbf{PRP}) \approx \frac{1}{L}Tr(\mathbf{PRP})\mathbf{I}_{LL}$ (3.13)

Assumption 3.3 can be justified as follows. Using the definition of the projection matrix and considering only a single linear constraint system

$$
\begin{aligned}
Diag(\mathbf{PRP}) &= Diag\left(\left(\mathbf{I}_{LL} - \frac{\mathbf{CC}^H}{\mathbf{C}^H\mathbf{C}}\right)\mathbf{R}_N\left(\mathbf{I}_{LL} - \frac{\mathbf{CC}^H}{\mathbf{C}^H\mathbf{C}}\right)\right) \\
&= Diag\left(\mathbf{R}_N - 2Re\left(\frac{\mathbf{R}_N\mathbf{CC}^H}{L}\right) + \frac{\mathbf{CC}^H\mathbf{R}_N\mathbf{CC}^H}{L^2}\right) \qquad (3.14)
\end{aligned}
$$

where $\mathbf{R}_N$ is the interference and noise correlation matrix.
In (3.14), the first and third term's diagonal components are all equal by virtue of their definition. The second term's diagonal components are not necessarily equal, they are approximately equal when the direction of arrival of the desired and interference signal are close. However, for L large the second term's components will be small. If L is not large, the assumption is still reasonable provided that the direction of arrival of the desired and interference signals are spaced such that they have steering vectors which are approximately orthogonal.

It is shown in In Appendix *B* that when the projected Time Multiplex sequence is used the following results can be derived.

Result 3.3.a Optimal Perturbation Step Size

If *Assumption 3.3* is satisfied the optimal perturbation step size for the single receiver system is given by

$$\hat{\gamma}(W(n)) = \left[\frac{\mathbf{W}^H(n)\mathbf{R}W(n)}{2Tr(\mathbf{PRP})}\right]^{1/2} \tag{3.15}$$

❑❑❑

Result 3.3.b Single Receiver System optimum gradient covariance

Let $\hat{\mathbf{V}}_{G3}(W(n))$ represent the value of $\mathbf{V}_{G3}(W(n))$ at $\gamma = \hat{\gamma}(W(n))$ then

$$\hat{\mathbf{V}}_{G3}(W(n)) = 2\mathbf{P}[Diag(\mathbf{PR}W(n)W^H(n)\mathbf{RP}) + 4W^H(n)\mathbf{R}W(n)Diag(\mathbf{PRP})]\mathbf{P} \tag{3.16}$$

❑❑❑

Note that *Result 3.3.a* and *b* is also true for the Hybrid perturbation approach.

To enhance the signal to noise ratio for a single receiver system the perturbation noise can be reduced by scaling the time varying perturbation step size in (3.15) according to

$$\gamma(W(n)) = c\left[\frac{\mathbf{W}^H(n)\mathbf{R}W(n)}{2Tr(\mathbf{PRP})}\right]^{1/2} \tag{3.17}$$

where c satisfies $0 \ll c < 1$.

The gradient covariance is then given by

$$\hat{\mathbf{V}}_{G3}(W(n)) = 2\mathbf{P}[Diag(\mathbf{PR}W(n)W^H(n)\mathbf{RP}) + aW^H(n)\mathbf{R}W(n)Diag(\mathbf{PRP})]\mathbf{P} \tag{3.18}$$

where $a = \left(c + \frac{1}{c}\right)^2$ (3.19)

3.3.2 Implementation of Optimum γ

From the previous expression for $\hat{\gamma}(W(n))$ it is clear that to evaluate $\hat{\gamma}(W(n))$ a knowledge of the covariance matrix $\mathbf{R}$ is required. Two approaches have been proposed to select a $\hat{\gamma}$. The first is to use a constant γ which is close to the optimal value and the second is to select a time varying γ.

A possible constant value that can be used is

$$\gamma = c\left[\frac{\hat{W}^H \mathbf{R} \hat{W}}{2Tr(\mathbf{PRP})}\right]^{1/2} \tag{3.20}$$

This requires a knowledge of $\hat{W}^H \mathbf{R} \hat{W}$ and $Tr(\mathbf{PRP})$ both of which are related to the input signal scenario of the processor.

A time varying γ that can be selected is

$$\gamma(W(n)) = c\left[\frac{W^H(n)\mathbf{R}W(n)}{2Tr(\mathbf{PRP})}\right]^{1/2} \tag{3.21}$$

This eliminates the need to know $\hat{W}$ but still requires a knowledge of $\mathbf{R}$ and $Tr(\mathbf{PRP})$.

The quantity $W^H(n)\mathbf{R}W(n)$ can be approximated by either

1) Using the instantaneous output power.

2) Using the average power over a perturbation cycle.

Estimating the diagonal elements of **PRP** is slightly more difficult. It is proposed that when *Assumption 3.3* holds then the diagonal elements of **PRP** will be similar to the diagonal elements of $\mathbf{R_N}$. Since the diagonal elements of $\mathbf{R_N}$ are all equal we obtain

$$Tr(\mathbf{PRP}) = Lr$$

where r is the diagonal element of $\mathbf{R_N}$ and can be found by suitably averaging the quantity $x^H{}_j(n)x_j(n)$.

3.4 Direct Analysis

In this section, we characterise the performance of the projected perturbation schemes by analysing the transient behaviour of the weight covariance matrix and determining the misadjustment. Although we use the same analysis approach as that developed in [6], the analysis is significantly different since the terms that occur in the weight covariance expression, which are different for different perturbation sequences or gradient estimation schemes, require special treatment.

In [6] the *Direct* misadjustment analysis is not applied to the single receiver system. Here we also demonstrate, with some modification, the application of the *Direct* misadjustment analysis to the single receiver system for the projected perturbation approach. The modified *Direct* misadjustment analysis can also be applied to the single receiver system with the non-projected perturbation approach.

The transient analysis could only be performed for the dual receiver systems. This is so since to derive a complete and closed form description of the transient behaviour it is necessary to simultaneously diagonalize the weight covariance matrix and **PRP** by the same unitary transformation [6]. This approach is intractable for the single receiver system.

Assumption 3.3 is required for the transient analysis of the dual receiver systems since it enables the simultaneous diagonalization of the weight covariance matrix and **PRP** by the same unitary transformation. *Assumption 3.3* is also required to perform the misadjustment analysis for all receiver structures. In a subsequent section for the dual receiver systems a different analysis approach, the *Bounds* approach, which does not require *Assumption 3.3* in order to obtain bounds on the misadjustment, is presented.

Details of the application of the *Direct* analysis can be found in Appendix *B*.

The expressions for the transient analysis and misadjustment are evaluated for the projected time multiplex sequence.

3.4.1 Transient Behaviour

The transient behaviour and convergence properties of the weight covariance matrix for the dual receiver structures are derived in this section. The convergence properties of the weight covariance matrix are required to determine the misadjustment.

The weight covariance matrix is defined as

$$\mathbf{K}_{WW}(n) = E[(W(n) - \overline{W}(n))(W(n) - \overline{W}(n))^H] \tag{3.22}$$

where

$$\overline{W}(n) = E[W(n)] \tag{3.23}$$

As defined in [6],[10] $\mathbf{K}_{WW}(n)$ satisfies a recursive equation as stated in the following result.

Result 3.4.a. Weight Covariance

If $\mathbf{V}_G(W(n))$ denotes the covariance of the gradient used in (2.27) for a given $W(n)$ and $\mathbf{K}_{WW}(n)$ denotes the covariance of $W(n)$ then

$$\mathbf{K}_{WW}(n+1) = \mathbf{P}\mathbf{K}_{WW}(n)\mathbf{P} - 2\mu\mathbf{P}[\mathbf{R}\mathbf{K}_{WW}(n) + \mathbf{K}_{WW}(n)\mathbf{R}]\mathbf{P} + 4\mu^2\mathbf{P}\mathbf{R}\mathbf{K}_{WW}(n)\mathbf{R}\mathbf{P} + \mu^2\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P} \quad (3.24)$$

❑❑❑

where the expectation is taken over W. Derivation of (3.24) is presented in [6].

From (3.24) it can be observed that the gradient covariance is pre- and post-multiplied by the projection matrix **P**. Due to this, for each perturbation receiver structure, the weight covariance matrix is identical for the cases when the gradient estimate is derived using the projected perturbation approach and the hybrid perturbation approach.

3.4.1.1 Diagonalization of the Weight Covariance Matrix

In this section, the condition under which it is possible to diagonalize the weight covariance matrix is examined and where the condition is met the solution is presented.

The transient behaviour of the weight covariance matrix will be studied under the condition that $\mathbf{K}_{WW}(n)$ and **PRP** can be diagonalized by the same unitary transformation. The condition under which the diagonalization of the weight covariance matrix and **PRP** by the same unitary transformation can take place may be stated as, [6]

Result 3.4.b. Diagonalisation of the Weight Covariance Matrix

The necessary and sufficient condition for the diagonalization of $\mathbf{K}_{WW}(n+1), n \geq 0$ and **PRP** by the same unitary transformation is that the unitary transformation also diagonalizes $\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P}$ for all n.

❑❑❑

The gradient covariance expressions presented in Section 3.2 are now examined in the light of the above condition.

Taking expectation over $W(n)$, pre- and post- multiplying by $\mathbf{P}$ on both sides of equations (3.8), (3.10), (3.18) and using *Assumption 3.3.*

$$\mathbf{P}E[\mathbf{V}_{G1}(W(n))]\mathbf{P} = \left(\frac{2}{L}\right)(Tr(\mathbf{PRP}))Tr(\mathbf{RR}_{WW}(n))\mathbf{P} \tag{3.25}$$

$$\mathbf{P}E[\mathbf{V}_{G2}(W(n))]\mathbf{P} = \mathbf{P}\left[\frac{2\gamma^2}{L}(Tr(\mathbf{PRP}))^2 + \frac{2}{L}Tr(\mathbf{RR}_{WW}(n))Tr(\mathbf{PRP})\right]\mathbf{P} \tag{3.26}$$

$$\mathbf{P}E[\hat{\mathbf{V}}_{G3}(W(n))]\mathbf{P} = 2\mathbf{P}Diag(\mathbf{PRR}_{WW}(n)\mathbf{RP})\mathbf{P} + \frac{a}{L}Tr(\mathbf{RR}_{WW}(n))\mathbf{P} \tag{3.27}$$

$$\text{where } \mathbf{R}_{WW}(n) = E[W(n)W^H(n)] \tag{3.28}$$

$\mathbf{P}$ can be diagonalized by the same unitary transformation as **PRP** since **PRP** is Hermitian and $\mathbf{P}$ commutes with **PRP**. In the single receiver case, even though the first matrix in (3.27) is Hermitian, it does not commute with **PRP**. Hence from Theorem A.3, the weight covariance matrix in this instance cannot be diagonalized by the same unitary transformation. The transient analysis for the single receiver system is thus not possible with this approach.

Let $\mathbf{Q}$ be the unitary transformation. Then

$$\mathbf{Q}^H\mathbf{PRPQ} = \Lambda$$

$$\mathbf{Q}^H\mathbf{PQ} = \Gamma \tag{3.29}$$

where Λ and Γ are diagonal matrices with the diagonal elements being the eigenvalues of **PRP** and $\mathbf{P}$ respectively. Pre- and post- multiplying (3.25), (3.26) by $\mathbf{Q}^H$ and $\mathbf{Q}$, and applying (3.29) gives

$$\mathbf{Q}^H\mathbf{P}E[\mathbf{V}_{G1}(W(n))]\mathbf{PQ} = \left(\frac{2}{L}\right)Tr(\mathbf{PRP})Tr(\mathbf{RR}_{WW}(n))\Gamma \tag{3.30}$$

$$\mathbf{Q}^H\mathbf{P}E[\mathbf{V}_{G2}(W(n))]\mathbf{PQ} = \left[\frac{2\gamma^2}{L}(Tr(\mathbf{PRP}))^2 + \frac{2}{L}Tr(\mathbf{RR}_{WW}(n))Tr(\mathbf{PRP})\right]\Gamma \tag{3.31}$$

These equations imply that (3.8) and (3.10) satisfy the conditions specified in *Result 3.4.b.* Thus for the dual receiver systems the weight covariance matrix can be represented by

$$\Sigma(n) = \mathbf{Q}^H\mathbf{K}_{WW}(n)\mathbf{Q} \tag{3.32}$$

where the variance of the gradient estimate is given by (3.8) or (3.10).

Similar to the weight covariance matrix a recursive equation for the diagonalized weight covariance matrix can be established. Pre- and post- multiplying (3.24) by $\mathbf{Q}^H$ and $\mathbf{Q}$, and using (3.32) gives

$$\begin{aligned}\Sigma(n+1) &= \Sigma(n) - 2\mu\Lambda\Sigma(n) - 2\mu\Sigma(n)\Lambda + 4\mu^2\Lambda\Sigma(n)\Lambda \\ &\qquad + \mu^2\mathbf{Q}^H\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{PQ} \\ &= (\mathbf{I}_{LL} - 4\mu\Lambda + 4\mu^2\Lambda^2)\Sigma(n) + \mu^2\mathbf{Q}^H\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{PQ}\end{aligned} \tag{3.33}$$

where the following substitution has been used

$$\mathbf{K}_{WW}(n) = \mathbf{R}_{WW}(n) - \overline{W}(n)\overline{W}^H(n) \tag{3.34}$$

It is shown in Appendix *B* that for the dual receiver systems a solution for the weight covariance equation can be established. This is done by determining the vector difference equation for the diagonal elements of the diagonalized weight covariance matrix (3.32) and solving it. The solutions are shown next.

Result 3.4.c. Weight Covariance of Dual Receiver Dual Perturbation System

$$\mathbf{K}_{WW}(n) = \sum_{l=1}^{L} \boldsymbol{n}_1(n)Q_lQ_l^H \tag{3.35}$$

where the solution of $\boldsymbol{n}_1(n)$ is given by

$$\boldsymbol{n}_1(n) = (\mathbf{I}_{LL} - \boldsymbol{H}_1)^n\boldsymbol{n}_1(0) + \mu^2\frac{2}{L}Tr(\mathbf{PRP})\sum_{i=1}^{L} k_o(n-i)(\mathbf{I}_{LL} - \boldsymbol{H}_1)^{i-1}\delta_1 \tag{3.36}$$

and $\qquad \boldsymbol{H}_1 = 4\mu\Lambda - 4\mu^2\Lambda^2 - \frac{2}{L}Tr(\mathbf{PRP})\mu^2(\delta_1\lambda^T)$

$k_o(n) = \overline{W}^H(n)\mathbf{R}\overline{W}(n)$

Q_l l=1,2,...L are the eigenvectors of **PRP**

δ_1 denotes the $L \times 1$ dimensional vector of eigenvalues of **P**

$\boldsymbol{n}_1(n)$ denotes the $L \times 1$ dimensional vector of eigenvalues of $\mathbf{K}_{WW}(n)$

λ denotes the $L \times 1$ dimensional vector of eigenvalues of **PRP**

❑❑❑

In all cases considered here $\lim_{n \to \infty} \overline{W}(n) = \hat{W}$ under suitable conditions.Thus $\lim_{n \to \infty} k_o(n) = \hat{W}^H\mathbf{R}\hat{W}$.

Result 3.4.d. Weight Covariance of Dual Receiver Reference Receiver System

$$\mathbf{K}_{WW}(n) = \sum_{l=1}^{L} \boldsymbol{n}_2(n) Q_l Q_l^H \tag{3.37}$$

where $\boldsymbol{n}_2(n)$ denotes the $L \times 1$ dimensional vector of eigenvalues of $\mathbf{K}_{WW}(n)$ and $\boldsymbol{Q}_l$ l=1,2,...L are the eigenvectors of **PRP**.

The solution of $\boldsymbol{n}_2(n)$ is given by

$$\boldsymbol{n}_2(n) = (\mathbf{I}_{LL} - \boldsymbol{H}_1)^n \boldsymbol{n}_2(0) + \mu^2 \frac{2}{L} Tr(\mathbf{PRP}) \sum_{i=1}^{L} (k_o(n-i) + \gamma^2 Tr(\mathbf{PRP}))(\mathbf{I}_{LL} - \boldsymbol{H}_1)^{i-1} \delta_1 \tag{3.38}$$

❑❑❑

3.4.1.2 Convergence of the Weight Covariance

In this section, we complete the characterisation by considering the convergence of the weight covariance. The convergence of the weight covariance is required in the following section to derive an expression for the misadjustment.

It is shown in Appendix *B* that equations $\boldsymbol{n}_1(n+1)$ and $\boldsymbol{n}_2(n+1)$ can be reduced to a set of $(L-1)$ difference equations since one of the components in each of the vectors is zero. This is due to the rank deficiency of **P**. For the single look direction constraint as **P** has a zero eigenvalue, **PRP** and $\mathbf{K}_{WW}(n)$ each have a zero eigenvalue.

In Appendix *B* proofs are presented that show the $\lim_{n \to \infty} \boldsymbol{n}'_1(n)$ and $\lim_{n \to \infty} \boldsymbol{n}'_2(n)$ exist, under the conditions specified there. Assuming that these conditions hold it follows that:

For the Dual Receiver Dual Perturbation System with a 4L Length perturbation sequence

$$\lim_{n \to \infty} \boldsymbol{n}'_1(n) = \frac{\frac{\mu}{2L} Tr(\mathbf{PRP}) \hat{W}^H \mathbf{R} \hat{W}}{1 - \frac{\mu}{2L} Tr(\mathbf{PRP}) \sum_{i=1}^{L-1} \frac{1}{1-\lambda_i}} \begin{bmatrix} \frac{1}{\lambda_1(1-\mu\lambda_1)} \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{\lambda_{L-1}(1-\mu\lambda_{L-1})} \end{bmatrix} \tag{3.39}$$

and *for the Dual Receiver Reference Receiver System with a 4L Length perturbation sequence*

$$\lim_{n \to \infty} n'_2(n) = \frac{[\hat{W}^H \mathbf{R} \hat{W} + \gamma^2 Tr(\mathbf{PRP})] \frac{\mu Tr(\mathbf{PRP})}{L}}{1 - \frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1 - \mu\lambda_i}} \begin{bmatrix} \frac{1}{\lambda_1(1-\mu\lambda_1)} \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{\lambda_{L-1}(1-\mu\lambda_{L-1})} \end{bmatrix} \tag{3.40}$$

The steady state expressions for the weight covariance matrices can now be established by substituting (3.39) and (3.40) into (3.35) and (3.37) respectively

Result 3.4.e. Steady State Weight Covariance of the Dual Receiver Dual Perturbation System with a 4L length perturbation sequence

$$\lim_{n \to \infty} \mathbf{K}_{WW1}(n) = \frac{\frac{\mu}{2L} Tr(\mathbf{PRP}) \hat{W}^H \mathbf{R} \hat{W} \sum_{i=1}^{L-1} \frac{1}{\lambda_i(1-\mu\lambda_i)} Q_i Q_i^H}{1 - \frac{\mu}{2L} Tr(\mathbf{PRP}) \sum_{i=1}^{L-1} \frac{1}{1-\lambda_i}} \tag{3.41}$$

❑❑❑

Result 3.4.f. Steady State Weight Covariance of the Dual Receiver Reference Receiver System with a 4L length perturbation sequence

$$\lim_{n \to \infty} \mathbf{K}_{WW2}(n) = \frac{[\hat{W}^H \mathbf{R} \hat{W} + \gamma^2 Tr(\mathbf{PRP})] \frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{\lambda_i(1-\mu\lambda_i)} Q_i Q_i^H}{1 - \frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i}} \tag{3.42}$$

❑❑❑

3.4.2 Misadjustment

In this section, we briefly review the *Direct* misadjustment analysis and summarise the results. Unlike the previous transient analysis, the *Direct* misadjustment analysis can be applied to all the receiver structures under study and is presented for the LMS algorithm with a single linear constraint.

As shown in [6] the misadjustment can be expressed as

$$M = \lim_{n \to \infty} \frac{Tr[\mathbf{K}_{WW}(n)\mathbf{R}]}{\hat{\mathbf{W}}^H \mathbf{R} \hat{\mathbf{W}}} \tag{3.43}$$

Note that for each receiver structure the resulting misadjustment is identical when the projected perturbation approach or the hybrid perturbation approach is used. This occurs because for each receiver structure the weight covariance matrix is identical for the two cases.

3.4.2.1 Misadjustment for the Dual Receiver Structures

By using the properties of the weight covariance matrix as stated in Appendix *B* it can be shown that the misadjustment can be represented in vector notation as

$$M = \frac{\lim_{n \to \infty} \lambda^T \boldsymbol{d}(n)}{\hat{W}^H \mathbf{R} \hat{W}} \tag{3.44}$$

where λ represents the $(L-1) \times 1$ dimensional vector of eigenvalues of **PRP** and the $(L-1) \times 1$ dimensional vector $\boldsymbol{d}(n)$ is defined by

$$\boldsymbol{d}(n) = Diag[Q^H \mathbf{K}_{WW}(n) Q] \tag{3.45}$$

In (3.45) $\mathbf{Q}$ is the unitary transformation.
Note that $\lambda^T \boldsymbol{d}(n)$ is the excess power in the output power due to fluctuations in the weight vector.

In Appendix *B*, for the dual receiver systems a vector difference equation for $\lambda^T \boldsymbol{d}(n)$ is established and expressed in a similar form to Theorem A.1. By applying Theorem A.1 the convergence analysis of $\lambda^T \boldsymbol{d}(n)$ could then be made and the steady state expression for $\lambda^T \boldsymbol{d}(n)$ and hence the misadjustment derived. Theorem A.1 is proved in [6] and located in Appendix A.
The main results of the misadjustment analysis for the dual receiver structures are summarised below:

Result 3.4.g. Misadjustment of the Dual Receiver Dual Perturbation System - Direct Approach

When a 4*L* length projected time multiplex sequence is used such that the gradient covariance is given by (3.8), when

$$0 < \mu < \frac{1}{\lambda_{max} + \frac{Tr(\mathbf{PRP})}{2L}} \quad \text{and} \quad \frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1 - \mu \lambda_i} < 1 \tag{3.46}$$

$\lim_{n \to \infty} \lambda^T \boldsymbol{d}(n)$ converges and the corresponding misadjustment is given by

$$M = \frac{\frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1 - \mu \lambda_i}}{1 - \frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1 - \mu \lambda_i}} \tag{3.47}$$

❑❑❑

Note that a similar transient and misadjustment analysis can be carried out for the dual receiver dual perturbation system using the minimum length perturbation sequence. In most instances due to the covariance expression for the two perturbation sequences being related by a scalar, results can be easily derived. Below, only the final misadjustment expression is shown.

When a a $2L$ length projected time multiplex sequence is used such that the gradient covariance is given by (3.9), when the conditions for convergence are satisfied, the misadjustment is given by

$$M = \frac{\frac{\mu Tr(\mathbf{PRP})}{L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}}{1-\frac{\mu Tr(\mathbf{PRP})}{L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}} \tag{3.48}$$

Result 3.4.h. Misadjustment of the Dual Receiver Reference Receiver System - Direct Approach

When a a $4L$ length odd symmetry projected time multiplex sequence is used such that the gradient covariance is given by (3.10), when

$$0<\mu<\frac{1}{\lambda_{max}+\frac{Tr(\mathbf{PRP})}{2L}} \text{ and } \frac{\mu Tr(\mathbf{PRP})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}<1 \tag{3.49}$$

$\lim_{n\to\infty}\sum_{i=1}^{L}\lambda^T d(n)$ exists and the corresponding misadjustment is given by

$$M = \frac{\left[1+\frac{\gamma^2 Tr(\mathbf{PRP})}{\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}}\right]\frac{\mu Tr(\mathbf{PRP})}{L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}}{1-\frac{\mu Tr(\mathbf{PRP})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}} \tag{3.50}$$

❑❑❑

3.4.2.2 Misadjustment for the Single Receiver Structure

The *Direct* misadjustment analysis approach is also possible for the single receiver system and is contained in Appendix *B*. The analysis for the single receiver system had to be handled differently from the dual receiver systems' analysis. Essentially a re-ordering of steps is required whereby it is possible to identify and eliminate obvious transient components in the weight covariance matrix. Eliminating these transient components does not affect the final misadjustment as they have no contribution to the final result. It was possible to identify the transient components by examining terms in the weight covariance matrix which have no contribution in the limit as $n \to \infty$. It is possible to do this as it is initially established that the mean of the weights approach their optimum value, $\lim_{n \to \infty} \overline{W}(n) = \hat{W}$. This modified procedure can also be applied to the dual receivers' analysis.

The main result of the misadjustment analysis for the single receiver structures is summarised below.

Result 3.4.i. Misadjustment of the Single Receiver System - Direct Approach

When a a $4L$ length odd symmetry projected time multiplex sequence is used such that the gradient covariance is given by (3.18), when

$$0 < \mu < \frac{1}{\lambda_{max} + \frac{aTr(\mathbf{PRP})}{4L}} \text{ and } \frac{\mu aTr(\mathbf{PRP})}{4L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i} < 1 \tag{3.51}$$

$\lim_{n \to \infty} \lambda^T d(n)$ converges and the corresponding misadjustment is given by

$$M = \frac{\frac{\mu aTr(\mathbf{PRP})}{4L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i}}{1 - \frac{\mu aTr(\mathbf{PRP})}{4L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i}} \tag{3.52}$$

❑❑❑

3.4.3 Summary

In Section 3.4 we have provided new results for the gradient covariance of all three receiver systems, the transient behaviour of the weight covariance matrix for the dual receiver systems, and new results for the *Direct* misadjustment analysis for all three receiver systems using the projected perturbation approach.

Comparing the covariance of the gradient estimate and misadjustment expressions with the expressions for the non-projected case, which are summarised in Appendix *F*, one observes that in general the weight covariance and the misadjustment for the projected cases are smaller due to the terms involving **PRP** removing the effect of the desired signal contribution. The difference in the performance depending on the look direction signal power. These observations are confirmed in the simulation studies section.

The expressions for misadjustment depend heavily on *Assumption 3.3* being true. Due to this another misadjustment analysis using the *Bounds* approach described in [5] is used to reinforce the results derived here.

3.5 Bounds Analysis

The derivation of the misadjustment expressions presented in the previous analysis rely on *Assumption 3.3*. For the dual receiver systems, an alternative approach for obtaining the misadjustment that does not require this assumption has been investigated. A bounding techniques described in [5],[7] has been used to derive the misadjustment bounds. For the single receiver system *Assumption 3.3* is still required to make the bound analysis tractable.

The convergence characteristics derived with the two techniques differ, but for suitably small step sizes the misadjustment derived using the two approaches are shown to be identical. It is also shown that the misadjustment bounds obtained using the projected perturbation sequence are tighter than those obtained using non-projected perturbation sequences.

The main difference between the two misadjustment analyses is that the *Direct* technique considers the output power due to the fluctuation of the weight vector while the *Bounds* technique examines the excess mean square output power. This difference, $\overline{W}^H(n)\mathbf{R}\overline{W}(n) - \hat{W}^H\mathbf{R}\hat{W}$, converges to zero as $n \to \infty$. Hence both quantities will converge to the same steady state excess mean square power.

3.5.1 Misadjustment

The excess mean square output power at the n^{th} iteration can be defined as

$$emsp(n) = E[(W(n) - \hat{W})^H X(n+1) X^H(n+1)(W(n) - \hat{W})] \tag{3.53}$$

Let $V(n)$ be the weight error vector defined by

$$V(n) = W(n) - \hat{W} \tag{3.54}$$

Substituting (3.54) into (3.53), taking expectation over X and using *Assumption 3.1*, yields the following expression for the steady state excess mean square output power

$$ssemsp = \lim_{n \to \infty} E[V^H(n)\mathbf{R}V(n)] = \lim_{n \to \infty} Tr[E[V(n)V^H(n)]\mathbf{R}] \tag{3.55}$$

It can be shown that when $E[||V(n)||^2]$ converges, bounds on *ssemsp* can be established and hence bounds for the misadjustment [5]. It can also be shown that *ssemsp* is equal to the numerator of the misadjustment expression in (2.53). The bounds for *ssemsp* may be written as

$$\hat{W}^H \mathbf{R} \hat{W} b_l \leq \lim_{n \to \infty} E[V^H(n) \mathbf{R} V(n)] \leq \hat{W}^H \mathbf{R} \hat{W} b_h \tag{3.56}$$

and the corresponding misadjustment is bound by

$$b_l \leq M \leq b_h \tag{3.57}$$

The constants b_l and b_h will be derived later.

Tests for the conditions of convergence of $E[\|V(n)\|^2]$ are detailed in Appendix *C*. It is interesting to note that the recursive equation for the norm of the weight error vector is in a similar form to the recursive equation for the weight covariance matrix, as expected. Also note that the covariance of the gradient estimate is pre- and post-multiplied by the projection matrix. Hence during the gradient estimation there is no advantage to correlating the instantaneous output power sequence of the single receiver system (or the instantaneous power difference sequence of the dual receiver systems) with the projected perturbation sequence, in an effort to reduce the weight error vector. It is sufficient to only apply the projected perturbation to the weights.

The recursive equation for the norm of the weight error vector *V(n)* is given by

$$\begin{aligned} B(n+1) &= B(n) - 2\mu \mathbf{P}[\mathbf{R}B(n) + B(n)\mathbf{R}]\mathbf{P} \\ &\quad + \mu^2 \mathbf{P} E[\mathbf{V}_G(W(n))]\mathbf{P} + 4\mu^2 \mathbf{P}\mathbf{R}B(n)\mathbf{R}\mathbf{P} \end{aligned} \tag{3.58}$$

where $V_G(W(n))$ denotes the covariance of the gradient estimate used in (2.27) and $B(n) = E[V(n)V^H(n)]$. *B(n)* has the property whereby

$$\mathbf{P}B(n)\mathbf{P} = \mathbf{P}B(n) = B(n)\mathbf{P} = B(n) \tag{3.59}$$

or equivalently $\mathbf{P}V(n) = V(n)$ (3.60)

In the convergence analysis the expressions for the gradient covariance of the three receiver structures are expressed in terms of the weight error vector, and the norm of the weight error vector is expressed in a similar form to Theorem A.2. Conditions for the convergence of the norm of the weight error vector could then be derived. Whereupon it is possible to evaluate the bounds for the steady state excess mean square power. The bounds for the steady state excess mean square power were derived by taking the trace of the norm of the weight error vector and solving it for $Tr[\mathbf{R}B(n)]$.

In the following results λ_{max} denotes the maximum eigenvalue of **R.**

Result 3.5.a. Misadjustment of the Dual Receiver Dual Perturbation System - Bounds Approach

When a $4L$ length projected time multiplex sequence is used such that the gradient covariance is given by (3.8), if

$$E[\|V(0)\|^2] < \infty \text{ and } 0 < \mu < \frac{1}{\frac{Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2} + \lambda_{max}} \tag{3.61}$$

then the steady state excess mean square power is bounded and the corresponding misadjustment is bounded by $b_l \le M \le b_h$ where

$$b_l = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu Tr(\mathbf{P}Diag(\mathbf{PRP}))} \tag{3.62}$$

$$\text{and } b_h = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu[Tr(\mathbf{P}Diag(\mathbf{PRP})) + 2\lambda_{max}]} \tag{3.63}$$

❑❑❑

Result 3.5.b. Misadjustment of the Dual Receiver Reference Receiver System - Bounds Approach

When a $4L$ length odd symmetry projected time multiplex sequence is used such that the gradient covariance is given by (3.10), if

$$E[\|V(0)\|^2] < \infty \text{ and } 0 < \mu < \frac{1}{\frac{Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2} + \lambda_{max}} \tag{3.64}$$

then the steady state excess mean square power is bounded and the corresponding misadjustment is bound by $b_l \le M \le b_h$

$$\text{where} \quad b_l = \frac{\mu\gamma^2 L Tr[\mathbf{P}(Diag(\mathbf{PRP}))^2] + \mu\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}Tr[\mathbf{P}Diag(\mathbf{PRP})]}{(2 - \mu Tr[\mathbf{P}Diag(\mathbf{PRP})])\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}} \tag{3.65}$$

$$\text{and} \quad b_h = \frac{\mu\gamma^2 L Tr[\mathbf{P}(Diag(\mathbf{PRP}))^2] + \mu\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}Tr[\mathbf{P}Diag(\mathbf{PRP})]}{(2 - \mu[Tr[\mathbf{P}Diag(\mathbf{PRP})] + \lambda_{max}])\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}} \tag{3.66}$$

❑❑❑

Result 3.5.c. Misadjustment of the Single Receiver System - Bounds Approach

When a $4L$ length odd symmetry projected time multiplex sequence is used such that the gradient covariance is given by (3.18), if

$$E[\|V(0)\|^2] < \infty \text{ and } 0 < \mu < \frac{4}{6\lambda_{max} + aTr[\mathbf{P}Diag(\mathbf{PRP})]} \tag{3.67}$$

then the steady state excess mean square power is bounded and the corresponding misadjustment is bound by $b_l \le M \le b_h$ where

$$b_l = \frac{\mu a\left(\frac{L-1}{L}\right)Tr(\mathbf{PRP})}{4-\mu(6\lambda_{max}+a\lambda_{max}(L-1))} \tag{3.68}$$

and $$b_h = \frac{\mu a\left(\frac{L-1}{L}\right)Tr(\mathbf{PRP})}{4-\mu(6\lambda_{max}+4\lambda_{max}(L-1))} \tag{3.69}$$

❑❑❑

3.5.2 Comparison of Misadjustment Analyses

From the previous equations, for a suitably small gradient step size it can be observed that for each of the three receiver structures the upper and lower misadjustment bounds are asymptotic. The asymptotic misadjustments are given by:

Result 3.5.d. Asymptotic Misadjustment of the Dual Receiver Dual Perturbation System

$$M = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2} \quad \text{when } \mu[Tr(\mathbf{P}Diag(\mathbf{PRP}))+2\lambda_{max}] \ll 2 \tag{3.70}$$

❑❑❑

Result 3.5.e. Asymptotic Misadjustment of the Dual Receiver Reference Receiver System

$$M = \frac{\frac{\mu\gamma^2 L Tr[\mathbf{P}(Diag(\mathbf{PRP}))^2]}{\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}} + \mu Tr[\mathbf{P}Diag(\mathbf{PRP})]}{2}$$

$$\text{when } \mu[Tr[\mathbf{P}Diag(\mathbf{PRP})]+\lambda_{max}] \ll 2 \tag{3.71}$$

❑❑❑

Result 3.5.f. Asymptotic Misadjustment of the Single Receiver System

$$M = \frac{\mu a\left(\frac{L-1}{L}\right)Tr(\mathbf{PRP})}{4} \quad \text{when } \mu(6\lambda_{max}+a\lambda_{max}(L-1)) \ll 4 \tag{3.72}$$

❑❑❑

From these expressions the following can be concluded:

- Using Lemma A.1 in (3.70), (3.71) and (3.72) the asymptotic misadjustment derived using the *Bounds* technique can be seen to provide an upper bound to those derived using the *Direct* technique.

- If the substitution (3.13) is allowed in (3.70), (3.71) and (3.72) it can also be observed that for a suitably sized gradient step size the two misadjustment expressions derived using the *Bounds* and *Direct* approaches are equal.

- If we compare the previous expressions for the misadjustment bounds to those for the non projected case, which are summarised in Appendix *F*, we can also observe that the bounds in the projected case are smaller than the non projected case. This can be more easily observed by using (3.13) in the expressions. We can also observe that the asymptotic values of the misadjustment bounds are smaller in the projected case. For the single receiver system there is significant reduction in the misadjustment as there is no term that is directly proportional to the optimum power, for the dual receiver cases the difference is not as significant and does depend on the desired signal power. The convergence characteristics for the norm of the weight error vector in the projected and non projected cases are different.

3.6 New Bounds Analysis

A new method to estimate the misadjustment is presented in this section. The new method is based on the techniques developed in [5], but in this instance the bounds on the weight covariance matrix are established to determine the misadjustment. Motivation for this new approach resulted from the fact that the recursive equations for the weight covariance matrix and the norm of the weight error vector are similar, indicating that the *Direct* analysis could be applied to the norm of the weight error vector or the *Bounds* analysis could be applied to the weight covariance matrix. Although the asymptotic misadjustment derived using this approach is shown to be identical to the misadjustment derived with the other techniques, the application of this new approach is more straightforward.

3.6.1 Misadjustment Estimation

The new Bounds approach determines the misadjustment by estimating bounds for $\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}]$. It is a simplified approach compared to the *Bounds* approach as we only consider the system to be in a converged state and do not derive the conditions for convergence. The conditions for convergence of the weight covariance matrix having already been determined in Section 3.4.

In Appendix *D*, a generic expression for the weight covariance matrix is defined and it is solved to obtain the bounds on $\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}]$. The analysis assumes that the conditions for convergence of the weight covariance matrix are satisfied.

The bounds for $\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}]$ are given by:

$$\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}] \le \frac{\lim_{n \to \infty} [\mu b Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \mu Tr(\mathbf{D})]}{4 - g\mu Tr(\mathbf{P}Diag(\mathbf{PRP})) - 4\mu\lambda_{max}(\mathbf{PR})} \tag{3.73}$$

and

$$\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}] \ge \frac{\lim_{n \to \infty} [\mu b Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \mu Tr(\mathbf{D})]}{4 - g\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))} \tag{3.74}$$

where the parameters g, b and $\mathbf{D}$ are defined in Appendix D.

It can be noted from the analysis in Appendix D that when the gradient step size μ is suitably chosen such that the upper and lower bounds approach each other the misadjustment is given by

$$M = \frac{\lim_{n \to \infty} [\mu b Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \mu Tr(\mathbf{D})]}{4\hat{W}^H\mathbf{R}\hat{W}} \tag{3.75}$$

The following components in the weight covariance matrix expression have little or no contribution to the final misadjustment:

$4\mu^2 Tr(\mathbf{PRK}_{WW}(n)\mathbf{RP})$ and $\mu^2 a Tr(\mathbf{K}_{WW}(n)\mathbf{R})Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P})$

These components can thus be described as transient components of the weight covariance matrix as they may only have significant contribution to the misadjustment prior to convergence.

These misadjustment expressions are now evaluated for the three receiver structures where it is assumed that $\lim_{n \to \infty} \overline{\mathbf{W}}(n) = \hat{\mathbf{W}}$.

Result 3.6.a. Misadjustment of the Dual Receiver Dual Perturbation System - New Bounds Approach

Substituting the expressions for the parameters as defined in Appendix D the misadjustment is bounded by $b_l \le M \le b_h$, where

$$b_h = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu Tr(\mathbf{P}Diag(\mathbf{PRP})) - 2\mu\lambda_{max}(\mathbf{PR})} \tag{3.76}$$

and $b_l = \dfrac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu Tr(\mathbf{P}Diag(\mathbf{PRP}))}$ (3.77)

For a suitably small gradient step size, the misadjustment is given by

$$M = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2} \quad \text{when } \mu Tr(\mathbf{P}Diag(\mathbf{PRP})) + 2\mu\lambda_{max}(\mathbf{PR}) \ll 2 \tag{3.78}$$

❑❑❑

Result 3.6.b. Misadjustment of the Dual Receiver Reference Receiver System - New Bounds Approach

Substituting the expressions for the parameters as defined in Appendix D the misadjustment is bounded by $b_l \le M \le b_h$, where

$$b_h = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP})) + \dfrac{\mu\gamma^2 L Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2)}{\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}}}{2 - \mu Tr(\mathbf{P}Diag(\mathbf{PRP})) - 2\mu\lambda_{max}(\mathbf{PR})} \tag{3.79}$$

$$b_l = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP})) + \dfrac{\mu\gamma^2 L Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2)}{\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}}}{2 - \mu Tr(\mathbf{P}Diag(\mathbf{PRP}))} \tag{3.80}$$

For a suitably small gradient step size the misadjustment is given by

$$M = \frac{\mu Tr(\mathbf{P}Diag(\mathbf{PRP})) + \dfrac{\mu\gamma^2 L Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2)}{\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}}}{2}$$

$$\text{when } \mu Tr(\mathbf{P}Diag(\mathbf{PRP})) + 2\mu\lambda_{max}(\mathbf{PR}) \ll 2 \tag{3.81}$$

❑❑❑

Result 3.6.c. Misadjustment of the Single Receiver System - New Bounds Approach

Substituting the expressions for the parameters as defined in Appendix D the misadjustment is bounded by $b_l \le M \le b_h$, where

$$b_h = \frac{\mu a Tr(\mathbf{P}Diag(\mathbf{PRP}))}{4 - \mu a Tr(\mathbf{P}Diag(\mathbf{PRP})) - 4\mu\lambda_{max}(\mathbf{PR})} \tag{3.82}$$

$$b_l = \frac{\mu a Tr(\mathbf{P}Diag(\mathbf{PRP}))}{4 - \mu a Tr(\mathbf{P}Diag(\mathbf{PRP}))} \tag{3.83}$$

For a sufficiently small gradient step size the misadjustment is given by

$$M = \frac{\mu a Tr(\mathbf{P}Diag(\mathbf{PRP}))}{4} \text{ when } \mu a Tr(\mathbf{P}Diag(\mathbf{PRP})) + 4\mu\lambda_{max}(\mathbf{PR}) \ll 4 \tag{3.84}$$

❑❑❑

3.7 Comparison of Misadjustment Expressions

Comparing the *New Bounds* expressions derived in the previous section with those obtained with the *Direct* and *Bounds* analysis the following can be observed.

- For the dual receiver systems the lower misadjustment bounds of the New Bounds and the Bounds approach are identical. While the upper misadjustment bounds of the New Bounds approach are in general higher due to the eigenvalues of **PR** being smaller than the eigenvalues of **R**.

- For the dual receiver structures *Assumption 3.3* is not required in the New Bounds Approach. For the single receiver case *Assumption 3.3* is not applied in the New Bounds approach but it is still used to determine the optimum perturbation step size. Hence the bounds derived with the New Bounds approach are more accurate in the sense that they rely less on *Assumption 3.3* being satisfied.

- For each receiver structure the asymptotic misadjustment expressions obtained in the *New Bounds* and the *Bounds* approach are identical.

- For each receiver structure, by using the following approximations in the *Direct*, *Bounds* and *New Bounds* misadjustment expressions the misadjustment expressions can be shown to be equal.

$$
\begin{aligned}
&Diag(\mathbf{PRP}) \cong \frac{1}{L} Tr(\mathbf{PRP})\mathbf{I}_{LL} \\
&\Rightarrow Tr(\mathbf{P}Diag(\mathbf{PRP})) \cong \frac{L-1}{L} Tr(\mathbf{PRP}) \\
&\Rightarrow Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2) \cong \frac{L-1}{L^2}(Tr(\mathbf{PRP}))^2
\end{aligned}
\tag{3.85}
$$

3.8 Simulation Studies

In this section, results of computer studies performed to confirm the accuracy of the expressions derived in the previous sections are presented. In the studies, simulations of the narrowband signals were performed using the techniques discussed in [11],[23]. The misadjustment was calculated for various scenarios as described in the Figures.

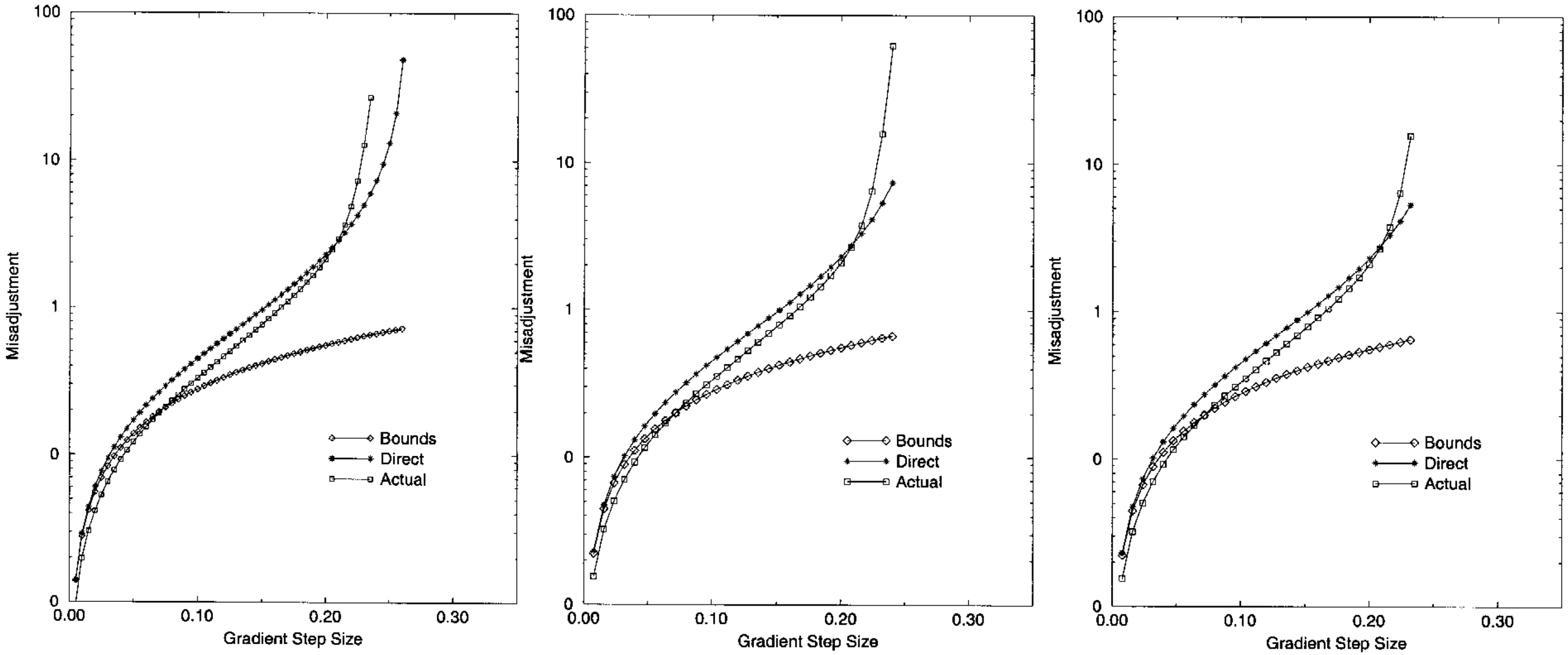
For the dual receiver system simulations a linear array with four equally spaced omnidirectional array elements with a quarter wavelength spacing at a nominal operating frequency was used. For the single receiver system simulations a linear array with four or six equally spaced omnidirectional array elements with a quarter wavelength spacing at a nominal operating frequency was used.

In the Figures the signal scenarios are described. Three signals are incident on each of the arrays, two signals are interference signals and the third is the desired signal. The desired signal's direction of arrival corresponds to the look direction. We assume the signals are coplanar such that they arrive in the xy plane where $\theta = 90$ and the elements of the array lie along the x axis. The direction of arrival of the signals are described in terms of the parameter ϕ as defined in Figure 1 and expressed in degrees.

White noise power of 10^{-4} relative to a unit reference power was added to each array element. The misadjustment was calculated at convergence by averaging the excess mean square output power in 20 blocks of 100 iterations. The powers of each signal are defined relative to the same unit reference power used to set the white noise level. The power of each signal is also described in the figures.

In Figure 8 to Figure 10 for the projected perturbation approach we show the misadjustment for the three receiver structures for different gradient step sizes. In Figure 8 and Figure 9 the curves labelled "*Bounds*" corresponds to the computed asymptotic misadjustment as derived in the *Bounds* and *New Bounds* analysis and are given by (3.70) or (3.71) as appropriate. The curve labelled "*Direct*" corresponds to the computed misadjustment as derived in the *Direct* analysis assuming μ is small and $\mu\lambda_i \ll 1$ and are given by (3.47) or (3.50) as appropriate. In Figure 10 the curves labelled "*Upper and Lower Bound*" corresponds to the computed misadjustment as derived in the *Bounds* analysis and are given by (3.68) and (3.69) as appropriate. The curve labelled "Actual" represents the results of the simulation. These simulations have been performed with a time invariant optimum perturbation step size given by (3.20). In each of these figures the trends suggested in the previous sections are confirmed. That is, at a suitably small gradient step size the *Direct* and *Bounds* misadjustment estimation approach eachother, the asymptotic bounds provides a lower bound to the *Direct* misadjustment expression and the theoretical expressions provide a good estimate of the real misadjustment.

Note that when examining the misadjustment at very small gradient step sizes, where we would expect close agreement, there is still some difference between the theoretical results (*Direct* and *Bounds*) and the actual misadjustment. This was found to occur due to the method used in simulating the narrowband signals. It was found that the variance in the white noise generator didn't produce an exact covariance of 1 and in fact the expected variance in the white noise generator was of the order of the parameter being measured. The simulation was shown to produce correct results by determining where the actual misadjustment becomes unbounded using the *Direct* misadjustment expressions.

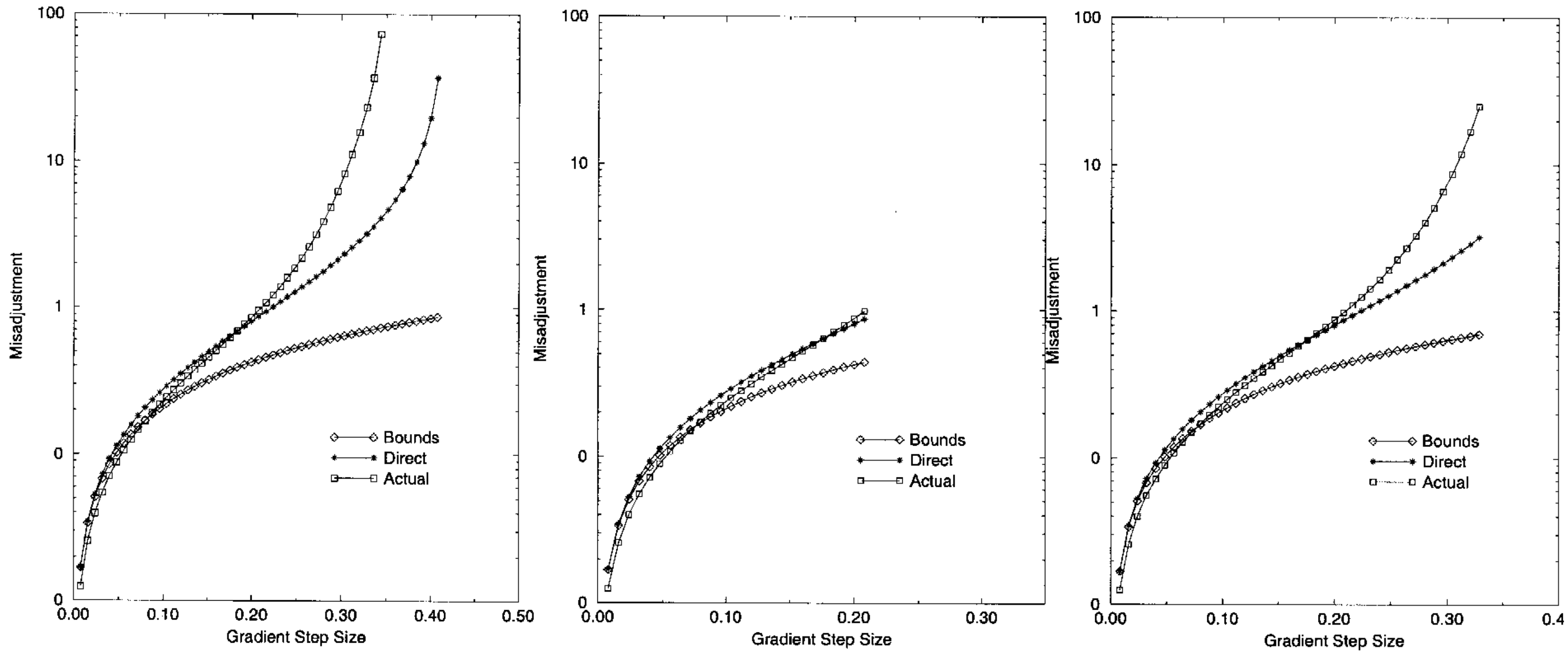


Signal	Direction of Arrival φ	Power
Desired	90	0.01
Interference 1	45	0.5
Interference 2	180	0.5

Signal	Direction of Arrival φ	Power
Desired	90	0.25
Interference 1	45	0.5
Interference 2	180	0.5

Signal	Direction of Arrival φ	Power
Desired	90	0.5
Interference 1	45	0.5
Interference 2	180	0.5

Figure 8 Misadjustment, Dual Receiver Dual Perturbation System, Projected Perturbation Approach

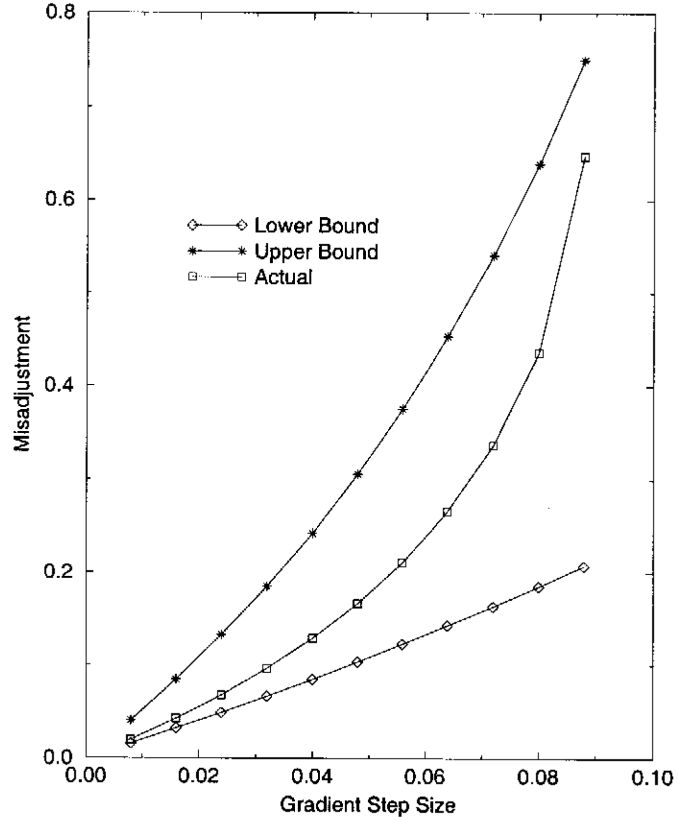


Signal	Direction of Arrival ϕ	Power
Desired	90	0.01
Interference 1	45	0.5
Interference 2	180	0.5

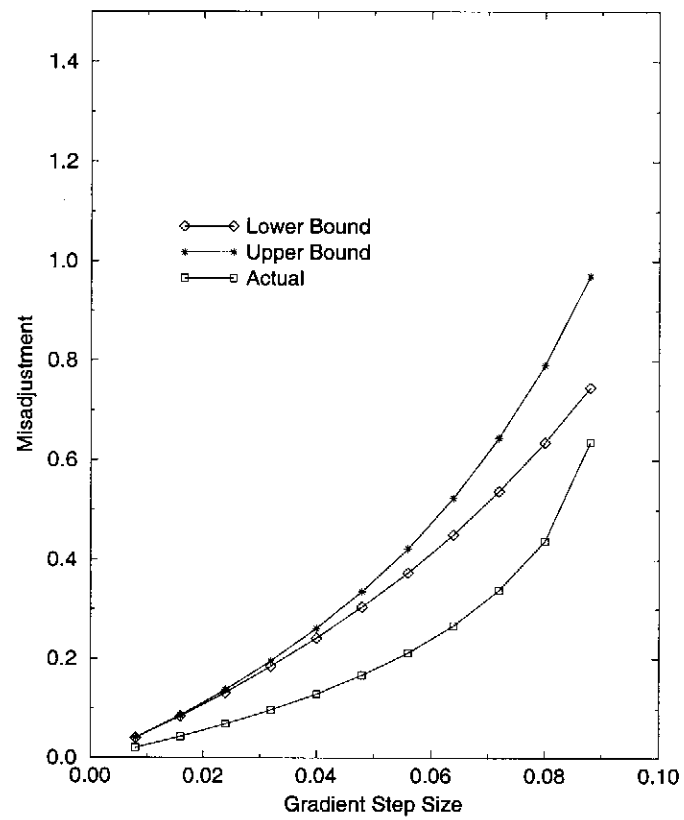
Signal	Direction of Arrival ϕ	Power
Desired	90	0.25
Interference 1	45	0.5
Interference 2	180	0.5

Signal	Direction of Arrival ϕ	Power
Desired	90	0.5
Interference 1	45	0.5
Interference 2	180	0.5

Figure 9 Misadjustment, Dual Receiver Reference Receiver System, Projected Perturbation Approach



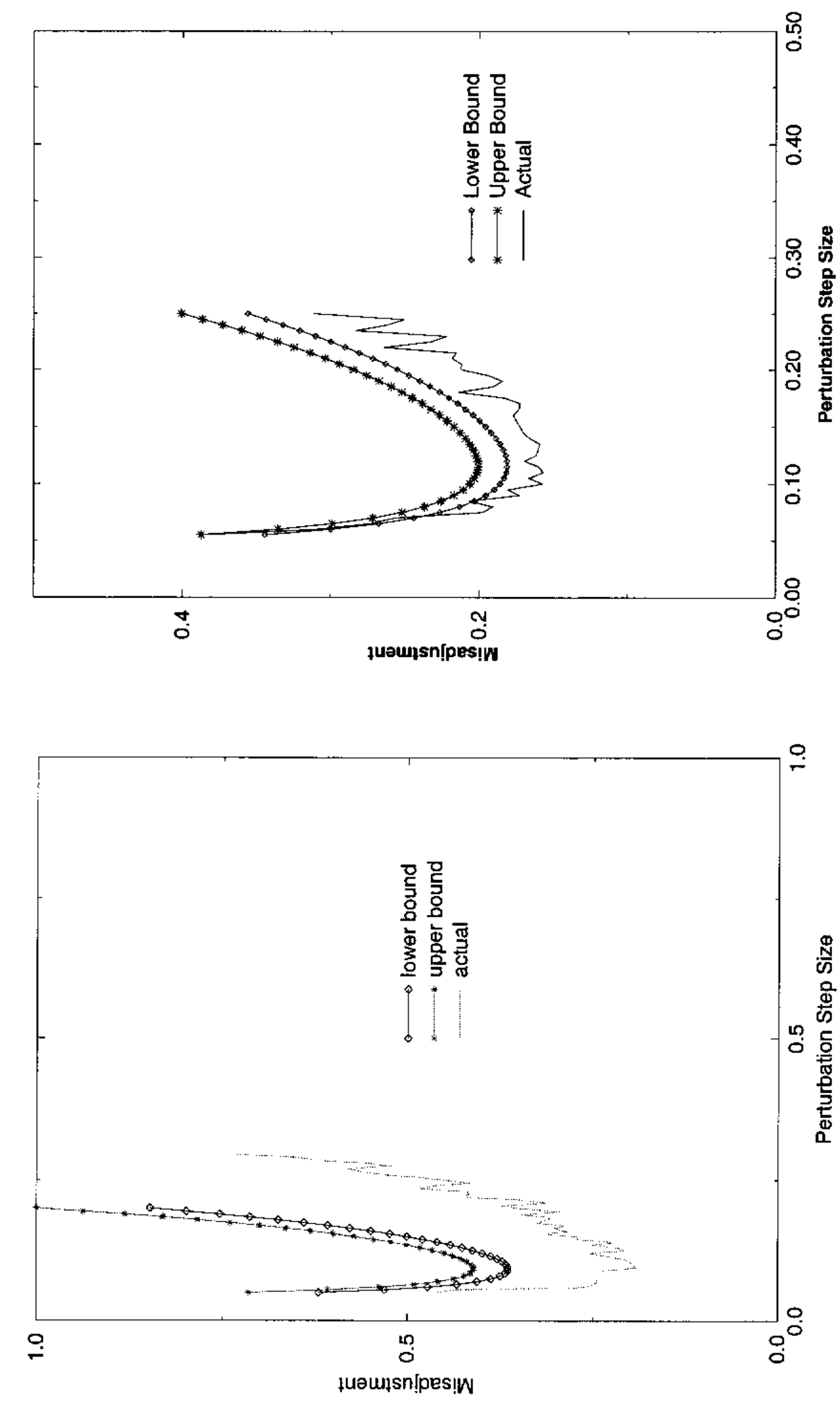
Signal	Direction of Arrival ϕ	Power
Desired	90	0.01
Interference 1	45	0.5
Interference 2	180	0.5



Signal	Direction of Arrival ϕ	Power
Desired	90	0.5
Interference 1	45	0.5
Interference 2	180	0.5

Figure 10 Misadjustment, Single Receiver System, Projected Perturbation Approach

Figure 11 shows the misadjustment versus the perturbation step size for the single receiver system. This figure confirms that an optimum perturbation step size exists for the single receiver system and that there is close agreement with the expression derived for the optimum perturbation step size. Note that the simulated values for the misadjustment do not fall within the derived bounds because a time invariant perturbation step size is used in the simulations.



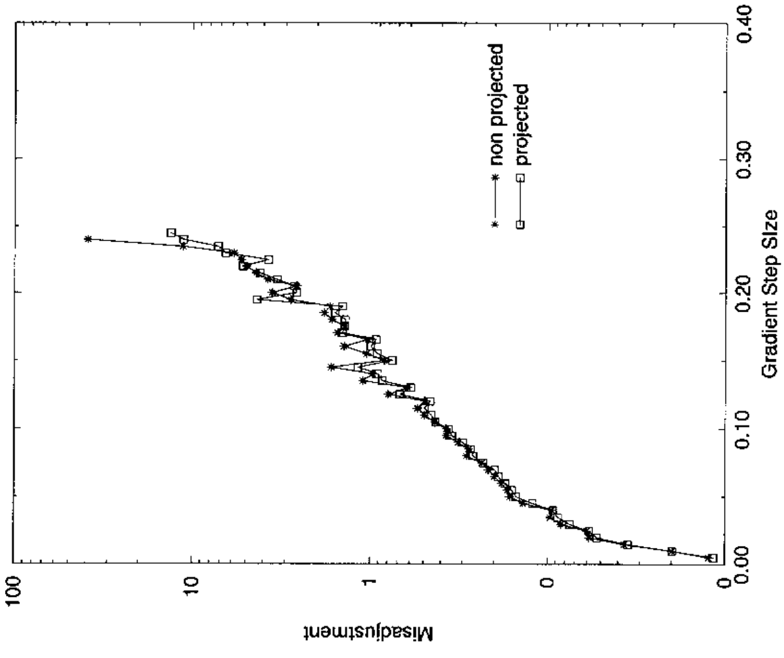
6 element array, gradient step size =0.055

4 element array, gradient step size=0.055

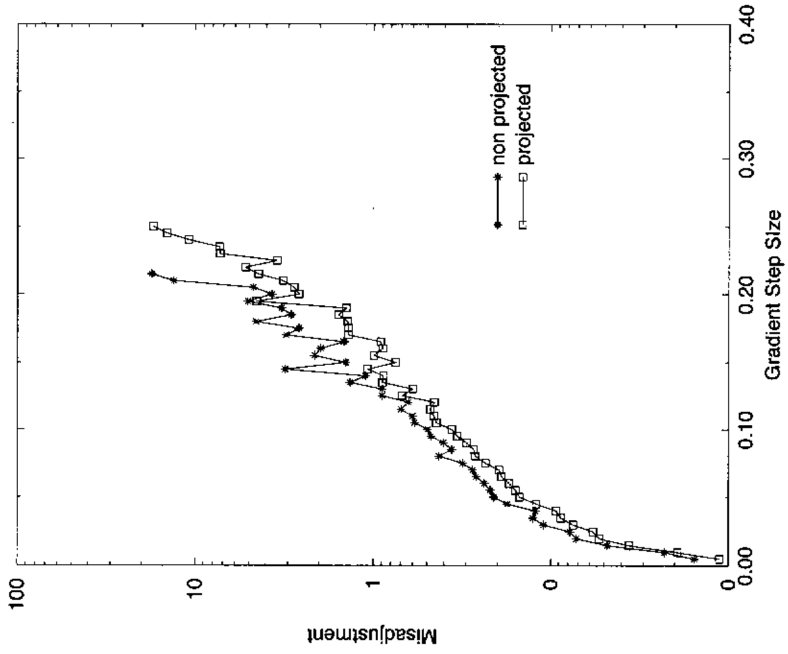
Signal	Direction of Arrival φ	Power
Desired	90	0.1
Interference 1	45	0.5
Interference 2	180	0.5

Figure 11 Single Receiver System Optimum Perturbation Step Size

Figure 12, Figure 13 and Figure 14 show the misadjustment versus gradient step size for the projected and non projected perturbation approaches for all receiver structures. One can observe from these figures that as the desired signal power increases there is a noticeable difference in the systems performance.



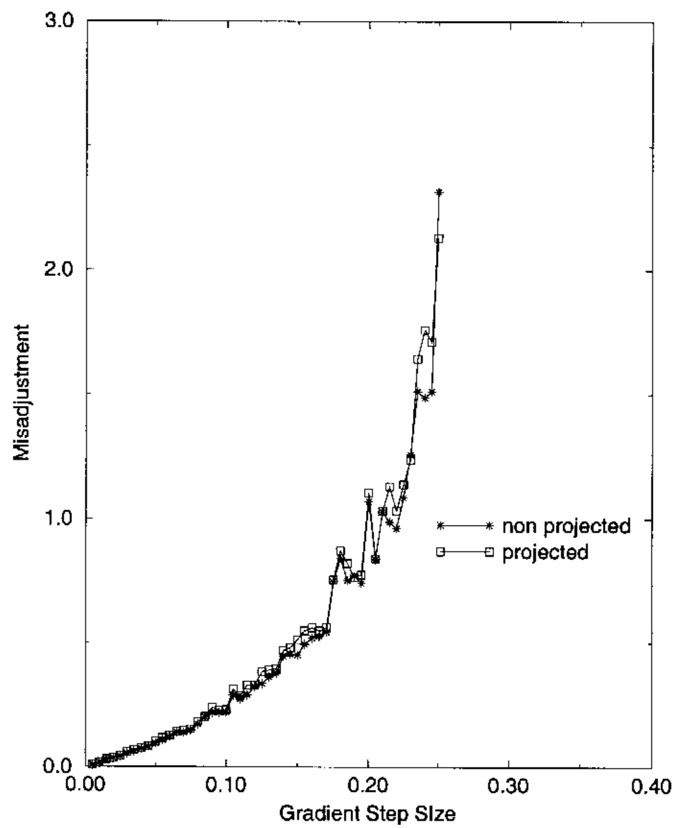
Signal	Direction of Arrival φ	Power
Desired	90	0.01
Interference 1	45	0.5
Interference 2	180	0.5



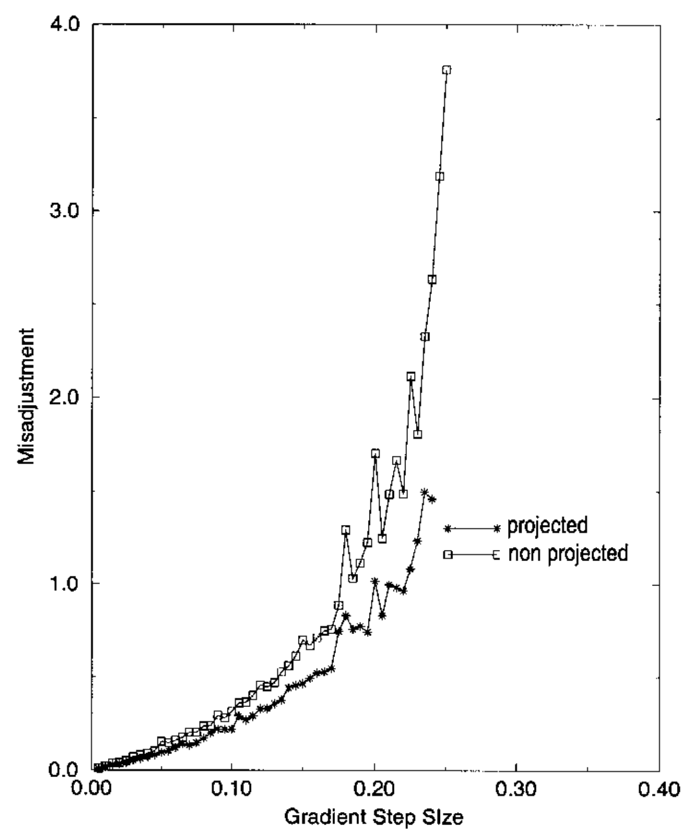
Signal	Direction of Arrival φ	Power
Desired	90	0.25
Interference 1	45	0.5
Interference 2	180	0.5

Dual Receiver Dual Perturbation System

Figure 12 Misadjustment, Projected Approach vs Non Projected Approach



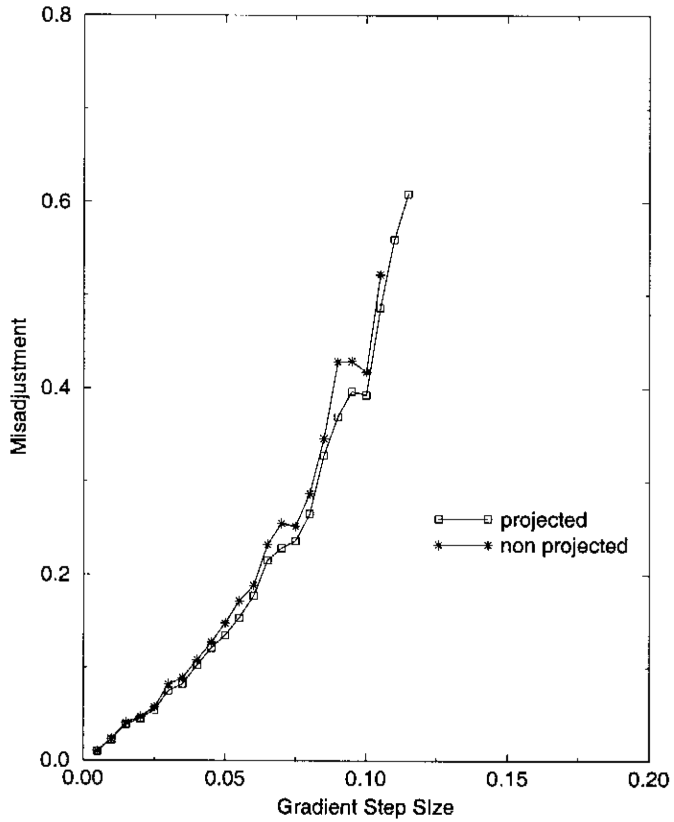
Signal	Direction of Arrival φ	Power
Desired	90	0.01
Interference 1	45	0.5
Interference 2	180	0.5



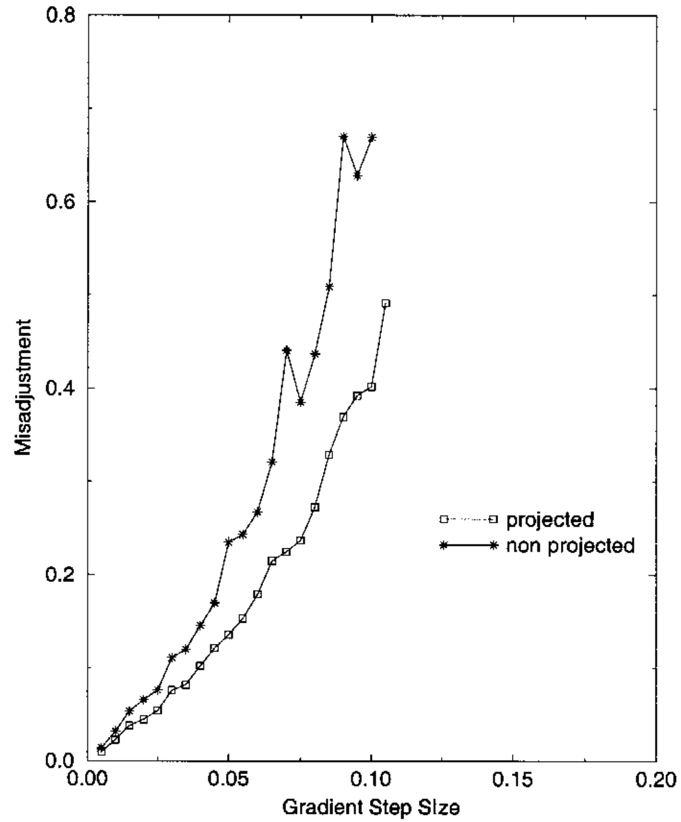
Signal	Direction of Arrival φ	Power
Desired	90	0.25
Interference 1	45	0.5
Interference 2	180	0.5

Dual Receiver Reference Receiver System

Figure 13 Misadjustment, Projected Approach vs Non Projected Approach



Signal	Direction of Arrival ϕ	Power
Desired	90	0.01
Interference 1	45	0.5
Interference 2	180	0.5



Signal	Direction of Arrival ϕ	Power
Desired	90	0.25
Interference 1	45	0.5
Interference 2	180	0.5

Single Receiver System

Figure 14 Misadjustment, Projected Approach vs Non Projected Approach

3.9 Summary

In this chapter we have analysed the performance of three adaptive array structures using projected perturbation sequences by deriving the covariance of the gradient estimate and the misadjustment expression using various approaches. It has been shown that the projected perturbation sequences generate smaller variance in the gradient estimate and hence smaller misadjustment compared to non projected perturbation sequences. The difference in performance depends on the look direction signal power. It has been shown that for a single receiver system an optimum perturbation step size exists. All results have been confirmed by computer simulation.

Chapter 4

Quantisation Effects in the Implementation of Projected Perturbation Sequences

4.1 Introduction

In the implementation of digital systems practical issues such as processor wordlength, signal and coefficient quantisation and computational times can cause the system performance to deviate from the ideal. For the conventional Least Mean Square adaptive algorithm processor wordlength and signal and coefficient quantisation effects have been well documented and widely studied [13], [14], [20], [21], [22], [33], [45]. However these effects have not been studied for the least Mean Square adaptive algorithm when the gradient estimate is obtained using perturbation based techniques.

For a perturbation based adaptive array there are three main sources of quantisation errors: input quantisation where the signals are quantised, coefficient quantisation where the array weights and the weight update algorithm constants are quantised, and quantisation in arithmetic operations. The errors introduced by quantisation can be minimised by design. For example, during adaptation, depending on the perturbation sequence, the effect of weight quantisation on the look direction response can be eliminated by designing the perturbation step size such that a weight perturbation is equal to an integral number of weight quantisation levels [5]. However, this is only possible when the perturbation sequences are similar to the Time Multiplex sequences and in general it is not possible for projected perturbation sequences. In this Chapter we study the effects of weight quantisation on the performance of an adaptive array processor that uses the projected perturbation approach.

A contribution of this thesis is that we determine the level of loss of performance due to weight quantisation and the limited dynamic range of the array weights for an adaptive array that uses the projected perturbation approach. In particular we develop new expressions for the gradient covariance and the misadjustment in the presence of weight quantisation for all the receiver structures under study and we determine conditions under which the properties of the projected perturbation sequence are preserved. The results presented here have been published by the author in [49].

This Chapter is organised as follows. In Section 4.2 we briefly review common digital implementation effects and their impact on the projected perturbation approach. In Section 4.3 the model for the quantisation process is introduced. In Section 4.4 we derive two performance measures, the interference rejection capability of the processor and the misadjustment. For the misadjustment analysis it is necessary to first establish the gradient covariance in the presence of weight quantisation. The misadjustment is then derived using the *Bounds* analysis approach. Finally in Section 4.5 simulation results and concluding remarks are presented.

For the sake of conciseness, an outline of the major results are presented here. A detailed derivation of the intermediate results are given in Appendix *E*.

4.2 Digital Implementation Effects

In this section, we briefly review some of the effects due to digital implementation with specific reference to a perturbation based adaptive array.

Basically the effects of digital implementation can be studied in terms of:

- computational time
- finite precision effects

Non zero computational time gives rise to computational delay. By computational delay we mean the period of time between the instant a new measurement sample is taken at the processor input, and the instant the updated weight is available at the output.

We broadly class finite precision effects as quantisation effects, computational round off error effects and saturation effects. The finite precision effects are interrelated and result from system limitations introduced by:

- the finite dynamic range of data representation where by data we mean the signals, weights and co-efficients such as the gradient and perturbation step sizes etc.
- the finite number of representable values (or quantisation levels) available.
- errors resulting from arithmetic computation.

Quantisation effects occur when data is approximated to one of the representable values. We refer to the truncation or the rounding-off of data to a finite precision as quantisation. In Figure 15 we show where quantisation can take place in the adaptive array processor. The figure only depicts the quantisation of the inphase signal from the l^{th} array element. Quantisation can take place at the analog to digital converters, and within the processor where there may be co-efficient quantisation and quantisation in arithmetic calculations. Quantisation of co-efficients such as the gradient step size μ, occur due to the finite wordlength of the processor. Quantisation of the array signal and weights can also occur due to the finite-wordlength of the processor and the analog to digital converters. The quantisation process generally results in an error component in the represented value, and for a continuous signal this error results in an additional noise component [13], [14], [26], [30], [58]. In a successful implementation, the impact of quantisation on performance has to be small.

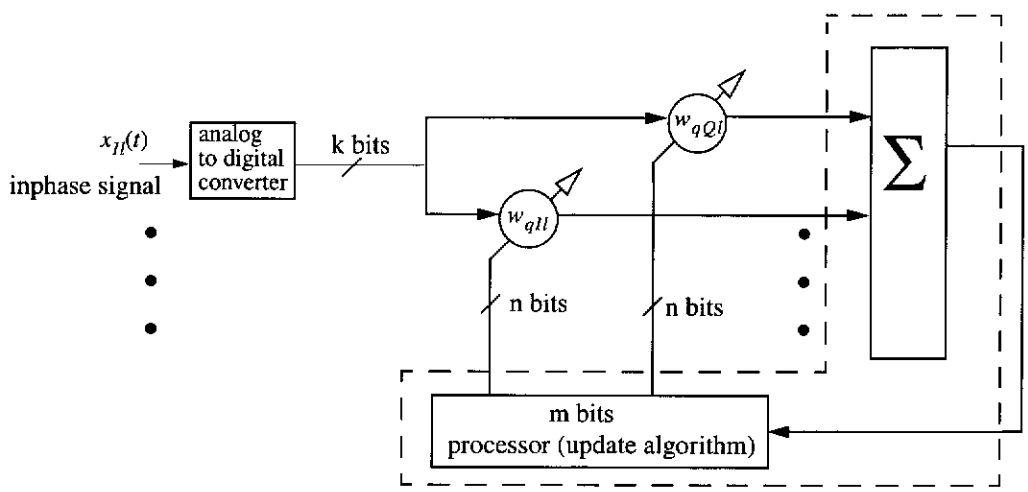


Figure 15 Quantisation of the l^{th} Array Element Inphase Signal

Due to the errors introduced by the quantisation of the weights, the properties of a perturbation sequence can be altered. The alteration can result in a biased estimate of the gradient being obtained, and in the case of a projected perturbation sequence not preserving the look direction constraint during adaptation. As mentioned in the introduction these effects can be minimised by designing the quantisation step size to be equal to a weight perturbation [5]. This strategy is possible with perturbation sequences whose vector components are integer multiples of each other but it is rarely possible for the projected perturbation sequence. Also a property of the projected perturbation approach is the reduction in perturbation noise. When the weights are quantised this benefit can be reduced.

We refer to computational round off errors as the errors resulting from arithmetic computation. Computational round off errors occur when intermediate calculations, such as multiplication, require greater precision than the processor makes available. The computational round off errors affect the numerical accuracy of the internal calculations and in turn the quality of the resultant beam. Computational round off errors are generally well understood [21], [22].

Saturation effects occur when data falls outside the available data range or intermediate calculations within the processor cause registers to overflow or underflow. The data range is generally limited by the wordlength of the processor. The finite range constraint implies that data cannot take arbitrarily large or small magnitudes. A successful implementation will require both proper data scaling, and data overflow and underflow management.

Computational round off errors can cause the weight update algorithm given by (2.27), to experience premature digital stopping. Equation (2.27) is repeated below for a single look direction constraint system. Digital stopping can occur when there is insufficient precision in the arithmetic calculation such that the error term $\mu G(W(n))$, is negligible compared to the nominal weight $W(n)$ and also when the update algorithm produces successive values whose difference is small relative to the size of a quantisation level. Using a larger gradient step size μ, can delay the effects of digital stopping, however using a larger μ will result in a larger misadjustment. As compared to the analog case, when adaptation terminates due to digital stopping it can result in a larger residual mean square error [20]. Premature digital stopping can be overcome in part by the dynamic scaling of weights [36] or with the use of virtual weights [9].

Weight Update Algorithm

$$W(n+1) = \mathbf{P}[W(n) - \mu G(W(n))] + \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1} \qquad (4.1)$$

With the use of virtual weights two sets of weights are used in the system. One set of weights known as the hardware weights represents the actual quantised array weights. The second set of weights maintained in software are known as the virtual weights and represents a higher precision of the hardware weights. The virtual weights are restricted by the finite wordlength of the processor [21], [22]. The hardware weights are obtained from the virtual weights by quantisation methods. In our quantisation model for the implementation we consider the use of virtual weights.

Using virtual weights, the weight update algorithm defined by (4.1) can be implemented as follows where $W_q(n)$ is the new hardware weight computed at the n^{th} iteration and where $W(n)$ is the new virtual weight computed at the n^{th} iteration

Hardware Weight Update Algorithm

1. Estimate the gradient by applying perturbations about the nominal hardware weight $W_q(n)$

2. Update the virtual weight vector,

$$W(n+1) = \mathbf{P}[W(n) - \mu G(W_q(n))] + \frac{C}{C^H C}$$

3. Obtain the next hardware weight, $W_q(n+1) = Q(W(n+1))$ (4.2)

In (4.2) Q denotes the quantisation operation where

$$Q(W(i)) = W(i) + \eta(i) \tag{4.3}$$

and $\eta(i)$ represents the $L \times 1$ dimensional vector of weight quantisation errors at the i^{th} instant of time.

The virtual weights act as a vector integrator accumulating the relatively small weight updates. The hardware weights will change when the virtual weights have been displaced sufficiently to cause a jump in the quantised weights. Stability aspects of using two sets of weights has been studied in [9]. The results indicate that this method makes a good attempt at finding a feasible solution closest to the optimum.

In the rest of this Chapter we concentrate only on the effects of weight quantisation on the performance of the array since for the conventional LMS algorithm it has been shown that the weights are the most sensitive of the parameters to quantisation [45]. Also we do not consider the computational time as the high speed of modern processors renders the computational time effect less significant.

4.3 Quantisation Error Model

In this section, we introduce the quantisation model and detail the assumptions made to model the quantisation effects.

We only consider the effects of real weight quantisation and we assume that the digital signal representation is of a high precision such that it can be considered ideal and that other arithmetic processes required in the weight update algorithm maintain the characteristic of an infinite precision system. The effective system being analysed is analogous to an analogue system with digitally controlled weights.

In our model of the implementation, we assume that virtual weights are used and that the complex weighting is realised by splitting the outputs from each array element into two phase quadrature channels and passing each through variable stepped attenuators. Methods to produce the quadrature signals are outlined in Chapter 2. The amplitude gain of each attenuator is set by a binary word so that each array weight can then be regarded as a complex variable whose real and imaginary parts have finite precision. In Figure 15 the complex weighting of the inphase signal from the l^{th} array element is shown.

Modelling the effects of weight quantisation in any process poses some difficulties since quantisation errors may be time invariant at system equilibrium and are generally not treatable as random. These quantisation errors may be related to correlations present within the system. To simplify the task we make the following general assumptions:

Assumption 4.1: Round-off quantisers are used and the resulting components of the weight quantisation error vector η_l $l = 1, \ldots L$, are zero mean, uncorrelated and uniformly distributed over the quantiser bin width. Furthermore, the quantiser inphase and quadrature bin widths are all uniform and equal to ΔW.

Using the above assumption the quantisation errors satisfy

$$E[\eta(i)\eta^H(i)] = \sigma_\eta^2 \mathbf{I}_{LL} \text{ and } E[\eta(i)\eta^H(j)] = 0 \text{ , } i \neq j \qquad (4.4)$$

It can be shown that [26],

$$\sigma_{\eta}^2 = \frac{(\Delta W_{real})^2}{12} + \frac{(\Delta W_{Imag})^2}{12} = \frac{(\Delta W)^2}{6} \tag{4.5}$$

where $\Delta W_{real} = \Delta W_{imag} = \Delta W$ is the magnitude of the quantisation increment.

Assumption 4.2: The quantisation is fine enough to prevent signal correlated patterns in the quantisation errors such that the distortion produced by quantisation affects the performance of the system as if it were an additive independent source of noise

Conditions that allow the quantisation errors to be modelled as an additive uniform white noise are examined in [58]. The additional errors that are introduced in the model when these conditions are not met and signal correlated patterns in the quantisation errors are also examined. For the example signal scenarios in [14], [30],[50] if the wordlength is greater than six bits the vector of quantisation errors could be accurately regarded as being uncorrelated with the signals and its elements as being mutually uncorrelated.

During adaptation there are two methods that can be used to obtain the hardware weights they are:

Method 1. Virtual Weights and Weight Perturbations are quantised separately

$$W_q(n) = Q(W(n)) \pm Q(\gamma\delta_p(i))$$

Method 2. Virtual Weights and Weight Perturbations are quantised together.

$$W_q(n) = Q(W(n) \pm \gamma\delta_p(i))$$

It is shown in Appendix *E* that by using Method 2 with the quantisation modelling assumptions made above, an unbiased estimate of the gradient can be obtained for each of the receiver structures under study. Note that the appropriate perturbation sequence must still be used. To obtain an unbiased gradient estimate using Method 1 the following assumption is also required

Assumption 4.3: The quantised error vector of a perturbation sequence has zero mean over a perturbation cycle.

Note that in Appendix E the analysis for either method requires the appropriate perturbation sequence to be used in order to obtain an unbiased gradient estimate. For simplicity in future analysis we will assume that during adaptation, quantisation takes place using Method 2.

4.4 Performance Measures

In this section, we examine criteria that can be used to gauge the performance of perturbation based adaptive arrays implemented with quantised weights. We consider the interference rejection capability of the array and develop new misadjustment expressions for the three receiver structures under study.

The interference rejection capability of the processor gives an insight of the effect of quantisation on the optimum weights. The new misadjustment expressions were developed as they also indicate the effect that the weight quantisation has on the gradient estimation process.

4.4.1 Effect of Quantisation on Interference Rejection

The level of loss of performance due to quantisation can be determined from the array's interference rejection capability. This can be established by examining the array's gain in the direction of a plane wave interference which lies in the sidelobes of a conventional array pattern.

When the virtual weight vector can be exactly set to null an interference and assuming there is no other noise present in the system it can be shown that the variance of the output signal power is given by [14]

$$E[S^H \eta \eta^H S] = L\sigma_\eta^2 = L\frac{(\Delta W)^2}{6} = \frac{2}{3}L\left(\frac{W_{max}}{Z-1}\right)^2 \tag{4.6}$$

In (4.6) the expectation is taken over all possible quantisation errors, S is the steering vector of the interference and ΔW has been defined by finding the range of the virtual weight vector. For Z equal levels of quantisation ΔW is given by

$$\Delta W = \frac{2W_{max}}{Z-1} \tag{4.7}$$

Equation (4.6) corresponds to the depth of the null. The number of bits required to achieve a desired null depth in the interference direction can be determined using (4.6). One method to determine the range of the virtual weight vector is by solving for the optimum weight vector using the expected signal scenario. Alternative methods as described in [9] could also be applied.

It can be observed from (4.6) that the effect of quantisation is to limit the degree of rejection of interference sources, hence by choosing the desired rejection capability of the processor a predictable interference scenario can be expected. It also indicates that as the number of quantisation levels increase the interference rejection capability improves. This performance measure though gives no indication on the weight quantisation effects during gradient estimation and does not distinguish between the different gradient estimation schemes under study.

4.4.2 Covariance of the Gradient Estimate

In this section, we present the expressions for the gradient covariance in the presence of weight quantisation. These expressions are required in the next section to determine the misadjustment. The expressions indicate the effect of weight quantisation on the gradient estimation process.

The covariance expressions are derived in Appendix *E* where the derivation is based on the approach used in [2], [5]. In the analysis, in addition to the assumptions of the quantisation model discussed earlier, we require the same assumptions used in Chapter 3, i.e *Assumption 3.1* and *3.2*. The assumptions of the quantisation model are required so that an unbiased gradient estimate can be obtained.

In Appendix *E*, the ***Generic Expressions*** for the gradient covariance have been derived for the projected perturbation and hybrid perturbation approaches. We refer to the expressions as being generic since they are applicable with any projected orthogonal perturbation sequence. Below we have evaluated these expressions for the projected perturbation approach that uses the projected time multiplex sequence. Note that the expressions do not take into consideration the reduced length perturbation sequences and, for simplicity of notation we write $\boldsymbol{W}(n)$ as $\boldsymbol{W}$.

Result 4.4.a Gradient Covariance of Dual Receiver Dual Perturbation System with Projected Time Multiplex Sequence

$$\mathbf{V}_{G1}(W_q(n)) = \mathbf{V}_{G1_{orig}}(W) + \sigma_\eta^2 \mathbf{P}(Diag(\mathbf{P}\mathbf{R}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{P}\mathbf{R}\mathbf{P}))\mathbf{P} + \frac{\sigma_\eta^2}{\gamma^2 2L}\left(W^H\mathbf{R}^2 W + Tr(\mathbf{R})W^H\mathbf{R}W + \frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\right)\mathbf{P}$$

for a 4L length sequence (4.8)

$$\mathbf{V}_{G1}(W_q(n)) = \mathbf{V}_{G1_{orig}}(W) + 2\sigma_\eta^2 \mathbf{P}(Diag(\mathbf{P}\mathbf{R}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{P}\mathbf{R}\mathbf{P}))\mathbf{P} + \frac{\sigma_\eta^2}{\gamma^2 L}\left(W^H\mathbf{R}^2 W + Tr(\mathbf{R})W^H\mathbf{R}W + \frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\right)\mathbf{P}$$

for a 2L length sequence (4.9)

❑❑❑

Result 4.4.b. Gradient Covariance of Dual Receiver Reference Receiver System with Projected Time Multiplex Sequence

$$\mathbf{V}_{G2}(W_q(n)) = \mathbf{V}_{G2_{orig}}(W) + 2\sigma_\eta^2 \mathbf{P}(Diag(\mathbf{P}\mathbf{R}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{P}\mathbf{R}\mathbf{P}))\mathbf{P} + \frac{2\sigma_\eta^2}{\gamma^2 L}\left(W^H\mathbf{R}^2 W + Tr(\mathbf{R})W^H\mathbf{R}W + \frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\right)\mathbf{P}$$

for a 4L length sequence, Approach 1 (4.10)

$$\mathbf{V}_{G2}(W_q(n)) = \mathbf{V}_{G2_{orig}}(W) + \sigma_\eta^2 \mathbf{P}(4Diag(\mathbf{P}\mathbf{R}^2\mathbf{P}) + 2Tr(\mathbf{R})Diag(\mathbf{P}\mathbf{R}\mathbf{P}))\mathbf{P} + \frac{\sigma_\eta^2}{\gamma^2 L}(W^H\mathbf{R}^2 W + Tr(\mathbf{R})W^H\mathbf{R}W + \sigma_\eta^2(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2))\mathbf{P}$$

for a 4L length sequence, Approach 2 (4.11)

❑❑❑

For the dual receiver reference receiver system two approaches can be taken to determine the gradient covariance. It can be assumed that the reference receiver quantisation errors have the same model as the perturbed receiver quantisation errors, Approach 1, or the reference receiver quantisation errors are constant in the gradient estimation period and that this error is small and can be ignored, Approach 2. In latter derivations we assume the modelling of Approach 1 is used.

Result 4.4.c. Gradient Covariance of Single Receiver System with Projected Time Multiplex Sequence

$$\mathbf{V}_{G3}(W_q(n)) = \mathbf{V}_{G3_{orig}}(W) + 2\sigma_\eta^2\mathbf{P}(2Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$
$$+ \frac{\sigma_\eta^2}{\gamma^2 L}\left(W^H\mathbf{R}^2W + Tr(\mathbf{R})W^H\mathbf{R}W + \sigma_\eta^2\left(Tr(\mathbf{R}^2) + \frac{1}{2}(Tr(\mathbf{R}))^2\right)\right)\mathbf{P} \quad (4.12)$$

for a 4L length sequence

❑❑❑

In (4.8), (4.9), (4.10), (4.11) and (4.12) $\mathbf{V}_{G1_{orig}}(W(n))$, $\mathbf{V}_{G2_{orig}}(W(n))$ and $\mathbf{V}_{G3_{orig}}(W(n))$ are the covariance expressions with no weight quantisation effects, they are defined by (3.8), (3.9), (3.10) and (3.11) respectively.

Examining (4.8), (4.9), (4.10), (4.11) and (4.12) and comparing them to the equivalent expressions developed in Chapter 3 the following observations can be made:

- All the additional terms are proportional to the variance of the quantisation errors. The effect of quantisation can be limited by making the variance of quantisation errors small.

- The majority of the additional covariance terms are proportional to the number of array elements. It can be expected that as the number of array elements is increased the misadjustment will increase. To make this observation we note that for a simple, single signal scenario using (2.14) in (2.18) the approximation $\mathbf{R}^2 = L\mathbf{R}$ is valid.

- Some of the additional covariance terms are inversely proportional to γ, indicating that the quantisation process favours a larger perturbation step size. However the perturbation noise performance of a system favours a smaller perturbation size.

- As in Chapter 3 for the dual receiver dual perturbation system the gradient covariance for the $4L$ length perturbation sequence is half that of the gradient covariance for the $2L$ length sequence.

We will latter examine the accuracy of these observations in the simulation studies section.

4.4.3 Misadjustment

In this section we derive the misadjustment using the *Bounds* approach.

Using the definition for misadjustment (2.53), when the weights are quantised, the misadjustment is given by

$$M = \lim_{n \to \infty} \frac{E[(W(n)+\eta)^H X(n+1) X^H(n+1)(W(n)+\eta)] - \hat{W}^H \mathbf{R} \hat{W}}{\hat{W}^H \mathbf{R} \hat{W}} \tag{4.13}$$

In (4.13), η represents the quantisation errors in the virtual weights $W(n)$.

We have assumed previously that $X(.)$ is an independent and identically distributed process (*Assumption 3.1*). Hence $W(n)$ is independent of $X(k+1)$ for $k \geq n$, and the expectation over X and W can be taken separately. We have also assumed that η is independent of $X(.)$ and W *(Assumption 4.1)*. Hence the expectation over η, $X(.)$ and W can also be taken separately.

Expanding (4.13), taking expectation with respect to η and using (4.4), Lemma A.9 and the zero mean property of η, gives

$$M = \lim_{n \to \infty} \frac{E[\ W^H(n) X(n+1) X^H(n+1) W(n) + Tr(\sigma_\eta^2 X(n+1) X^H(n+1))] - \hat{W}^H \mathbf{R} \hat{W}}{\hat{W}^H \mathbf{R} \hat{W}}$$

Now taking expectation with respect to $X(.)$ gives

$$M = \lim_{n \to \infty} \frac{E[W^H(n) \mathbf{R} W(n)] - \hat{W}^H \mathbf{R} \hat{W}}{\hat{W}^H \mathbf{R} \hat{W}} + \frac{\sigma_\eta^2 Tr(\mathbf{R})}{\hat{W}^H \mathbf{R} \hat{W}} \tag{4.14}$$

Comparing (4.14) to (2.53) we can observe that the additional term in (4.14) represents the effect of weight quantisation. This additional term only represents the effect of the nominal virtual weight quantisation. It does not represent the quantisation effects that take place in the gradient estimation. Due to this, the misadjustment in the presence of quantisation cannot be simply derived by adding the additional term in (4.14) to the expressions derived in Chapter 3. To derive the misadjustment the first term on the right hand side of (4.14) must be calculated. This term depends on the covariance of the gradient estimate.

To calculate the first term on the right hand side of (4.14) we use the Bounds approach. The analysis is contained in Appendix E. As the derivation is similar to the analysis in Chapter 3 we do not perform the convergence analysis of the norm of the weight error vector. We assume that for a suitably dimensioned system, the additional terms in the gradient covariance due to quantisation can be made sufficiently small such that they have no effect on the convergence. We thus assume that the conditions for convergence established in Chapter 3 are true here. In Appendix E we express the gradient covariance in terms of the weight error vector and using the recursive equation of the norm of the weight error vector determine the converged bounds on $V^H(n)\mathbf{R}V(n)$. The first term on the right hand side of (4.14) can be determined from the converged bounds.

Following are the asymptotic misadjustments of the three receiver systems. Note that in the results λ_{max} denotes the maximum eigenvalue of $\mathbf{R}$ and we have not listed quantisation terms with a order higher than σ_η^2.

Result 4.4.d. Misadjustment of the Dual Receiver Dual Perturbation System

$$M_1 = \frac{\mu(L-1)}{2L}Tr(\mathbf{PRP}) + \frac{\mu\sigma_\eta^2}{4\hat{W}^H\mathbf{R}\hat{W}}\frac{(L-1)}{L}\{Tr(\mathbf{R})Tr(\mathbf{PRP}) + Tr(\mathbf{PR}^2\mathbf{P})\} + \frac{\mu\sigma_\eta^2}{8\gamma^2 L}(L-1)\left\{\frac{\hat{W}^H\mathbf{R}^2\hat{W}}{\hat{W}^H\mathbf{R}\hat{W}} + Tr(\mathbf{R})\right\} + \frac{\sigma_\eta^2 Tr(\mathbf{R})}{\hat{W}^H\mathbf{R}\hat{W}} \quad \textit{for a 4L length sequence} \tag{4.15}$$

❑❑❑

Result 4.4.e. Misadjustment of the Dual Receiver Reference Receiver System

$$M_2 = \frac{\mu(L-1)}{2L}\left\{Tr(\mathbf{PRP}) + \frac{\gamma^2(Tr(\mathbf{PRP}))^2}{\hat{W}^H\mathbf{R}\hat{W}}\right\} + \frac{\mu\sigma_\eta^2}{2\hat{W}^H\mathbf{R}\hat{W}}\frac{(L-1)}{L}\{Tr(\mathbf{R})Tr(\mathbf{PRP}) + Tr(\mathbf{PR}^2\mathbf{P})\} + \frac{\mu\sigma_\eta^2}{2\gamma^2 L}(L-1)\left\{\frac{\hat{W}^H\mathbf{R}^2\hat{W}}{\hat{W}^H\mathbf{R}\hat{W}} + Tr(\mathbf{R})\right\} + \frac{\sigma_\eta^2 Tr(\mathbf{R})}{\hat{W}^H\mathbf{R}\hat{W}} \quad \textit{for a 4L length sequence} \tag{4.16}$$

❑❑❑

Result 4.4.f. Misadjustment of the Single Receiver System

$$M_3 = \frac{a\mu(L-1)}{4L}(Tr(\mathbf{PRP})) + \frac{\mu\sigma_\eta^2}{2\hat{W}^H\mathbf{R}\hat{W}}\frac{(L-1)}{L}\{Tr(\mathbf{R})Tr(\mathbf{PRP}) + Tr(\mathbf{PR}^2\mathbf{P})\}$$

$$+\frac{\mu\sigma_\eta^2}{4\gamma^2 L}(L-1)\left\{\frac{\hat{W}^H\mathbf{R}^2\hat{W}}{\hat{W}^H\mathbf{R}\hat{W}}+Tr(\mathbf{R})\right\}+\frac{\sigma_\eta^2 Tr(\mathbf{R})}{\hat{W}^H\mathbf{R}\hat{W}} \textit{ for a 4L length sequence} \quad (4.17)$$

❑❑❑

In (4.15), (4.16) and (4.17) the gradient covariance used in the derivation is given by (4.9), (4.10) and (4.12) respectively.

Examining (4.15), (4.16) and (4.17) and comparing them to the equivalent expressions developed in Chapter 3 the following observations can be made:

- All the additional misadjustment terms are proportional to the variance of the quantisation errors. The effect of quantisation can be limited by making the variance of quantisation errors small.
- The additional misadjustment terms are proportional to the number of array elements. Therefore it can be expected that as the number of array elements is increased the misadjustment will increase.
- Some of the additional terms are inversely proportional to γ, indicating that the quantisation process favours a larger perturbation step size.
- In the misadjustment expressions without quantisation a property of the projected perturbation sequence was to remove some of the dependence the misadjustment has on the desired signal power. In the above misadjustment expressions this dependence is re-introduced, to a degree.

The accuracy of these observations are examined in the simulation studies section.

4.5 Simulation Studies

In this section we present simulation study results. Simulations were performed using the same methods as described in Chapter 3. All arrays were considered to be linear with equally spaced omnidirectional array elements with a quarter wavelength spacing at a nominal operating frequency.

In the Figures the signal scenarios are described. We assume the signals are coplanar such that they arrive in the xy plane where $\theta = 90$ and the elements of the array lie along the x axis. The direction of arrival of the signals are described in terms of the parameter ϕ as defined in Figure 1 and expressed in degrees.

White noise power of 10^{-4} relative to a unit reference power was added to each array element and the powers of each signal are defined relative to the same unit reference power used to set the white noise level. The power of each signal is also described in the figures.

In all simulations unless otherwise stated the magnitude of a perturbation step was adjusted such that it was larger than the quantisation step size. It was not possible to specify a minimum size quantisation level as for each different scenario the projection matrix can change and this will influence the size of the vector components in the perturbation sequence.

In the simulations we examine the mean weight quantisation error, weight range effects, the array's interference rejection capability and the misadjustment. The following results were obtained from the simulations.

4.5.1 Weight Quantisation Error

The sensitivity of Assumption 4.3, was first tested to determine whether the weight quantisation errors that occurred during a perturbation cycle had zero mean. As to be expected this was not always the case. The average quantisation error over a perturbation cycle depended on the length of the perturbation sequence and the number of quantisation levels used.

4.5.2 Quantisation Method 1 vs Method 2

The difference in performance of the two weight quantisation methods, Method 1 and Method 2 as discussed in Section 4.3 was examined

Using Method 1 it was observed that for a single linear constraint system, when there was coarse quantisation such that the quantisation increment is of the order of the minor components of the perturbation vector, the projected Time Multiplex sequence reverted back to a scaled Time Multiplex sequence.The amount of scaling depended on the coarseness of the quantisation process.

The cause of this reversion can be observed by examining the components of a typical projected perturbation vector which are given by the following

$$\begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \begin{bmatrix} V_1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} P_{11}V_1 \\ P_{21}V_1 \\ P_{31}V_1 \end{bmatrix} \text{ where } P_{ii} = \frac{L-1}{L},\ P_{ij} = \frac{-e^{j(\tau_1 - \tau_2)}}{L} \quad (4.18)$$

Here, P_{ij} is a component of the projection matrix and V_l is a component of a typical Time Multiplex perturbation vector.

From (4.18) it can be observed that the projected perturbation vectors will contain one component which is dominant. It was observed that when the quantisation becomes coarser, the less significant components in the projected perturbation vector are rounded down to zero, and the projected Time Multiplex sequence begins to look like a scaled version of the original Time Multiplex sequence. This effect does not necessarily occur when there is more than one constraint placed on the system.

Under this coarse quantisation, one would expect to see a rise in the misadjustment due to the re-introduction of the perturbation noise in the look direction since the system is effectively operating with a non projected sequence. This is not always the case since the major components which are left in the quantised vector may be scaled up or down. This scaling corresponds to the scaling of the original perturbation step size in a non-projected system. The observed misadjustment may then be decreased due to a reduced perturbation step size γ.

When the system does not operate under coarse quantisation similar convergence times and misadjustments were obtained.

4.5.3 Weight Range Effects

In all the simulations, the dynamic range of the weights was determined by calculating the optimum weights and allowing for an expected misadjustment.

The convergence of the weight update algorithm appeared to be guaranteed when the initial weights were set close to the optimal weights. When this occurred, it was found that under differing scenarios the output power and the optimum output power as defined by (2.26) were in good agreement. The gain in the look direction during the convergence of the algorithm also remained constant.

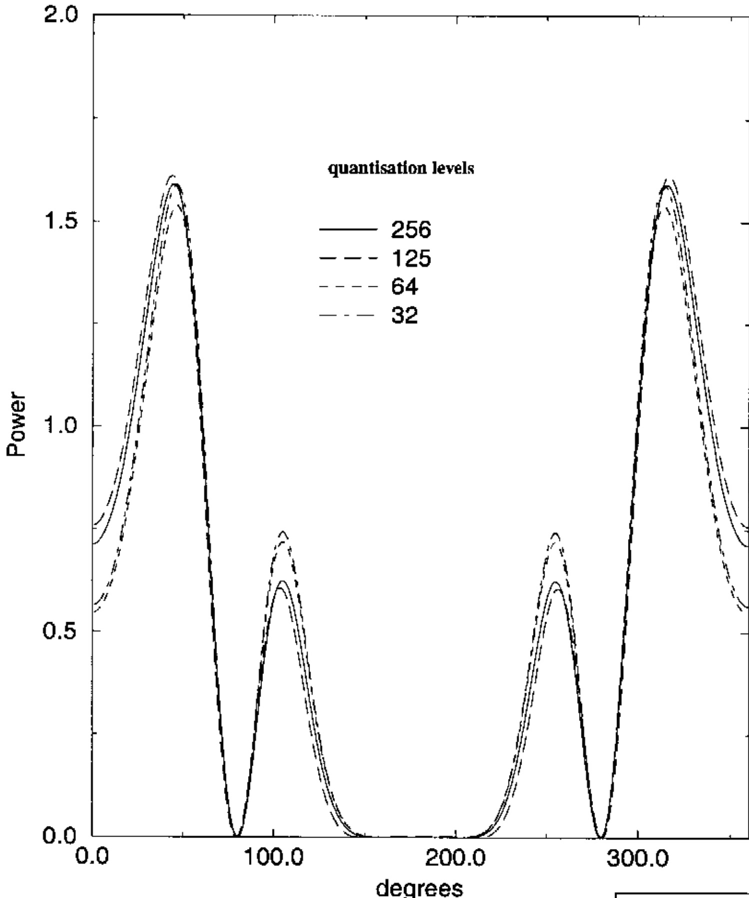
It was also observed that at times when the initial weights were not set close to the optimum weights the algorithm did not converge to the optimum weights. This was due to weight saturation. This effect could cause the final misadjustment to become smaller than the ideal predicted as not all the system constraints were satisfied.

This indicated that it was not appropriate to calculate the dynamic range of the weights by simply extrapolating from the optimum weights and allowing for a predicted misadjustment. Instead an additional buffer was required to stop the weights overflowing. This was necessary when the convergence of the algorithm occurs in a underdamped sense in which the weight values overshoot their final values.

For a fixed number of quantisation levels and a varying quantisation error variance, (i.e. the dynamic range of the weights is not fixed, the dynamic range of the weights is proportional to the optimum weights), by increasing the number of array elements L, the misadjustment at a larger L tended to be smaller than expected. This occurs since when L increases the dynamic range of the weights and the quantisation step size decreased which may favour minimizing power. Also as the number of array elements increase the number of degrees of freedom the array possess increases which aided in interference nulling.

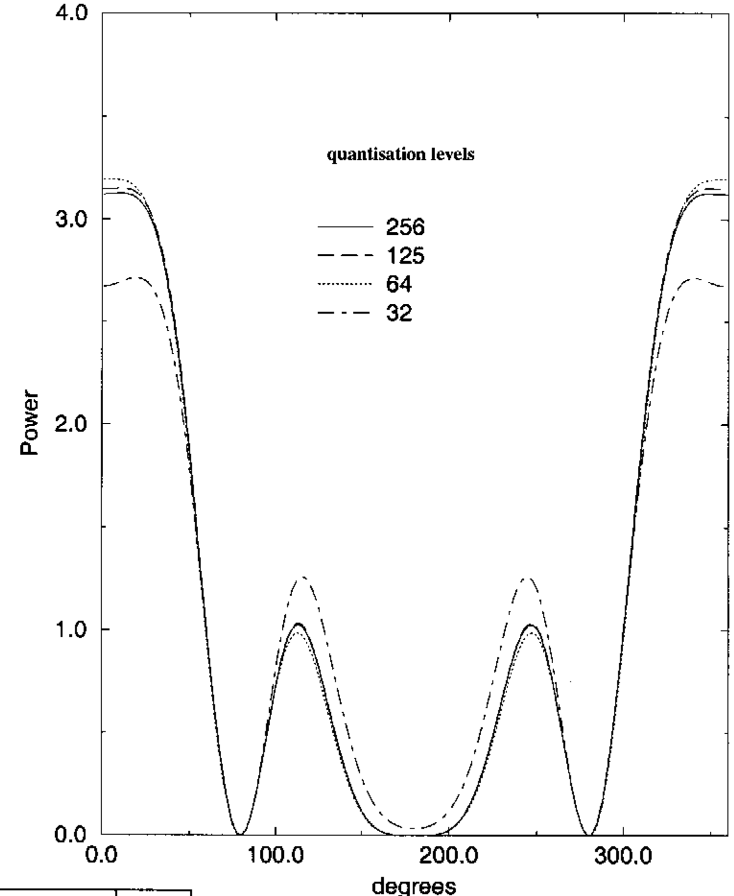
4.5.4 Interference Rejection Capability

As discussed in Section 4.4.1, we expect that with weight quantisation the gain in the interference signal directions may increase and the gain in the desired signal directions decrease for smaller number of quantisation levels. In Figure 16, we show the optimum power response to a signal of unity power for a 4 and 6 element linear array as the number of weight quantisation levels vary. In the figure "degrees" represents the direction of arrival (ϕ as defined in Figure 1) of the unity signal. The same dynamic range of the weights is used for each array and the signal scenario is described in the figure. From the figure we can observe that the 4 element array's interference rejection capability is more sensitive to the number of quantisation levels. And in general for increasing number of quantisation levels the interference rejection capability improves when the hardware weights realisation is more accurate.



6 element array

Signal	Direction of Arrival ϕ	Power
Desired	60	0.1
Interference 1	80	0.5
Interference 2	180	0.5



4 element array

Figure 16 Optimum Power Response

4.5.5 Misadjustment

In this section we present the simulated misadjustment results for the three array structures.

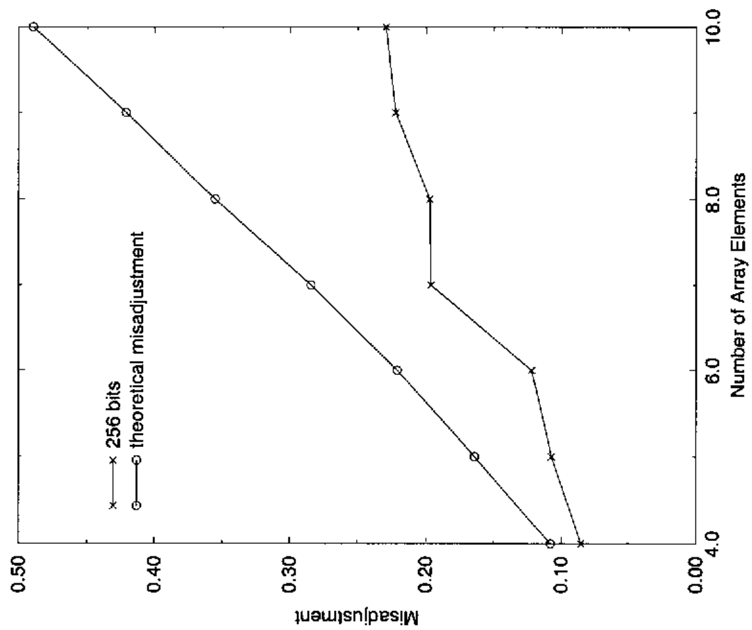
For the results presented here the quantisation variance is fixed by fixing the dynamic range of the weights and number of quantisation levels. During the simulations it was necessary to check that the narrowband assumption defined by (2.8) was still being satisfied since every time an element is added to the array, the size of the array increased.

Figure 17 shows the misadjustment for the three receiver structures for different numbers of quantisation levels as the number of array elements vary. The misadjustment for the non quantised case and the theoretical misadjustment as given by (4.15), (4.16) and (4.17) are also shown. These simulations have been performed with a gradient step size of 0.05 and with a time invariant optimum perturbation step size given by (3.20).

Figure 17 (i) and (ii) shows the results for the dual receiver dual perturbation system. In these simulations a $2L$ length projected time multiplex sequence is used. Figure 17 (iii) and (iv) shows the results for the dual receiver reference receiver system and Figure 17 (iv) shows the results for the single receiver system.

For a fixed signal and quantisation variance scenario, it is difficult to directly compare misadjustment simulation results as the number of array elements increase. This is due to the fact that for a fixed quantisation variance, as the number of array elements increase the number of degrees of freedom the array possesses increases which changes its interference nulling capability. Also the fixed quantisation variance may favour a particular array size. So it is possible that the optimum weight representation at say L_1 will be more accurate than the optimum weight representation at L_2. For example, as L increases the size of the optimum weights decrease and for a fixed number of quantisation levels it may be more difficult to accurately represent these optimum weights compare to the optimum weights at a smaller L.

From Figure 17 it is possible to observe that there is a definite increase in the misadjustment as the number of array elements increase and that whenever quantisation is applied an increase in misadjustment results. The theoretical misadjustment as given by (4.15), (4.16) and (4.17) form an upper bound to the actual/simulated misadjustment. The expected increase in misadjustment as predicted from the expressions does not occur due in part to the reasons mentioned previously and due to the quantisation model assumptions.



Signal	Direction of Arrival φ	Power
Desired	60	0.1
Interference 1	80	0.5
Interference 2	180	0.5

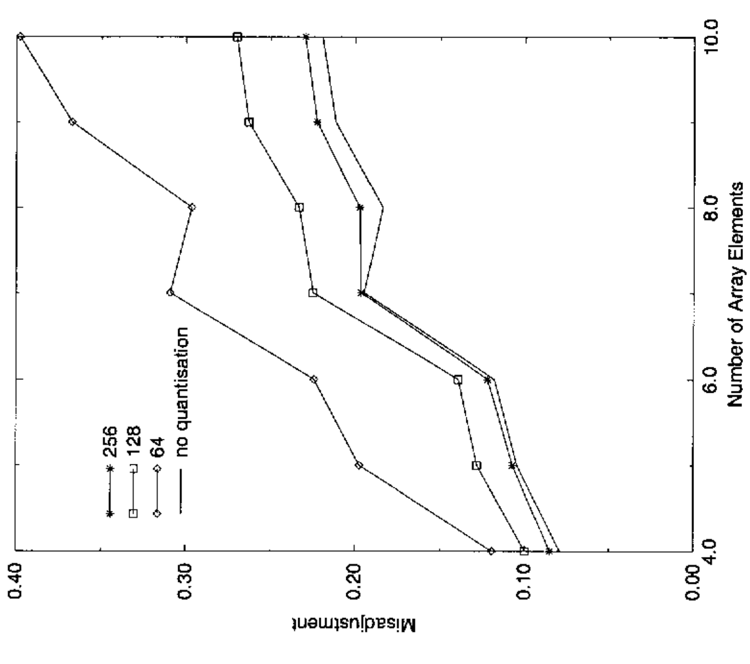
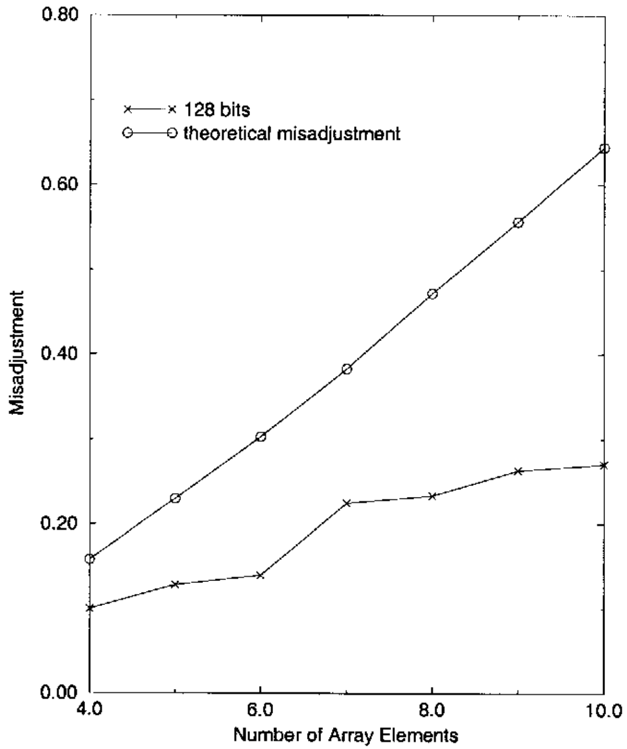
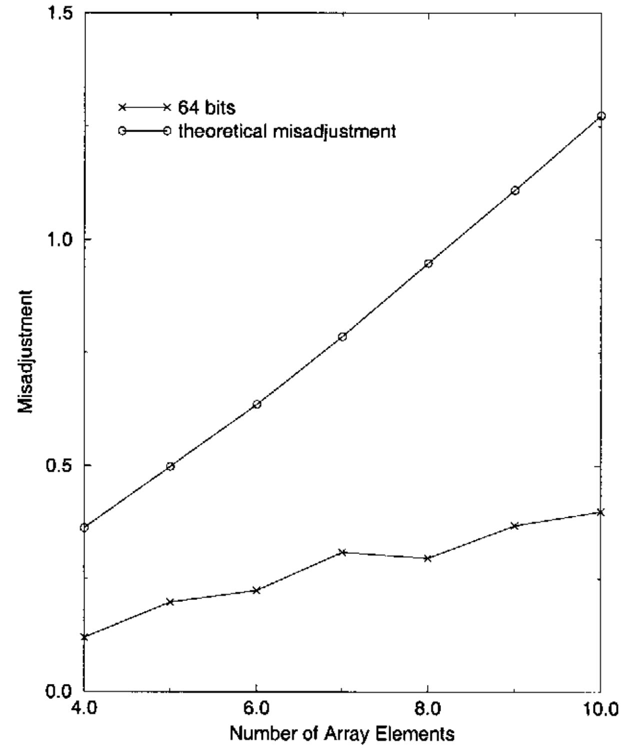


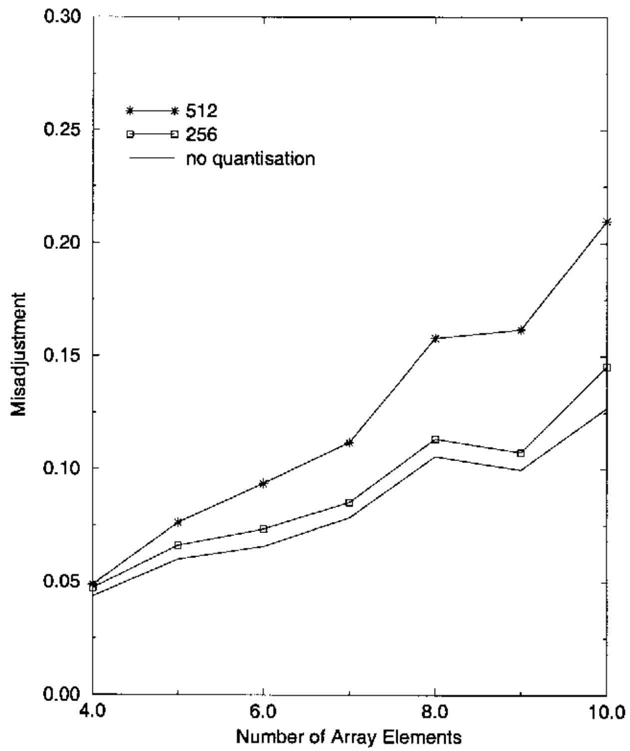
Figure 17 (i) Misadjustment, Dual Receiver Dual Perturbation System

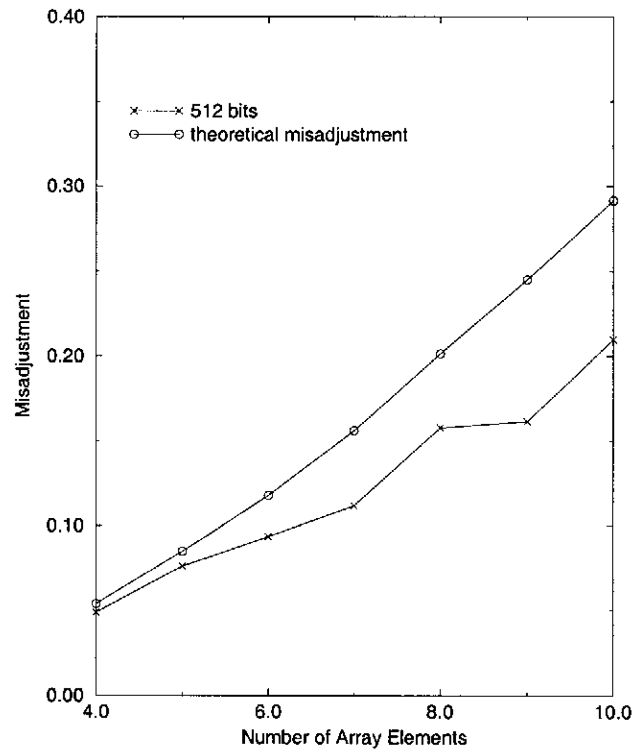




Signal	Direction of Arrival ϕ	Power
Desired	60	0.1
Interference 1	80	0.5
Interference 2	180	0.5

Figure 17 (ii) Misadjustment, Dual Receiver Dual Perturbation System





Signal	Direction of Arrival ϕ	Power
Desired	60	0.1
Interference 1	80	0.5
Interference 2	180	0.5

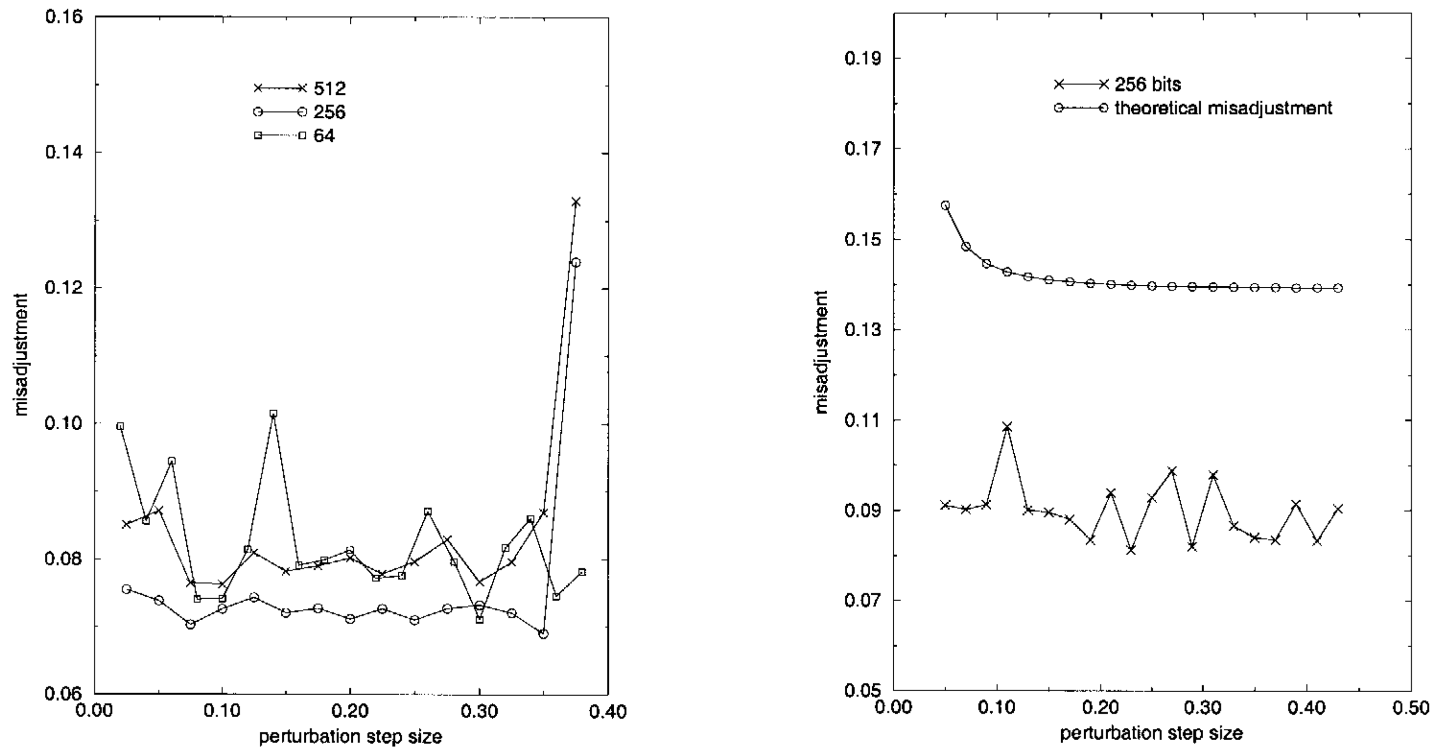
Figure 17 (iii) Misadjustment, Dual Receiver Reference Receiver System

Figure 18 shows the misadjustment for the three receiver structures for different numbers of quantisation levels as the perturbation step size varies. The misadjustment for the non quantised case is also shown.

In the dual receiver dual perturbation case, the misadjustment is approximately constant for a range of perturbation step sizes, the misadjustment increasing when the perturbation step size exceeds some maximum γ_{quan}. The rise in misadjustment when the perturbation step size exceeds γ_{quan} is due to the weights being forced outside the dynamic range of the weights.

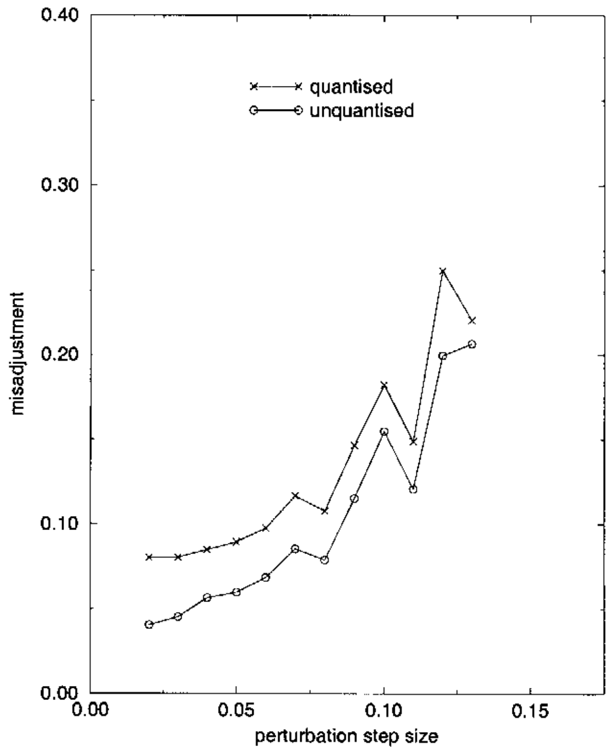
In the dual receiver reference receiver case the misadjustment is a monotonically increasing function of γ and in the single receiver case the misadjustment is a convex function.

In each of the receiver structures the misadjustment increased as predicted and agreed with the literature. The misadjustment also increased from the ideal case, but again not to the degree to that predicted by the derived expressions. As shown in Figure 18 (i) the theoretical misadjustment forms an upper bound to the actual misadjustment. Similar results were obtained for the other receiver structures.



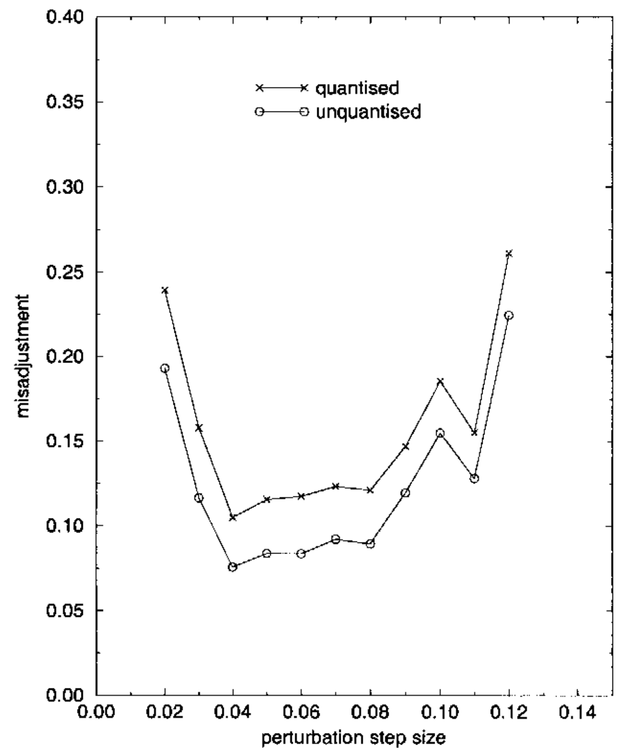
Signal	Direction of Arrival ϕ	Power
Desired	60	0.01
Interference 1	80	0.5
Interference 2	180	0.5

Figure 18 (i) Misadjustment, Dual Receiver Dual Perturbation System



Signal	Direction of Arrival ϕ	Power
Desired	60	0.01
Interference 1	80	0.5
Interference 2	180	0.5

Dual Receiver Reference Receiver System



Signal	Direction of Arrival ϕ	Power
Desired	60	0.01
Interference 1	80	0.5
Interference 2	180	0.5

Single Receiver System

Figure 18 (ii) Misadjustment

4.6 Summary

From this section one was able to observe that when the perturbation vectors can be accurately represented under quantisation, the implementation of the projected perturbation algorithm is generally well behaved and performs predictably as in the non projected case. The expressions developed in this section provide some insight into how an array's performance reacts to different number of quantisation levels, the number of array elements and the size of the perturbation step size. Though the modelling of the quantisation errors needs refinement.

Two criteria have been used to examine the performance of an adaptive array with digital implementation effects, they are the interference rejection capability of the processor and the misadjustment. With these criteria it is difficult to conclude what are the minimum wordlength requirements to implement a projected perturbation sequence. The difficulty arises as it is a scenario dependent problem and also because the wordlength requirements determined to give a desired performance may not be sufficient to accurately represent the perturbation sequence.

Chapter 5

Application of Spatial Derivative Constraints to Narrowband Adaptive Array Processing

5.1 Introduction

The projected perturbation approach in adaptive beamforming is used to eliminate the effect of perturbation noise during adaptation. However when there is a directional mismatch between desired signal's direction of arrival and the look direction, projection of the perturbation sequence onto the look direction constraint plane does not necessarily eliminate the perturbation noise in the output signal as reviewed earlier in Section 2.10.4. Under mismatched conditions the reception of the desired signal may also be degraded and the array may treat it as an unwanted interference signal.

The application of spatial derivative constraints or directional constraints to an array processor weights can be used to improve the reception of the desired signal when it is offset from the look direction [9], [13], [18], [28], [29], [35], [42]. Therefore, it can be expected that projection of the perturbation sequence onto these additional constraint planes can also be used to remove the perturbation noise in the output signal when directional mismatch occurs. Alternatively, when directional mismatch is not a concern, projection of the perturbation vectors onto the additional constraint planes can be used to reduce the perturbation noise contributed by signals within an angular region of the look direction.

A contribution of this thesis is that we examine the benefit of projecting the perturbation sequence onto the spatial derivative constraint planes. Little work has appeared about the application of spatial derivative constraints to narrowband perturbation based processors. We develop new expressions for an array's sensitivity to perturbation noise, and identify conditions on the array under which spatial derivative constraints applied to perturbation sequences are effective in suppressing perturbation noise. Applying the spatial derivative constraints to the perturbation sequence can unnecessarily increase the complexity of implementing the sequence.

We also briefly examine these implementation issues.

This Chapter is organised as follows. In Section 5.2 we review the derivation of the first and second order spatial derivative constraints and evaluate them for a linear array. In Section 5.3, we examine the adaptive beamforming problem with spatial derivative constraints. In Section 5.4 by examining the array's sensitivity to perturbation noise, the norm of the spatial derivative constraint and the array's directional sensitivity, we determine conditions under which spatial derivatives could be applied effectively to the perturbation vectors. Results of simulation studies are presented in Section 5.5.

5.2 Spatial Derivative Constraints

In this section, the derivation of the spatial derivative constraints of a narrowband processor are reviewed. Initially we present the generic power response derivative constraints and the frequency response spatial derivative constraints and evaluate them for a linear array. The optimisation problem with spatial derivative constraints is then presented.

5.2.1 Derivation of Spatial Derivative Constraints

In general, the n^{th} order spatial derivative constraints are derived by examining the conditions under which the n^{th} order spatial derivative of the array's power response or frequency response in an angular direction is zero. Here the spatial derivatives are taken with respect to the angular co-ordinates defined in Chapter 2 but other techniques vary in this respect, such as in [13], where the spatial derivatives are taken with respect to the sine of the angular co-ordinates. The zeroth order constraint corresponds to the look direction constraint defined by (2.56).

The power response of the array processor to a signal arriving from direction, (θ, ϕ), is defined by

$$\rho(\theta, \phi) = H(\theta, \phi)H^H(\theta, \phi) \tag{5.1}$$

where

$$H(\theta, \phi) = W^H(n)S(\theta, \phi) \tag{5.2}$$

is the frequency response to a signal arriving from direction (θ, ϕ).

Following are the first and second order power response derivative constraints. These constraints are necessary and sufficient constraints and their derivation can be found in [51].

First Order Power Response Derivative Constraints

$$Re\{W^H S_\theta(\theta, \phi)\} = 0, Re\{W^H S_\phi(\theta, \phi)\} = 0, Re\{W^H S_f(\theta, \phi)\} = 0 \quad (5.3)$$

Second Order Power Response Derivative Constraints

$$Re\{W^H S_{\theta\theta}(\theta, \phi)\} = -|Im\{W^H S_\theta(\theta, \phi)\}|^2$$

$$Re\{W^H S_{\phi\phi}(\theta, \phi)\} = -|Im\{W^H S_\phi(\theta, \phi)\}|^2$$

$$Re\{W^H S_{ff}(\theta, \phi)\} = -|Im\{W^H S_f(\theta, \phi)\}|^2$$

$$Re\{W^H S_{\theta\phi}(\theta, \phi)\} = -Im\{W^H S_\theta(\theta, \phi)\}Im\{W^H S_\phi(\theta, \phi)\}$$

$$Re\{W^H S_{\phi f}(\theta, \phi)\} = -Im\{W^H S_\phi(\theta, \phi)\}Im\{W^H S_f(\theta, \phi)\}$$

$$Re\{W^H S_{f\theta}(\theta, \phi)\} = -Im\{W^H S_f(\theta, \phi)\}Im\{W^H S_\theta(\theta, \phi)\} \quad (5.4)$$

where $S_\alpha(\theta, \phi) = \frac{\partial}{\partial\alpha} S(\theta, \phi)\Big|_{(\theta_o, \phi_o, f_o)}$ (5.5)

$$S_{\alpha\beta}(\theta, \phi) = \frac{\partial^2}{\partial\alpha\partial\beta} S(\theta, \phi)\Big|_{(\theta_o, \phi_o, f_o)} \quad (5.6)$$

Following are the first and second order frequency response spatial derivative constraints. These constraints are sufficient constraints and their derivation can be found in [35]. Note that for simplicity of notation throughout the rest of the chapter we use $S = S(\theta_0, \phi_0)$.

First Order Frequency Response Spatial Derivative Constraints

$$W^H(n)\frac{\partial\Lambda_\tau}{\partial\theta}\Big|_{(\theta_o, \phi_o)} S = 0 \quad (5.7)$$

$$W^H(n)\frac{\partial\Lambda_\tau}{\partial\phi}\Big|_{(\theta_o, \phi_o)} S = 0 \quad (5.8)$$

Second Order Frequency Response Spatial Derivative Constraints

$$W^H(n)\left(\frac{\partial\Lambda_\tau}{\partial\theta}\Big|_{(\theta_o, \phi_o)}\right)^2 S = 0 \text{ and } W^H(n)\frac{\partial^2\Lambda_\tau}{\partial\theta^2}\Big|_{(\theta_o, \phi_o)} S = 0 \quad (5.9)$$

$$W^H(n)\left(\frac{\partial\Lambda_\tau}{\partial\phi}\Big|_{(\theta_o, \phi_o)}\right)^2 S = 0 \text{ and } W^H(n)\frac{\partial^2\Lambda_\tau}{\partial\phi^2}\Big|_{(\theta_o, \phi_o)} S = 0 \quad (5.10)$$

$$W^H(n)\frac{\partial\Lambda_\tau}{\partial\phi}\bigg|_{(\theta_o,\,\phi_o)}\frac{\partial\Lambda_\tau}{\partial\theta}\bigg|_{(\theta_o,\,\phi_o)}S = 0 \quad \text{and } W^H(n)\frac{\partial\Lambda_\tau}{\partial\theta\partial\phi}\bigg|_{(\theta_o,\,\phi_o)}S = 0 \tag{5.11}$$

where Λ_τ is an $L \times L$ diagonal matrix defined by

$$\Lambda_\tau = Diag[\tau_1(\theta, \phi), \tau_2(\theta, \phi), \ldots\ldots, \tau_L(\theta, \phi)] \tag{5.12}$$

Note that in the derivation of the higher order constraints it is always assumed that lower order constraints are satisfied.

Examining the expressions for the power response derivative constraints and the frequency response spatial derivative constraints the following observations can be made.

- The power response derivative constraints are phase independent and non-linear.
- The frequency response spatial derivative constraints are linear but phase dependent. That is, the derivative constraints depend on the location of the phase centre [18], [28], [51], [52]. As a unique solution exists for the optimum weight vector when the constraints are linearly independent, the optimum weight vector is different for different phase centres.
- The frequency response spatial derivatives constraints can be obtained by removing the *Re* and *Im* complex operators in the power response spatial derivative constraints and setting each component to zero.
- A potential drawback of the sufficient frequency response spatial derivative constraints is that they are more restrictive when compared to the necessary and sufficient power response spatial derivative constraints [18], [51].

Because of their simpler form we only consider the frequency response spatial derivative constraints. For convenience in the rest of the thesis we refer to the frequency response spatial derivative constraints as spatial derivative constraints.

5.2.2 *Evaluation of Spatial Derivative Constraints for Linear Arrays*

Here we evaluate the first and second order spatial derivative constraints for a linear array having L equally spaced elements. Furthermore, signals are considered to arrive in the xy plane where $\theta = 90^o$, see Figure 1. The elements of the array are assumed to lie along the x axis and are symmetrical about the origin. Hence for each element of the array $y_i = z_i = 0$.

The look direction steering vector is now given by

$$S^T = \left[e^{j2\pi f_o x_1 \frac{\cos(\phi_o)}{\upsilon}}, e^{j2\pi f_o x_2 \frac{\cos(\phi_o)}{\upsilon}}, \ldots, e^{j2\pi f_o x_L \frac{\cos(\phi_o)}{\upsilon}} \right] \tag{5.13}$$

and the first and second order constraints reduce to:

Result 5.2.c. First Order Spatial Derivative Constraints, Linear Array

$$W^H(n)\Lambda_X S \sin(\phi_o) = 0 \tag{5.14}$$

where Λ_X is an $L \times L$ diagonal matrix defined by

$$\Lambda_X = Diag[x_1, x_2, \ldots\ldots, x_L] \tag{5.15}$$

❑❑❑

Result 5.2.d. Second Order Spatial Derivative Constraints, Linear Array

$$W^H(n)\Lambda_X^2 S \sin^2(\phi_o) = 0 \tag{5.16}$$

$$W^H(n)\Lambda_X^2 S \cos(\phi_o) = 0 \tag{5.17}$$

where Λ_X^2 is an $L \times L$ diagonal matrix defined by

$$\Lambda_X^2 = Diag[x_1^2, x_2^2, \ldots\ldots, x_L^2] \tag{5.18}$$

❑❑❑

5.3 The Adaptive Array Processor with Spatial Derivative Constraints

In this section we re-examine the array optimisation problem and the adaptive beamforming problem with spatial derivative constraints.

When the spatial derivative constraints are applied to the array, the optimisation problem is a multiple linear constraint problem. For the linear array we summarise the

First and *Second* order multiple linear constraint optimisation problems. The Zeroth order constraint optimisation problem corresponds to the single look direction optimisation problem.

First Order Optimisation Problem.

The first order optimisation problem is defined by

$$\begin{aligned} &\min_{W} \; P(W) \\ \text{subject to} \;\; &W^H(n)S = 1 \\ &W^H(n)\Lambda_X S = 0 \end{aligned} \tag{5.19}$$

where $P(W)$ is the output power defined by (2.17).

Second Order Optimisation Problem.

The second order optimisation problem is defined by

$$\begin{aligned} &\min_{W} \; P(W) \\ \text{subject to} \;\; &W^H(n)S = 1 \\ &W^H(n)\Lambda_X S = 0 \\ &W^H(n)\Lambda_X^2 S = 0 \end{aligned} \tag{5.20}$$

It is interesting to note that when we assume that the co-ordinate system origin is chosen at the phase centre of the array, the zeroth order and first order constraints are mutually orthogonal. Also the first order and the second order constraints are mutually orthogonal. However the zeroth and second order constraints are not necessarily orthogonal

The problems specified by (5.19) and (5.20) can be expressed equivalently as the linear constrained optimization problem.

$$\begin{aligned} &\min_{W} \; P(W) \\ \text{subject to} \;\; &W^H(n)C = F \end{aligned} \tag{5.21}$$

where C is $L \times 2$ or $L \times 3$ constraint matrix defined by

$$C = \begin{bmatrix} C_0 \; C_1 \; \dots \; C_2 \end{bmatrix} \tag{5.22}$$

and F is a $L \times 2$ or $L \times 3$ vector

$$F = \begin{bmatrix} 1 \; 0 \; \dots \; 0 \end{bmatrix} \tag{5.23}$$

The columns of C, i.e. C_0, C_1, C_2 correspond to S, $\Lambda_X S$ and $\Lambda_X^2 S$ respectively.

It is interesting to note that for the optimum weight vector satisfying only the zeroth order constraint the following are true.

$$Re(\hat{W}^H \Lambda_X S) = 0$$

and

$$Im(\hat{W}^H \Lambda_X^2 S) = 0$$

As discussed in Chapter 2, since **R** is unknown and must be estimated from the data to obtain the optimum weight vector $\hat{W}$ given by (2.25), a real time algorithm for determining $\hat{W}$ must be used. When C has full rank a real time algorithm for obtaining the optimum weight vector $\hat{W}$ is given by

$$W(n+1) = \mathbf{P}[W(n) - \mu G(W(n))] + \mathbf{C}(\mathbf{C}^H \mathbf{C})^{-1} F^H \tag{5.24}$$

In (5.24) the projection matrix **P** is not the same as the projection matrix for a single linear constraint problem, (2.27), since there is now more than one constraint placed on the system. When the linear constraints placed on the system are mutually orthogonal, [9], i.e. C is full rank, the projection matrix in (5.24) can be represented as

$$\mathbf{P} = \mathbf{P}_0 \mathbf{P}_1 \mathbf{P}_2 \ldots \text{ where } \mathbf{P}_i = \left[\mathbf{I}_{LL} - \frac{C_i C_i^H}{C_i^H C_i}\right] \tag{5.25}$$

Note that the projection matrix for projecting the perturbation vectors, defined here as $\mathbf{P}_s$, is not necessarily the same as that in (5.24) for the weight update algorithm. The selection of $\mathbf{P}_s$ is studied in the next section. It is possible to have a **P** of a higher order than $\mathbf{P}_s$. That is, the number of constraint planes the array weights are projected onto is greater than the number of constraint planes the perturbation vectors are projected onto. This will occur when it is not necessary to project the perturbation vectors onto all constraint planes to counter the mismatch conditions and to effect the desired perturbation noise response. Alternatively, it is also possible for $\mathbf{P}_s$ to be of a higher order than **P** although it is envisaged there is no application for this.

5.4 Spatial Derivative Constraints- Performance and Implementation Issues

In general the cost and complexity involved in generating and implementing projected sequences must be weighed against the array's improved performance. In this section, to determine the benefit of projecting the perturbation sequence onto additional spatial derivative constraint planes we examine an array's perturbation noise performance and some implementation issues.

Firstly we develop new expressions for the array's sensitivity to perturbation noise. The expressions give the array's perturbation noise sensitivity in relation to the signal's offset from the look direction, the number of array elements and the desired signal power. The expressions are developed for a linear array for the zeroth order and zeroth plus first order projections. A similar analysis can be applied to other array configurations but the expressions developed for more complicated arrays are complex and don't give an insight to an array's sensitivity to perturbation noise. For the linear array we do not develop expression for the zeroth plus first plus second order projections as it is shown later in the simulation results that the beam width for perturbation noise suppression gained by the use of the zeroth order constraint or the zeroth plus first order constraint was sufficient for narrowband applications.

Next we examine the effect the additional spatial derivative projections have on the perturbation vector size in implementation. Finally we consider the array's sensitivity to directional errors. For a linear array we identify some simple scenarios which are insensitive to small look direction errors.

5.4.1 Perturbation Noise with Zeroth Order Constraint

The main reason for projecting the perturbation sequences onto the spatial derivative constraint planes is to suppress perturbation noise. The improvement in the array's performance can be determined by comparing how much perturbation noise would be present in the system without and with the use of projected perturbations. The perturbation noise in the desired signal direction for a zeroth order constraint system and a first order constraint system will now be determined.

For a zero mean projected perturbation sequence the excess power due to perturbations about a nominal weight W is given by [4],

$$\xi = 2\gamma^2 Tr(\mathbf{PR}) \tag{5.26}$$

and the component of this noise due to the desired signal is given by

$$\xi_{ds} = 2\gamma^2 Tr(\mathbf{PR}_{ds}) \tag{5.27}$$

where $\mathbf{R}_{ds}$ is the correlation matrix due to the desired signal's component. Substituting for $\mathbf{R}_{ds}$

$$\xi_{ds} = 2\gamma^2 p_{ds} S_{ds}^H \mathbf{P} S_{ds} \tag{5.28}$$

where S_{ds} is the steering vector in the desired signal direction and p_{ds} is the desired signal's power. For the linear array,

$$S_{ds}^T = \left[e^{j2\pi f_o x_1 \frac{\cos(\phi_o + \Delta)}{\upsilon}}, e^{j2\pi f_o x_2 \frac{\cos(\phi_o + \Delta)}{\upsilon}}, \ldots\ldots\ldots\ldots, e^{j2\pi f_o x_L \frac{\cos(\phi_o + \Delta)}{\upsilon}} \right] \tag{5.29}$$

where Δ is the angular deviation the desired signal arrives from the look direction, it is given by

$$\Delta = \phi_o - \phi_{ds} \tag{5.30}$$

where ϕ_{ds} is the direction of arrival of the desired signal.

When the perturbation sequence is only projected onto the zeroth order constraint plane, it can be shown that by substituting the projection matrix for the zeroth order constraint, $\mathbf{P}_0$, into (5.28), that the excess power due to perturbations contributed by the desired signal is given by

$$\xi_{Zds} = 2\gamma^2 p_{ds} S_{ds}^H \left(\mathbf{I}_{LL} - \frac{SS^H}{L} \right) S_{ds} \tag{5.31}$$

On substitution for S_{ds} and S

$$\xi_{Zds} = 2\gamma^2 p_{ds} \left[L - \frac{1}{L} \left\{ \sum_i \cos\left(\frac{2\pi f_o x_i}{\upsilon} (\cos(\phi_o + \Delta) - \cos(\phi_o)) \right) \right\}^2 \right] \tag{5.32}$$

To determine the characteristics for small look direction errors, consider now the derivatives of the perturbation noise power ξ_{Zds} with respect to Δ

$$\frac{\partial \xi_{Zds}}{\partial \Delta} = -4\gamma^2 \frac{p_{ds}}{L} \left[\sum_i \cos\left(\frac{2\pi f_o x_i}{\upsilon} (\cos(\phi_o + \Delta) - \cos(\phi_o)) \right) \right] \times$$

$$\left[\sum_i \frac{2\pi f_o x_i}{\upsilon} \sin(\phi_o + \Delta) \sin\left(\frac{2\pi f_o x_i}{\upsilon} (\cos(\phi_o + \Delta) - \cos(\phi_o)) \right) \right] \tag{5.33}$$

$$\frac{\partial^2 \xi_{Zds}}{\partial \Delta} = -4\gamma^2 \frac{p_{ds}}{L}\Bigg[\Bigg[\sum_i \frac{2\pi f_o x_i}{\upsilon}\sin(\phi_o+\Delta)\sin\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg]^2$$

$$+\Bigg[\sum_i \cos\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg]\times$$

$$\Bigg[\sum_i\Bigg[\frac{2\pi f_o x_i}{\upsilon}\cos(\phi_o+\Delta)\sin\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg] -$$

$$\sum_i\Bigg[\Big(\frac{2\pi f_o x_i}{\upsilon}\Big)^2 \sin^2(\phi_o+\Delta)\cos\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg]\Bigg]\Bigg] \quad (5.34)$$

From (5.33) and (5.34) the following can be observed:

Observation 5.4.1.a

When $\Delta = 0$, ξ_{Zds} is at a minimum since

$$\left.\frac{\partial \xi_{Zds}}{\partial \Delta}\right|_{\Delta=0} = 0 \text{ and } \left.\frac{\partial^2 \xi_{Zds}}{\partial \Delta}\right|_{\Delta=0} = \left(\frac{4\gamma\pi f_o \sin(\phi_o)}{\upsilon}\right)^2 p_{ds}\sum_i x_i^2 > 0$$

❑❑❑

Observation 5.4.1.b

When $\phi_o = n\pi$, $n = 0,1,2,...$, since

$$\left.\frac{\partial \xi_{Zds}}{\partial \Delta}\right|_{\Delta=0} = 0 \text{ and } \left.\frac{\partial^2 \xi_{Zds}}{\partial \Delta}\right|_{\Delta=0} = 0$$

then $\Delta = 0$ is a point of inflection.

Hence it is expected that the array's perturbation noise performance degrades relatively slowly with Δ in the vicinity of these look directions.

❑❑❑

Observation 5.4.1.c

When $\phi_o = n\frac{\pi}{2}$, $n = 1,3,5,...$, $\frac{\partial^2}{\partial \Delta}\xi_{Zds}$ is at a maximum indicating that the array's perturbation noise performance will degrade most rapidly with Δ in the vicinity of these look directions.

❑❑❑

Observation 5.4.1.d

When $|\Delta|$ is sufficiently small, such that $\cos(\phi_o+\Delta) \approx \cos(\phi_o)$ for any look direction, the second order derivative of ξ_{Zds} is proportional to $\sum_i x_i^2$. Hence the rate of change of $\frac{\partial}{\partial \Delta}\xi_{Zds}$ will be proportional to the size of the array and the number of elements in the array. This indicates that for a given inter-element spacing the perturbation noise will increase as the number of elements increases. Also, for a given number of elements the perturbation noise will increase as the inter-element spacing increases.

❑❑❑

5.4.2 Perturbation Noise with the Zeroth plus First Order Constraints

The difference in excess power due to perturbations about a nominal weight $\boldsymbol{W}$ when the perturbation sequence is projected onto the zeroth order constraint plane and the zeroth plus first order constraint plane is defined by

$$\xi_D = 2\gamma^2 Tr((\mathbf{P}_0 - \mathbf{P}_0\mathbf{P}_1)\mathbf{R}) \tag{5.35}$$

where $\mathbf{P}_0$ and $\mathbf{P}_1$ are the projection matrices for the zeroth order and first order constraints respectively.

Substituting for $\mathbf{P}_0$ and $\mathbf{P}_1$ it can be shown that

$$\xi_D = 2\gamma^2 Tr\left(\frac{\Lambda_X S S^H \Lambda_X}{\sum_i x_i^2}\mathbf{R}\right) \tag{5.36}$$

The desired signal's contribution to, ξ_D, the difference in excess power is defined by

$$\xi_{Dds} = 2\gamma^2 p_{ds}\left(\frac{S_{ds}^H \Lambda_X S S^H \Lambda_X S_{ds}}{\sum_i x_i^2}\right) \tag{5.37}$$

On substitution for S_{ds} and S

$$\xi_{Dds} = 2\gamma^2 p_{ds}\left(\frac{\left(\sum_i x_i \sin\left(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o + \Delta) - \cos(\phi_o))\right)\right)^2}{\sum_i x_i^2}\right) \tag{5.38}$$

Consider now the derivatives of the difference in perturbation noise power ξ_{Dds} with respect to Δ to determine the characteristics for small look direction errors.

$$\frac{\partial \xi_{Dds}}{\partial \Delta} = \frac{-4\gamma^2 p_{ds}}{\sum_i x_i^2}\left[\sum_i x_i \sin\left(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o + \Delta) - \cos(\phi_o))\right)\right] \times$$
$$\left[\sum_i \frac{2\pi f_o x_i^2}{\upsilon}\sin(\phi_o + \Delta)\cos\left(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o + \Delta) - \cos(\phi_o))\right)\right] \tag{5.39}$$

$$\frac{\partial^2 \xi_{Dds}}{\partial \Delta} =$$

$$\frac{-4\gamma^2 p_{ds}}{\sum_i x_i^2}\Bigg[-\Bigg[\sum_i \frac{2\pi f_o x_i^2}{\upsilon}\sin(\phi_o+\Delta)\cos\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg]^2$$

$$+\Bigg[\sum_i x_i \sin\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg]\times$$

$$\Bigg[\Bigg[\sum_i \frac{2\pi f_o x_i^2}{\upsilon}\cos(\phi_o+\Delta)\cos\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg]+$$

$$\Bigg[\sum_i \frac{4\pi^2 f_o^2 x_i^3}{\upsilon^2}\sin^2(\phi_o+\Delta)\sin\Big(\frac{2\pi f_o x_i}{\upsilon}(\cos(\phi_o+\Delta)-\cos(\phi_o))\Big)\Bigg]\Bigg]\Bigg] \quad (5.40)$$

From (5.39) and (5.40) the following can be observed:

Observation 5.4.2.a

When $\Delta = 0$, ξ_{Dds} is at a minimum since

$$\left.\frac{\partial \xi_{Dds}}{\partial \Delta}\right|_{\Delta=0} = 0 \text{ and } \left.\frac{\partial^2 \xi_{Dds}}{\partial \Delta}\right|_{\Delta=0} = \left(\frac{4\gamma\pi f_o \sin(\phi_o)}{\upsilon}\right)^2 p_{ds}\sum_i x_i^2 > 0$$

❑ ❑ ❑

Observation 5.4.2.b

When $\phi_o = n\pi$, $n = 0,1,2,\ldots$, since

$$\left.\frac{\partial \xi_{Dds}}{\partial \Delta}\right|_{\Delta=0} = 0 \text{ and } \left.\frac{\partial^2 \xi_{Dds}}{\partial \Delta}\right|_{\Delta=0} = 0$$

then $\Delta = 0$ is a point of inflection.

Hence it is expected that the array's perturbation noise reduction is minimal in the vicinity of these look directions.

❑ ❑ ❑

Observation 5.4.2.c

When $\phi_o = n\frac{\pi}{2}$, $n = 1,3,5,\ldots$, $\frac{\partial^2}{\partial \Delta}\xi_{Dds}$ is at a maximum. Indicating that the rate at which the array's reduction in perturbation noise occurs, as Δ varies in the vicinity of these look directions, is at a maximum.

❑ ❑ ❑

Observation 5.4.2.d

When Δ is sufficiently small, such that, $\cos(\phi_o + \Delta) \approx \cos(\phi_o)$ for any look direction the second order derivative of ξ_{Dds} is proportional to $\sum_i x_i^2$. Hence the rate of change of $\frac{\partial \xi_{Dds}}{\partial \Delta}$ will be proportional to the size of the array and the number of elements in the array. This indicates that for a given inter-element spacing the reduction in perturbation noise will increase as the number of elements increases. Also, for a given number of elements the reduction in perturbation noise will increase as the inter-element spacing increases.

❑❑❑

When Δ is small the desired signal's contribution to ξ_D is approximately given by

$$\xi_{Dds} \approx 2\gamma^2 p_{ds} \frac{(\sum x_i \sin(2\pi f x_i \sin\phi \sin\Delta))^2}{\sum x_i^2} \tag{5.41}$$

For small $|\Delta|$, which from simulations may be up to 5^o, the value of ξ_{Dds} given by (5.41) will be relatively small and the projection of the perturbation sequence onto the first order constraint plane may be unnecessary.

When the perturbation sequence is projected onto the zeroth plus first order constraint planes, the excess power due to perturbations contributed by the desired signal is given by

$$\xi_{Fds} = \xi_{Zds} - \xi_{Dds} \tag{5.42}$$

It can be observed from (5.32) and (5.38) that the excess power due to signal's from the look direction is zero as the weights have been projected onto the zeroth order constraint plane while the excess power due to the desired signal is related to the desired signal's power p_{ds}, the angular deviation from the look direction Δ, and the size of the array. Similar observations to those made for ξ_{Zds} and ξ_{Dds} can also be made here:

Observation 5.4.2.e

When $\Delta = 0$, ξ_{Fds} is at a minimum.

❑❑❑

Observation 5.4.2.f

Since $\left.\frac{\partial^2}{\partial\Delta}\xi_{Fds}\right|_{\Delta = 0} = 0$, when $\left.\frac{\partial\xi_{Fds}}{\partial\Delta}\right|_{\Delta = 0} = 0$, $\Delta = 0$ is at a point of inflection independent of the look direction. Hence it is expected that the array's perturbation noise performance will degrade relatively slowly with Δ in the vicinity of the look direction.

❑❑❑

Observation 5.4.2.g

When Δ is sufficiently small, such that $\cos(\phi_o + \Delta) \approx \cos(\phi_o)$ the second order derivative of ξ_{Zds} and ξ_{Dds} are proportional to $\sum_i x_i^2$. Hence the rate of change of $\frac{\partial}{\partial\Delta}\xi_{Fds}$ is expected to be proportional to the size of the array and the number of elements in the array. This indicates that for a given inter-element spacing the perturbation noise will increase as the number of elements increases. Also for a given number of elements the perturbation noise will increase as the inter-element spacing increases.

❑❑❑

5.4.3 Perturbation Vector Size

In this section, issues related to the implementation of projected perturbation sequences are examined.

As the effect of the projection operation is to reduce the norm of a vector, successive projections of the perturbation vector onto the spatial derivative constraint planes will decrease the size of the vector's components. This has implications in realizing the perturbation sequence, since in the presence of quantisation it will become increasingly more difficult to represent the small components of the perturbation vector since it will be necessary for the hardware weights to have a larger dynamic range. Also, to avoid round off errors, a large perturbation step size γ may have to be employed which in turn may affect the misadjustment level [47].

Another effect of a projection operation is that it may increase the number of significant perturbation vector components. For example, in the case of the Time Multiplex sequence the number of non zero components in each vector changes from one to as many as L. Thus at the application of each weight perturbation, all weights in the array may need adjustment.

Considering a first order constraint system, the decrease in the size of the perturbation vector after projection onto the first order derivative constraint plane is given by

$$(\mathbf{P}_0 - \mathbf{P}_0\mathbf{P}_1)\delta(i) = \frac{\Lambda_X S S^H \Lambda_X}{\sum_j x_j^2}\delta(i) \qquad (5.43)$$

For a Time Multiplex sequence, the change in the size of the k^{th} component of the i^{th} perturbation vector, normalised with respect to the component's maximum possible original size is given by

$$\nabla(\delta_k(i)) = \left(\frac{\Lambda_X S S^H \Lambda_X}{\sqrt{2L}\sum_j x_j^2}\delta(i)\right)_k = \frac{x_i x_k e^{j\frac{2\pi f_o}{\upsilon}(x_k - x_i)\cos(\phi_o)}}{\sum_j x_j^2} \qquad (5.44)$$

Considering only the maximum change in the perturbation vector component which occurs when, $i = k = 1$ or L

$$\nabla(\delta_i(i))_{max} = \frac{(L-1)^2}{8\sum_j^{L/2}(j-1/2)^2} \quad \text{L even} \qquad (5.45)$$

$$\nabla(\delta_i(i))_{max} = \frac{(L-1)^2}{8\sum_j^{(L-1)/2} j^2} \quad \text{L odd} \qquad (5.46)$$

Some typical values for the maximum change in the perturbation vector components are shown in Table 1.

Table 1: Change in Perturbation Vector Component Size

no. elements	3	4	5	6	7	8	9	10	20
$\nabla(\delta_i(i))_{max}$	0.5	0.45	0.4	0.36	0.32	0.29	0.26	0.24	0.14

From Table 1 it can be observed that when L is small the change in the size of the perturbation vector components is quite high, the change decreases as L increases. From these numbers it is expected that a large hardware dynamic range is required to effectively implement the first order projected vectors for all the array sizes [49]. The limitation in accurately representing a perturbation vector component occurs when a component's size or the change in the component's size is too small to be represented by a quantisation step.

5.4.4 Array Directional Sensitivity

An array's sensitivity to directional errors will determine whether it is necessary to project the array's weights onto spatial derivative planes. When an array is insensitive to directional errors it is not necessary to project the perturbation sequence onto spatial derivative constraint planes. Here we identify some simple scenarios for the linear array which are insensitive to directional errors.

A gauge of an array's sensitivity to directional errors can be gained by examining the directional constraints and the range of the array's weights. Considering the linear array presented earlier, for a zeroth order constraint system the following is true

$$\mathbf{P}\Lambda_X \boldsymbol{S} = \Lambda_X \boldsymbol{S} \tag{5.47}$$

where $\mathbf{P}$ is the projection matrix for the zeroth order constraint system.

The vector $\Lambda_X \boldsymbol{S}$ lies in the zeroth order constraint subspace defined by

$$\Gamma = \{\boldsymbol{W}:\ \boldsymbol{W}^H \boldsymbol{S} = 0\} \tag{5.48}$$

Now, when the array operates with weights that are close to the optimum weight vector $\hat{\boldsymbol{W}}$ and the optimum weight vector is orthogonal to the constraint subspace defined by Γ, the first order constraint defined by (5.14) is nearly always satisfied. Accordingly in these scenarios the beam is either sufficiently wide in the look direction or that the first order derivative constraint is limited in broadening the beam. In either case, it is not necessary to project the weight vector or the perturbation vector onto the first order derivative constraint plane. In the latter case the second order derivative constraint would be more useful in broadening the beam and suppressing the perturbation noise. The same result is expected when the optimum weight vector is approximately orthogonal to the zeroth order constraint subspace.

Most of the scenarios where the first order constraints are automatically satisfied by the weights in use are trivial but they do depend on the number of signals present and the number of elements in the array. Some examples are shown below:

For a quarter wavelength linear array the first order constraint is always satisfied when:

Example 1. For any sized array and the look direction $\phi_o = n\pi$ n=0,1,2,...

Example 2. For a 2-element array and

i) only the look direction signal is present

ii) the look direction signal and interference signal arrive from the broadside direction

iii) interference signals arrive from the broadside direction.

Example 3. For a 3-element array with only the look direction signal present and the look direction $\phi_o = n\pi/2$ n=0,1,2,... .

Note that in general when an array has a limited number of degrees of freedom, such as when the incident number of interference signals on the array is close to or equal to the number of array elements, it is also possible for the array to be less sensitive to directional errors since it is less capable of nulling signals close to the look direction.

5.5 Simulation Studies

Simulations were performed for a linear planar array with quarter wavelength inter-element spacing. The perturbation noise response of the array was first examined. Without loss of generality, the perturbation noise or the derivative of the perturbation noise has been normalised with respect to p_{ds} and γ.

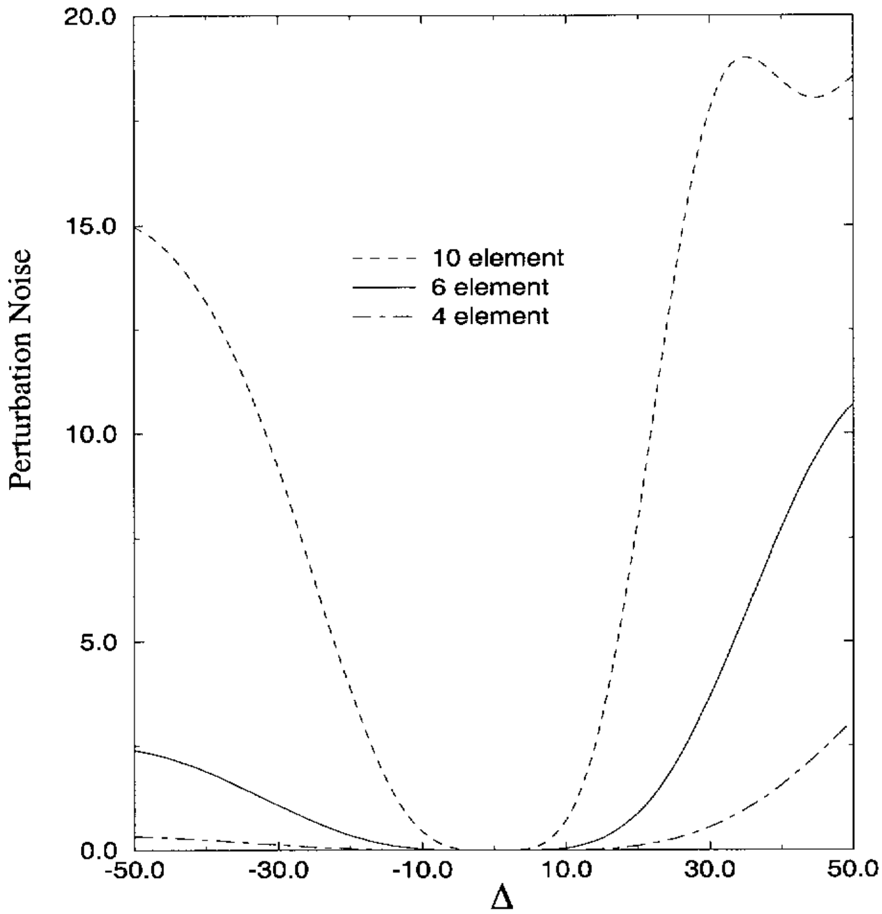


Figure 19 Perturbation Noise Response

Figure 19 shows the perturbation noise response for the following scenario:

a. 10-, 6- and 4-element array

b. perturbation sequence projected onto the zeroth plus first order constraint plane

c. look direction fixed at $\phi_o = 55^o$.

From Figure 19 it can be observed that as the number of elements in the array increases so too does the perturbation noise and the array's perturbation noise performance degrades. In all cases, as Δ varies away from the look direction, the perturbation noise response degrades. It can also be observed, that in the vicinity of the look direction, the smaller 4-element array has better perturbation noise suppression and is able to null perturbation noise within 20 degrees of the look direction. In all cases the beam width for perturbation noise suppression is significant for a narrowband application.

Figure 20 shows the perturbation noise response for the following scenario:

a. 10-element array

b. perturbation sequence projected onto the zeroth order constraint plane and the zeroth plus first order constraint plane

c. look direction varies

From Figure 20 it can be observed that for either projection the array's perturbation noise response degrades slowly when the look direction is close to the end fire direction, the degradation increasing as the look direction approaches broadside as predicted. It can also be seen from the zeroth order projection that when the look direction is close to the end fire direction the beam width for perturbation noise suppression is sufficiently large for narrowband applications. For all look directions, the beam width for perturbation noise suppression is significantly broadened when the perturbation vectors are projected onto the zeroth plus first order constraint planes. In all cases for a narrowband application it would be unnecessary to project the perturbation vectors onto a higher order constraint plane than the first.

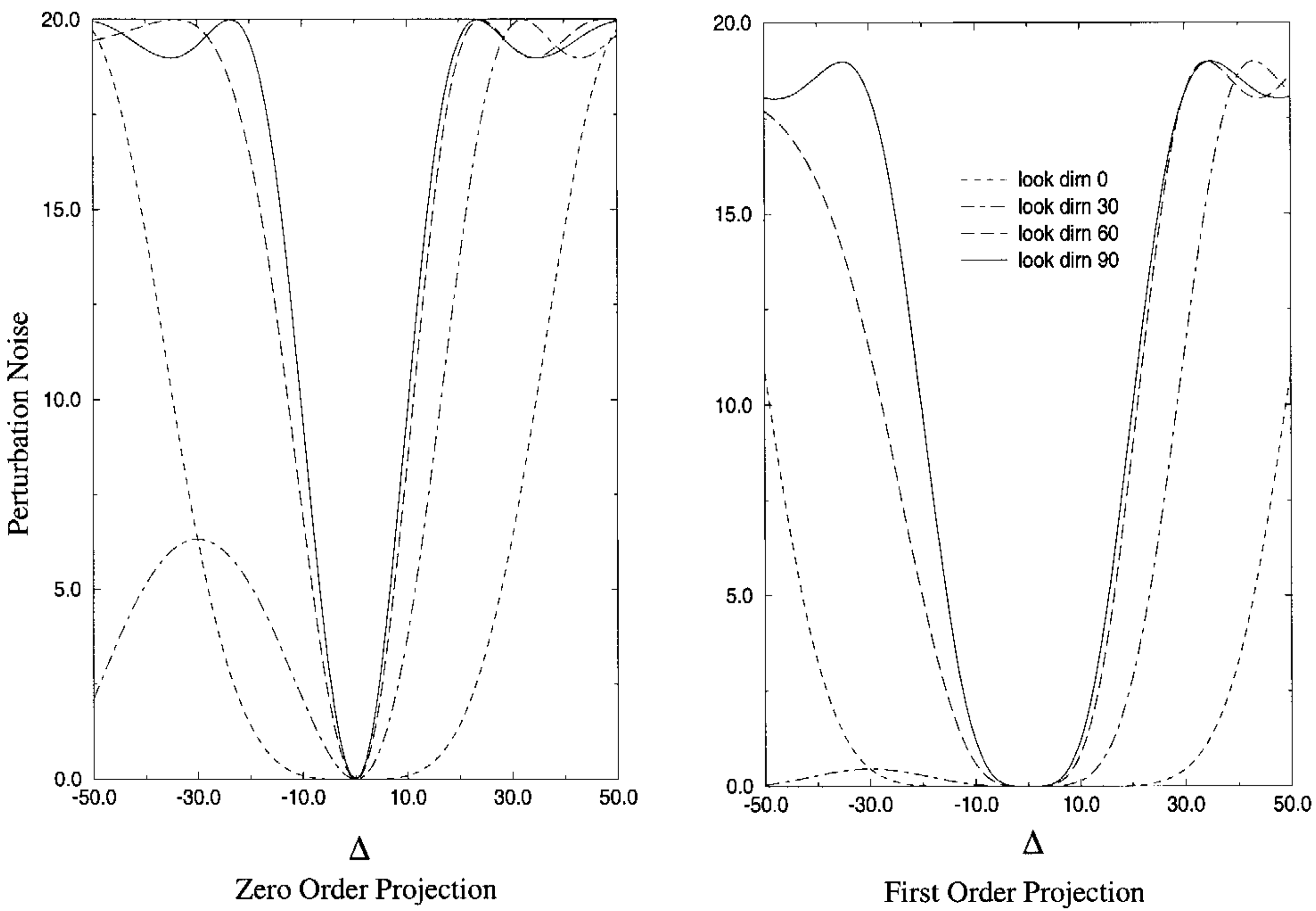


Figure 20 Perturbation Noise Response

Figure 21 shows the derivative of the perturbation noise response for the following scenario:

a. 10-, 6- and 4-element array

b. perturbation sequence projected onto the zeroth order constraint plane and the zeroth plus first order constraint plane

c. look direction varies

Similar observations to those made for the previous figures can be made here:

1) with the zeroth plus first order projection the array's perturbation noise performance degrades slowly with Δ in the vicinity of the look direction. This is clearly seen from the fact that $\Delta = 0$ is a point of inflection, independent of the look direction.

2) the larger arrays perturbation noise performance degrades more quickly with Δ as Δ varies from the look direction.

3) as the look direction approaches the broadside direction the sensitivity of the array's perturbation noise performance to look direction errors increases.

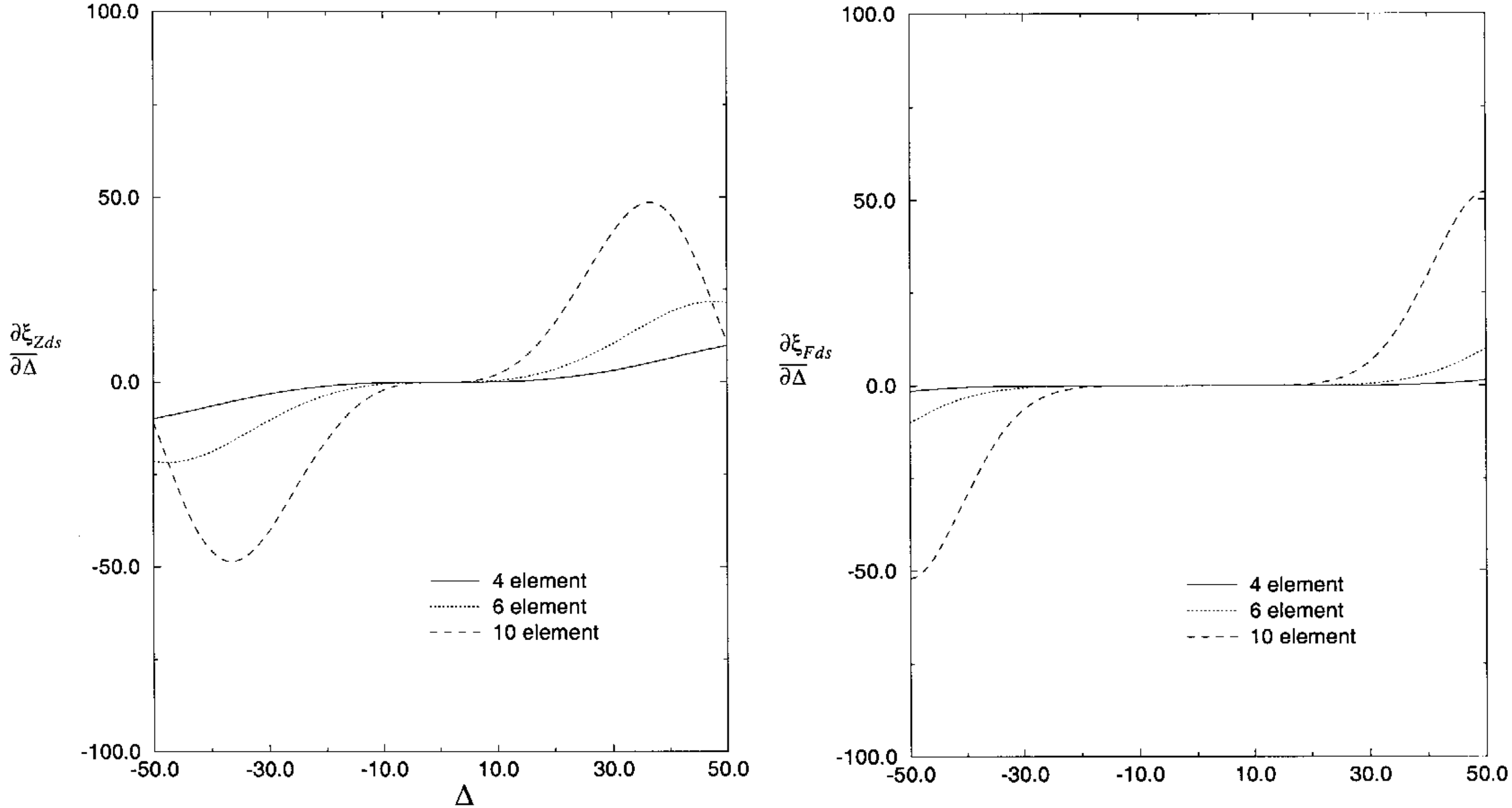


Figure 21a, look direction 0

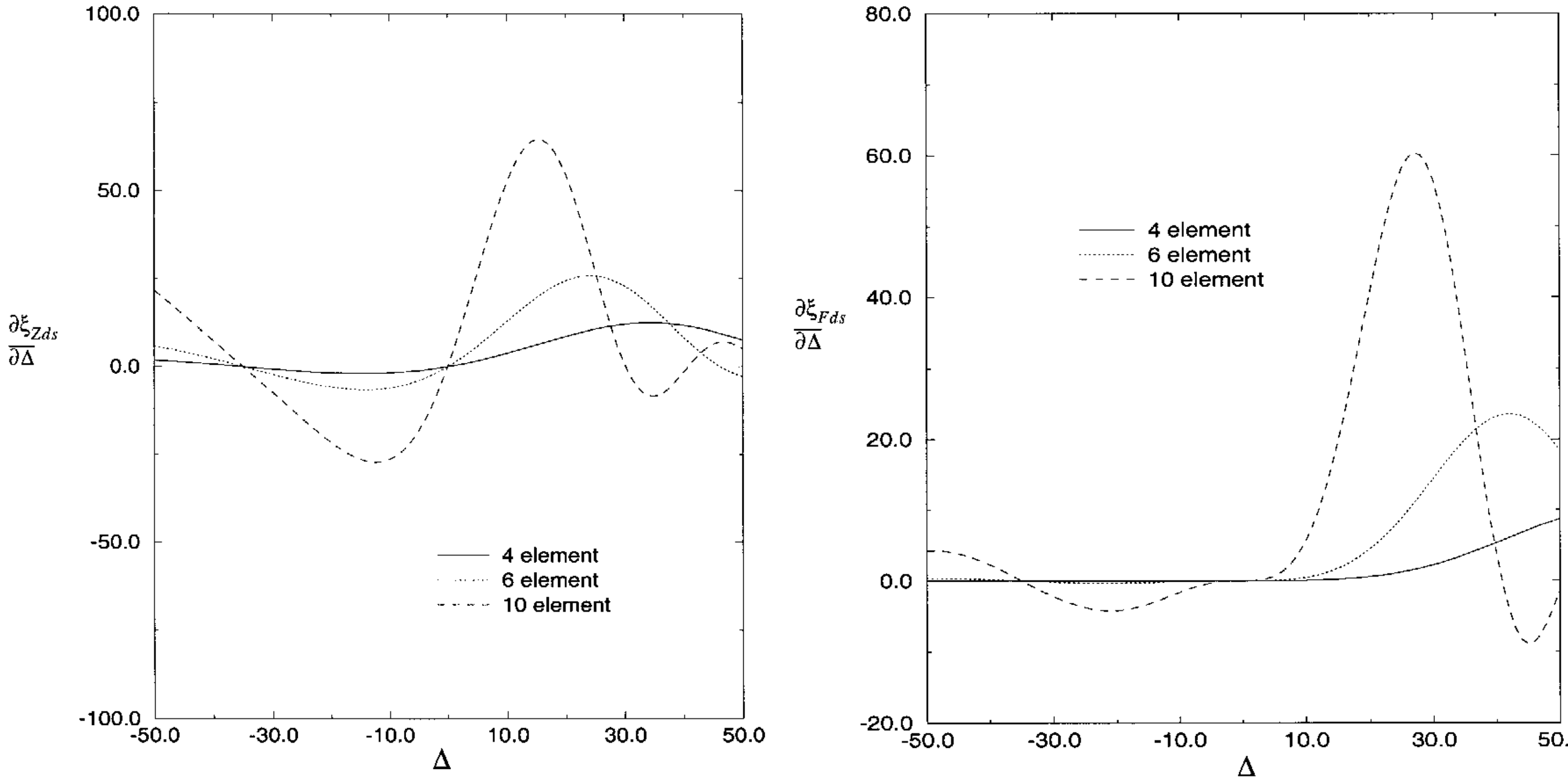


Figure 21b, look direction 35

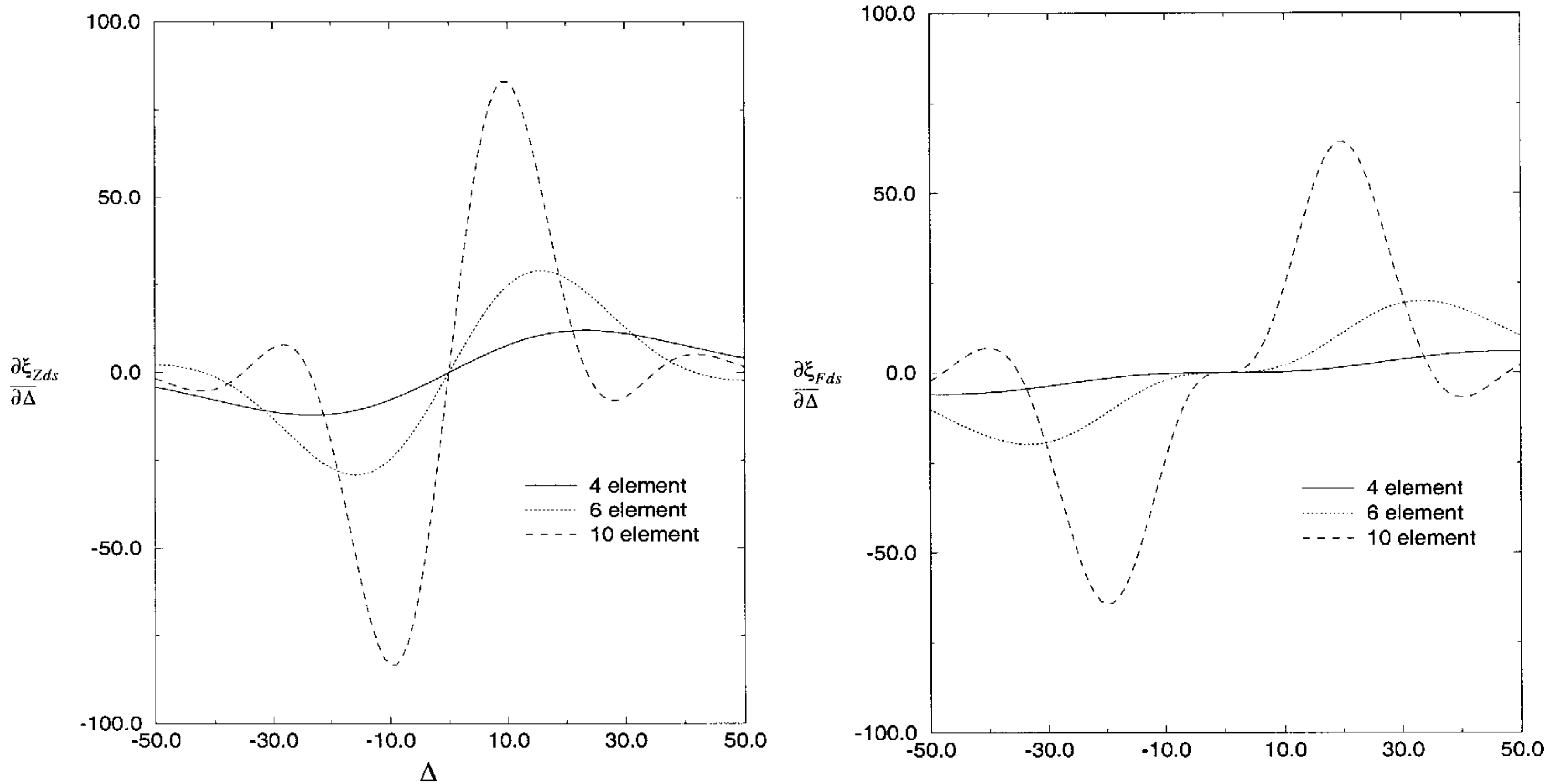
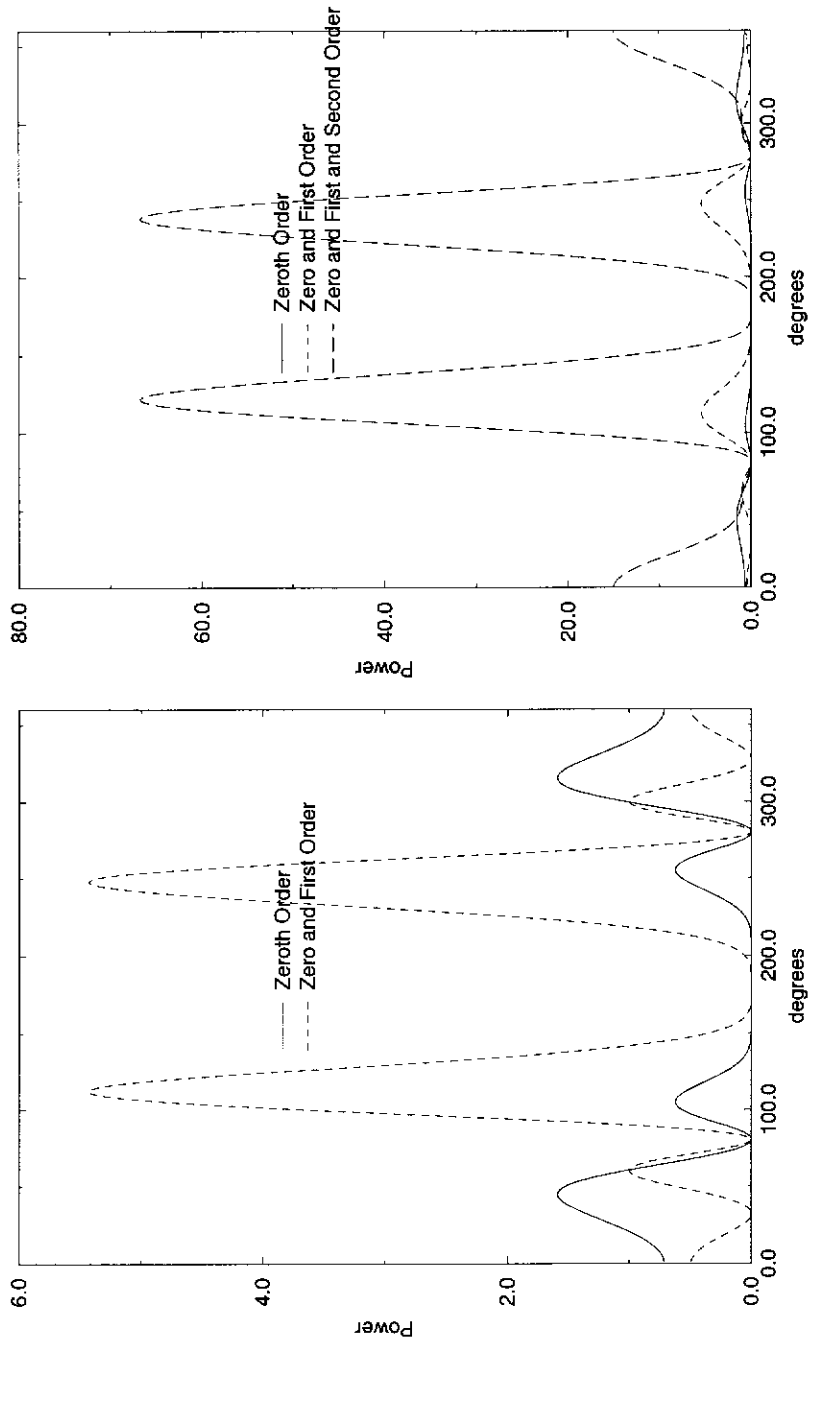


Figure 21c, look direction 90

In Figure 22 for the scenario described in the figure, we show the optimum power response of a 6 element array to a signal of unity power when zeroth, first and second order constraints are used. In the figure "degrees" represents the direction of arrival (ϕ as defined in Figure 1) of the unity signal. In the figure, one observes that particularly away from the look direction, the power level increases as more spatial derivative constraints are added. Hence a deficiency of using the spatial derivative constraints is a possible increase in interference power and loss in array gain.



Signal	Direction of Arrival ϕ	Power
Desired	60	0.1
Interference 1	80	0.5
Interference 2	180	0.5

Figure 22 Optimum Power Response

5.6 Summary

It has been shown that to determine whether spatial derivative constraints should be applied to perturbation vectors, the array's perturbation noise response should be examined. The performance gained with the use of the projected perturbation sequences must generally be weighed against the additional complexity and cost involved in implementing the sequences and this can only be done on a case by case basis.

For the linear array under study, expressions for the perturbation noise response have been developed. The expressions show that the perturbation noise in the desired signal direction is a function of the desired signal's offset from the look direction, the look direction, the size of the array and the number of elements in the array. It has been shown that for a small array, it was sufficient to project the perturbation sequence onto the zeroth order constraint plane to suppress perturbation noise and only in the larger arrays, depending on the look direction, it was necessary to project the perturbation sequence onto the zeroth plus first order constraint planes.

Projection of the perturbation sequence onto spatial derivative constraint planes higher than the first was not necessary as the beam width for perturbation noise suppression gained by the use of the zeroth order constraint or the zeroth plus first order constraint was sufficient for narrowband applications. Unlike the beam broadening in signal reception, the effectiveness of the spatial derivative constraints in suppressing perturbation noise is independent of the signal scenario.

Chapter 6

Summary and Extensions

6.1 Summary

The work presented in this thesis has focused on quantifying the performance of narrowband adaptive arrays that employ the projected perturbation technique. The performance of the narrowband arrays have been analysed both under idealised conditions and when practical implementation effects occur. The resulting analysis provides mechanisms by which the projected perturbation technique can be assessed against others.

In summary the contributions of this thesis are as follows:

- Identified the different approaches possible for extracting the required gradient when using the projected perturbation approach.
- Under idealised conditions we have characterised the performance of the projected perturbation approach by:
  - deriving expressions for the gradient covariance
  - determining the optimum perturbation step size
  - analysing the transient performance of the weight covariance matrix
  - determining the misadjustment.
- Introduced a new misadjustment analysis technique.
- Extended the system performance characterisation to include digital implementation effects by:
  -developing new expressions for the gradient covariance and the misadjustment in the presence of weight quantisation
  -determining the level of loss of performance due to weight quantisation and the limited dynamic range of the array weights.
- Developed new expressions to quantify an array's sensitivity to perturbation noise. These expressions can be used to determine the benefit of using spatial

derivative constraints in the projection operation to counteract directional mismatch and to improve an array's perturbation noise response.

- Presented simulation studies to verify the accuracy of the new expressions.

6.2 Extensions

There are a number of extensions to the work contained in this thesis, amongst these are:

- Perform validation of the new performance measures on a practical system with an emphasis on developing more accurate models for both signal scenarios and implementation factors.
- Some of the performance measures are complicated, derivation of higher order approximations to these measures would be useful.
- A number of important assumptions have been used in our modelling, for example *Assumption 3.3*. To improve the accuracy of the performance measures these models can be refined.

References

[1] B.Widrow and J.B.Mcool, "Comparison of Adaptive Algorithms Based on Methods of Steepest Descent and Random Search", IEEE Trans. Antennas Propag.,Vol. Ap-24, pp615-637, Sept. 1976.

[2] A.Cantoni, "Application of Orthogonal Perturbation Sequences to Adaptive Beamforming", IEEE Trans. Antennas Propag., Vol. Ap-28, pp191-202, March 1980.

[3] R.M.Davis, D.C.Farden and P.J-S Sher, "A Coherent Perturbation Algorithm", IEEE Trans. Antennas Propag., Vol. Ap-34, no.3, pp 350-387, March 1986.

[4] I.Webster, R.J.Evans and A.Cantoni, "Robust Perturbation Algorithms for Adaptive Antenna Arrays", IEEE Trans. Antennas Propag., Vol. Ap-38, no.2, pp195-201, February 1990.

[5] L.C.Godara and A.Cantoni, "Analysis of the Performance of Adaptive Beamforming using Perturbation Sequences", IEEE Trans. Antennas Propag., Vol. Ap-31, no.2, pp 268-279, March 1983.

[6] L.C.Godara and A.Cantoni, "Analysis of Constrained LMS algorithm with Application to Adaptive Beamforming using Perturbation Sequences", IEEE Trans. Antennas Propag., Vol. Ap-34, no.3, pp 368-379, March 1986.

[7] K.D.Senne, "Adaptive Linear Discrete Time Estimation", Technical Report no. 6778-5, Syst. Theory Lab. Stanford University, Stanford, California U.S.A.

[8] O.L.Frost, "An Algorithm for linearly Constrained Adaptive Array", Proc. IEEE, Vol. 68, no. 8, pp 926-935, August 1972.

[9] J.E.Hudson, Adaptive Array Principles, New York, Peter Peregrins Ltd, New York and London,1981.

[10] L.C.Godara and A.Cantoni, "On the Real and the Complex LMS Algorithms for Adaptive Array Processing", Technical Report no. EE8204, Dept. of Electrical and Computer Engineering, University of Newcastle, Newcastle, New South Wales, 2308 Australia.

[11] L.C.Godara and A.Cantoni, "Signal Representation for Array Processing", Technical Report no. EE8203, Dept. of Electrical and Computer Engineering, University of Newcastle, NSW, 2308, Australia.

[12] R.Bellman, Introduction to Matrix Analysis, McGraw-Hill 1970.

[13] R.A.Monzingo and T.W.Miller, Introduction to Adaptive Arrays, J.Wiley and Sons Inc.U.S.A 1980.

[14] J.E.Hudson, "The Effect of Signal and Weight Coefficient Quantisation in an Adaptive Array Processor", Aspects of Signal Processing. G.Tacconi, Ed., Nato Advance Study of Institute series, P.Readel Pub. pp 423-428, 1980

[15] A.Papoulis, Probability Random Variables and Stochastic Processes., McGraw-Hill, Inc., 1965.

[16] A. Papoulis, Signal Analysis, McGraw Hill Inc., 1977.

[17] B.Widrow et al., "Adaptive Antenna Systems", Proc. IEEE. Vol. 55 pp 2143-2158, December 1967.

[18] M.H.Er and A.Cantoni, "On an Adaptive Antenna Array under Directional Constraint", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-33 no. 4, pp 1326-1328, October 1985.

[19] L.C.Godara, "Performance Analysis of Structured Gradient Algorithm", IEEE Trans. Antennas Propag., Vol. Ap-38, no. 7, pp 1078-1083, July 1990.

[20] R.D.Gitlin, J.E.Mazo and M.G.Taylor, "On the Design of Gradient Algorithms for Digitally Implemented Adaptive Filters", IEEE Trans. on Circuit Theory, pp 125-136, March 1973.

[21] C.Caraisos and B.Liu, "A Roundoff Error Analysis of the LMS Adaptive Algorithm", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-32 no.1, pp 34-41, Feb.1984.

[22] J.Normile and F.M.Boland, "Adaptive Filtering with Finite Wordlength Constraints", IEEE Proc. Vol. 130 Pt. E no.2, pp 42-45, March 1983.

[23] K.D.Tocher, The Art of Simulation, The English Universities Press Ltd., London EC4 1963.

[24] S.Jaggi and B.Martinez, "Upper and Lower Bounds of the Misadjustment in the LMS Algorithm", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-38 no.1, pp 164-166, Jan. 1990.

[25] B.Widrow, "Adaptive Filters 1: Fundamentals", Stanford Electronics Laboratories, Stanford, California, Report Sel-66 - 126 (Tech. report 6764-8) December 1966.

[26] S.Haykin, Adaptive Filter Theory, 2nd edition, Prentice Hall Inc. 1986.

[27] B.Noble, Applied Linear Algebra 1969.

[28] M.C.Blasikiewicz, "Analysis of Optimum Arrays", Honours Thesis, Department of Electrical and Electronic Engineering, The University of Western Australia 1989.

[29] M.H.Er, "Optimum Antenna Array Processors with Linear and Quadratic Constraints", PhD Thesis, Department of Electrical and Computer Engineering, The University of Newcastle, New South Wales 1985.

[30] S.Haykin, Communication Systems, John Wiley and Sons Incorporated 1978.

[31] J.R.Magnus and H.Neudecker, Matrix Differential Calculus with Application in Statistics and Econometrics, John Wiley & Sons 1988.

[32] F.A.Grayhill, Introduction to Matrices with Application in Statistics., John Wiley & Sons 1983.

[33] Compton R.T Jr., Adaptive Antennas Concepts and Performance, Prentice Hall Inc., Englewood Cliff, N.J., 1988.

[34] S.A.Ivandich, "Bandwidth Performance of a Narrowband Adaptive Array", Masters Preliminary Report, Department of Electrical and Electronic Engineering, The University of Western Australia 1989.

[35] S.P.Applebaum and D.J.Chapman, "Adaptive Arrays with Main Beam Constraints", IEEE Trans. Antennas and Propagation Vol. AP-24, no. 5 pp 650-662

Sept. 1976.

[36] I.Webster, "Implementation Issues in Adaptive Narrowband beamforming", PhD Thesis, Department of Electrical Engineering and Computer Science, The University of Newcastle, New South Wales 1990.

[37] I.Webster and R.J.Evans, "Robust Perturbation Algorithms for Adaptive Antenna Arrays", Technical Report no. EE 8811, Department of Electrical Engineering and Computer Science, The University of Newcastle, New South Wales 1988.

[38] A.Cantoni, "A Study of Element Space Antenna Array Processors", Technical Report no. EE8105, Department of Electrical Engineering and Computer Science, The University of Newcastle, New South Wales 1981.

[39] L.C.Godara and A.Cantoni, "Analysis of the Performance of Adaptive Beamforming using Perturbation Sequences", Technical Report EE 8015, Department of Electrical and Computer Engineering, The University of Newcastle, New South Wales, 1980.

[40] L.C.Godara and A.Cantoni, "Analysis of Constrained LMS Algorithm with Applications to Adaptive Beamforming using Perturbation Sequences", Technical Report EE 8403, Department of Electrical and Computer Engineering, The University of Newcastle, New South Wales, 1984.

[41] T.A.Whitelaw, Linear Algebra, Blackie & Son Ltd., Glasgow and London, 1983.

[42] A.K.Steele, "Comparison of Directional and Derivative Constraints for Beamformers Subject to Multiple Linear Constraints", IEEE Proc., Vol. 130, Pts.F&H.

no.1 , pp 41-45, February 1983.

[43] Shen -De Lin and M. Barkat, 'The Performance of the SMI Method in the Constrained LMS Array and the Griffiths Array", IEEE Trans. on Antennas and Prop., Vol. 38, no. 11, pp. 1878-1882, November 1990.

[44] L.J.Griffiths, "An Alternative Approach to Linearly Constrained Adaptive Beamforming", IEEE. Trans. on Antennas and Prop., Vol. AP-30, no.1, pp. 27-34, Jan. 1982.

[45] S.T.Alexander, "Transient Weight Misadjustment Properties of the Finite Precision LMS Algorithm", IEEE. Trans. on Antennas, Speech, and Signal Processing, Vol. ASSP-35, no. 9, pp. 1250-1258, September 1987.

[46] J.E.Hudson, A.Cantoni and A.S.Ratner, "A Coefficient Perturbation HF Adaptive Array", IEE Jnl. Communication and Navigation, Vol. 130, Pts F & H, no. 1, pp 91-97, February 1983.

[47] S.Ivandich and A.Cantoni, "Performance Analysis of Narrowband Adaptive Arrays Using Projected Perturbation Sequences", IEEE Trans. Antennas Propagat., Vol. 41, no. 5,pp 625-634, May 1993.

[48] S.Ivandich and A.Cantoni, "Performance Analysis of Narrowband Adaptive Arrays Using Projected Perturbation Sequences", ISSPA 92, Signal Processing and its Applications. Gold Coast, Australia, pp 581-585, August 16-21 1992.

[49] S.Ivandich, "Quantisation Error Modelling of Narrowband Adaptive Arrays Using Projected Perturbation Sequences". International Conference On Acoustics, Speech and Signal Processing, Adelaide Australia, Volume 2 pp 309-312, April 1994.

[50] B Liu, "Effect of Finite Word-length on the Accuracy of Digital Filters", IEEE Trans. CT-18, pp 670-677, Nov. 1971.

[51] C.Y.Tseng, "Minimum Variance Beamforming with Phase Independent Derivative Constraints", IEEE Trans. on Antennas and Propagat., AP-40, no.3, pp 285-294, Mar. 1992.

[52] K.M. Buckley and L.J.Griffiths, "Linearly- Constrained Beamforming: A Relationship Between Phase Centre Locations and Beampattern", 19th Asilomar Conference on Circuits, Systems and Computers, Pacific Grove, California, session TA- 2, pp. 234-238.

[53] B.Widrow and S.D.Stearns, Adaptive Signal Processing, Prentice-Hall Inc. Englewood Cliffs N.J 1985.

[54] A.J.Viterbi, Principles of Coherent Communications, New York, McGraw Hill, 1966

[55] M.Bellanger, Digital Processing of Signals, New York, Wiley, 1984.

[56] Hahn, Stefan L, Hilbert Transforms in Signal Processing, Boston: Artech House 1996

[57] W. Yates and A.Bateman, Digital Signal Processing Design, London: Pitman, 1988

[58] A.B.Sripad and D.L.Snyder, "A Necessary and Sufficient Condition for Quantization Errors to be Uniform and White", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-25, no 5, October 1977

[59] D.H.Johnson and D.E.Dudgeon, Array Signal Processing: Concepts and Techniques, Prentice Hall: Englewood Cliffs, NJ, 1993.

[60] J.Litva and Titus Kwok-Yeung Lo, Digital Beamforming in Wireless Communications, Artech House: Boston, 1996.

[61] Bernard D. Steinberg, Principles of Aperture and Array System Design : Including Random and Adaptive Arrays, John Wiley & Sons, 1976.

[62] B.Widrow and J.B.Mcool, et al, "Stationay and Non-stationary Learning Characteristics of the LMS Adaptive Filter", Proc IEEE, Vol. 64, no. 8, pp 1151-1162, Aug. 1976.

[63] B.Widrow, J.M.Mcool and M.Ball, "The Complex LMS Algorithm", Proc. IEEE, Vol. 63, no. 4, pp. 75-76, April 1975.

[64] P.E.Green Jr, E.J.Kelly Jr, and M.J.Levin, "A Comparison of Seismic Array Processing Methods", Geophys. J. R. Astron. Soc. Vol. 11. pp 67-84 1966

[65] W.Vanderkulk, "Optimum Processing of Acoustic Arrays", J. Br. Inst. Radio Eng. Vol. 34, pp 285-292, 1963

[66] J.W.R.Griffiths and J.E.Hudson, "An Introduction to Adaptive Processing in a Passive Sonar System", Aspects of Signal Processing, Pt. 1 NATO Advanced Study Institute Series, D. Reidel Publishing Co. Dordrecht, Holland 1977

[67] L.E.Brennan and I.S.Reed, "Theory of Adaptive Radar", IEEE Trans. Aerospace and Electronic Systems, AES-9 pp 237-252, Nov. 1973

[68] R.T.Compton Jr, "Multiplier Offset Voltages in Adaptive Arrays", IEEE Trans. Antennas and Propagation, Vol. AP-24, no. 5, pp 616-627, Sept 1976

[69] L.J.Griffiths, "A Simple Adaptive Algorithm for Real-time Processing in Antenna Arrays", Proc. IEEE, Vol. 57, no. 10, pp. 1696-1704, Oct. 1969.

[70] L.L. Horrowitz and K.D.Senne, "Performance Advantage of Complex LMS Algorithm for Controllling Narrowband Adaptive Arrays", IEEE Tran. Circuit System., Vol. CAS-28, no. 6, pp. 562-576, June 1981

[71] K.M.Buckley and L.J.Griffiths, "An Adaptive Generalised Sidelobe Canceller with Derivative Constraints", IEEE Trans. Antennas and Propagation, Vol. AP-34, pp. 311-319, March 1986.

[72] J.M.Cioffi, "Limited-Precision Effects in Adaptive Filtering", IEEE Trans. Circuits and Systems, Vol. CAS-34, no. 7, pp. 821-833, July 1987.

[73] M.H.Er and A.Cantoni, "Derivative Constraints for Broadband element space antenna array processors", IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-31, pp. 1378-1393, Dec. 1983

[74] B.Fisher and N.J.Bershad, "The Complex LMS Adaptive Algorithm - Transient Weight Mean and Covariance with Applications to ALE", IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-31, no 1, pp 34-44, Feb. 1983.

[75] W.C.Knight, R.G.Pridham and S.M.Kay, "Digital Signal Processing for Sonar", Proc. IEEE, Vol. 69, no. 11, pp. 1451-1506, Nov. 1981.

[76] D.T. Sherwood and N.J.Bershad, "Quantisation Effects in the Complex LMS

Adaptive Algorithm: Linearisation using Dither Theory", IEEE Trans. Circuits and Systems, Vol. CAS-34, no. 7, pp. 848-854, July 1987.

[77] S. Anderson, M.Millnert, M.Viberg, and B.Wahlberg, "An Adaptive Array for Mobile Communication Systems", IEEE Transactions on Vehicular Technology, Vol. 40, pp. 230-236, Feb. 1991.

[78] D.Asztely, "On Antenna Arrays in Mobile Communication Systems: Fast Fading and GSM Base Station Receiver Algorithms", Tech. rep., Royal Institute of Technology, Stockholm, Sweden, Mar. 1996. Royal Institute of Technology.

[79] P.Chevalier, F.Pipon, J.J.Monot, and C.Demeure, "Smart Antennas for the GSM System Experimental Results for a Mobile Reception", IEEE Vehicular Technology Conference, 1997.

[80] J.Miller and S.Miller, "An Adaptive Antenna Array for Multiple Received Signals in Direct-Sequence CDMA Communications Systems", IEEE Milcom, pp. 733-737, 1994.

[81] B.Suard, A.F.Naguib, G.Xu, and A.Paulraj, "Performance of CDMA Mobile Communication Systems using Antenna Arrays", IEEE ICASSP, pp. IV: 153-156, 1993.

[82] G.Tsoulos, M.Beach, and S.Swales, "Adaptive Antennas for Third Generation DS-CDMA Cellular Systems", Proc., IEEE Veh. Tech. Conf., Vol. 1, (USA), pp. 45-49, IEEE; New York, NY, USA, 1995.

Appendix A

This Appendix contains Lemmas and Theorems that are required in the derivation of the results contained in this thesis.

Lemma A.1

$$\lambda_{min}(\mathbf{A})Tr[\mathbf{B}] \le Tr[\mathbf{AB}] \le \lambda_{max}(\mathbf{A})Tr[\mathbf{B}]$$
$$\forall \quad (\mathbf{A} \in H) \text{ and } \forall \quad (\mathbf{B} \in H^{+}) \qquad \text{(a.1)}$$

where $\lambda_{min}(\mathbf{A})$ and $\lambda_{max}(\mathbf{A})$ represents the minimum and maximum eigenvalues of a square matrix $\mathbf{A}$

Proof of this Lemma can be found in [39].

Lemma A.2

a) $Tr(\mathbf{AB}) = Tr(\mathbf{BA})$ for every matrix $\mathbf{A}$ and $\mathbf{B}$ (a.2)

b) $Tr(\mathbf{A}XY^{H}) = Y^{H}\mathbf{A}X$ for any matrix $\mathbf{A}$ and vectors X and Y (a.3)

c) $\lambda_{min}(\mathbf{A})Tr(\mathbf{AB}) \le Tr(\mathbf{ABA}) \le \lambda_{max}(\mathbf{A})Tr(\mathbf{BA})$

$$\forall \quad ((\mathbf{A},\mathbf{B}) \in H^{+}) \qquad \text{(a.4)}$$

Proof of this Lemma can be found in [39].

Lemma A.3

If $\{X(n)\}$ is a complex gaussian process representing the narrowband signals as defined in Chapter 2, [11], for any matrix $\mathbf{A}$

$$E[X(n)X^{H}(n)\mathbf{A}X(n)X^{H}(n)] = \mathbf{RAR} + Tr(\mathbf{RA})\mathbf{R} \qquad \text{(a.5)}$$

Derivation of (a.5)

Let $\mathbf{A}$ be represented by the product of two vectors such that $\mathbf{A} = CD^{H}$, then the $(i,j)^{th}$ component of the expectant matrix can be represented as

$$E[X(n)X^{H}(n)\mathbf{A}X(n)X^{H}(n)]_{i,j} = E\Big[\sum_{l}\sum_{k} x_i x_l^H c_l d_k^H x_k x_j^H\Big] \quad l,k=1,2,...L$$

Taking the expectant operator inside the summation and re-arranging

$$= \sum_{l}\sum_{k} c_l d_k^H E[x_i x_l^H x_k x_j^H]$$

using the fourth order moment of complex gaussian variates [9], [11], [41].

$$= \sum_{l}\sum_{k} c_l d_k^H \{E[x_i x_l^H]E[x_k x_j^H] + E[x_i x_k]E[x_l^H x_j^H] + E[x_i x_j^H]E[x_l^H x_k]\}$$

$$= \sum_{l}\sum_{k} c_l d_k^H \{\mathbf{R}_{il}\mathbf{R}_{kj} + \mathbf{R}_{ij}\mathbf{R}_{lk}^H\}$$

$$= \sum_{l}\sum_{k} c_l d_k^H \{\mathbf{R}_{il}\mathbf{R}_{kj} + \mathbf{R}_{ij}\mathbf{R}_{kl}\}$$

Now, using (a.3) the expectant value of the first term in (a.5) can be written as

$$E[X(n)X^H(n)\mathbf{A}X(n)X^H(n)] = \mathbf{R}CD^H\mathbf{R} + \mathbf{R}D^H\mathbf{R}C = \mathbf{RAR} + Tr(\mathbf{RA})\mathbf{R}$$

Lemma A.4

Let $\mathbf{V}_T(W)$ denote the covariance of T for a given W that is

$$\mathbf{V}_T(W) = E[TT^H|W] - E[T|W]E[T^H|W] \tag{a.6}$$

where: T is a random vector given by

$$T = \frac{1}{\gamma m}\sum_{i=1}^{m} d_i(W)\delta(i) \tag{a.7}$$

$\delta(i)$ is a vector in the complex vector sequence S given by

$$S = \{\delta(1), \delta(2), ...\delta(m)\} \tag{a.8}$$

and $d_i(W)$ is a random process which has the property that

$$E[d_i(W)d_j^H(W)|W] = E[d_i(W)|W]E[d_j^H(W)|W] \text{ for } i \neq j \tag{a.9}$$

then

$$\mathbf{V}_T(W) = \frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}\{E[d_i d_i^H|W] - E[d_i|W]E[d_i^H|W]\}\delta(i)\delta^H(i) \tag{a.10}$$

Proof of this Lemma is given in [2].

This Lemma is used to determine the gradient covariance when a perturbation approach is used to estimate the gradient.

Lemma A.5

Let $r(X)$ be an odd scalar function defined over the space of L dimensional complex vectors. That is $r(X) = -r(-X)$ for any $L \times 1$ complex column vector X. If the vector sequence S as defined by (a.8) has odd symmetry, then

$$\sum_{i=1}^{m} r(\delta(i))\delta(i)\delta^H(i) = 0 \tag{a.11}$$

The proof of this Lemma can be found in [6].

Lemma A.6

If the sequence S as given by (a.8) is the odd symmetry Time Multiplex sequence of length $4L$, then for any complex $L \times 1$ column vector $\boldsymbol{H}$

a) $$\frac{1}{m}\sum_{i=1}^{m} \delta(i)\delta^H(i) = 2\mathbf{I}_{LL} \qquad \text{(a.12)}$$

b) $$\frac{1}{m^2}\sum_{i=1}^{m} \delta^H(i)\mathbf{R}\delta(i)\delta(i)\delta^H(i) = Diag(\mathbf{R}) \qquad \text{(a.13)}$$

c) $$\frac{1}{m^2}\sum_{i=1}^{m} \boldsymbol{H}^H\delta(i)\delta^H(i)\boldsymbol{H}\delta(i)\delta^H(i) = Diag(\boldsymbol{H}\boldsymbol{H}^H) \qquad \text{(a.14)}$$

d) $$\frac{1}{m^2}\sum_{i=1}^{m} [\boldsymbol{H}^H\delta(i) + \delta^H(i)\boldsymbol{H}]^2\delta(i)\delta^H(i) = 2Diag(\boldsymbol{H}\boldsymbol{H}^H) \qquad \text{(a.15)}$$

e) $$\frac{1}{m^2}\sum_{i=1}^{m} [\delta^H(i)\mathbf{R}\delta(i)]^2\delta(i)\delta^H(i) = 2L(Diag\mathbf{R})^2 \qquad \text{(a.16)}$$

If the sequence S is a minimum length Time Multiplex sequence of length $2L$ then for any complex $L \times 1$ column vector $\boldsymbol{H}$

f) $$\frac{1}{m}\sum_{i=1}^{m} \delta(i)\delta^H(i) = 2\mathbf{I}_{LL} \qquad \text{(a.17)}$$

g) $$\frac{1}{m^2}\sum_{i=1}^{m} \delta^H(i)\mathbf{R}\delta(i)\delta(i)\delta^H(i) = 2Diag(\mathbf{R}) \qquad \text{(a.18)}$$

h) $$\frac{1}{m^2}\sum_{i=1}^{m} \boldsymbol{H}^H\delta(i)\delta^H(i)\boldsymbol{H}\delta(i)\delta^H(i) = 2Diag(\boldsymbol{H}\boldsymbol{H}^H) \qquad \text{(a.19)}$$

i) $$\frac{1}{m^2}\sum_{i=1}^{m} [\boldsymbol{H}^H\delta(i) + \delta^H(i)\boldsymbol{H}]^2\delta(i)\delta^H(i) = 4Diag(\boldsymbol{H}\boldsymbol{H}^H) \qquad \text{(a.20)}$$

j) $$\frac{1}{m^2}\sum_{i=1}^{m} [\delta^H(i)\mathbf{R}\delta(i)]^2\delta(i)\delta^H(i) = 4L(Diag\mathbf{R})^2 \qquad \text{(a.21)}$$

Proofs of (a.12) to (a.16) can be found in [6], [39].

Derivation of Lemma A.6 f) to j).

For the purpose of the summation in Lemma A.6 (f) to (j) the Time Multiplex sequence, $S = \{\delta(1), \delta(2)..., \delta(2L)\}$, can be broken into two sequences Q_1, Q_2 such that $Q_l = \{\boldsymbol{n}_l(1), \boldsymbol{n}_l(2), ..., \boldsymbol{n}_l(L)\}$ $l = 1,2$. (a.22)

Where $\boldsymbol{n}_1(i) = \delta(2i-1)$ and $\boldsymbol{n}_2(i) = \delta(2L+2i-1)$ then $\boldsymbol{n}_2(i) = j\boldsymbol{n}_1(i)$

Derivation of (a.17)

Using (a.22) in the first term of (a.17) and expanding

$$\frac{1}{m}\sum_{i=1}^{2L}\delta(i)\delta^H(i) = \frac{1}{m}\sum_{i=1}^{L}\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i) + \frac{1}{m}\sum_{i=1}^{L}\boldsymbol{n}_2(i)\boldsymbol{n}_2^H(i) = \frac{2}{m}\sum_{i=1}^{L}\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i) \quad \text{(a.23)}$$

by using the orthogonality properties of the Time Multiplex sequence as defined by (2.31) - (2.34) it can be shown that (a.23) is equal to $2\mathbf{I}_{LL}$

Derivation of (a.18)

Using (a.22) in the first term of (a.18) and expanding

$$\frac{1}{m^2}\sum_{i=1}^{2L}\delta^H(i)\mathbf{R}\delta(i)\delta(i)\delta^H(i) = \frac{1}{m^2}\sum_{i=1}^{L}\boldsymbol{n}_1^H(i)\mathbf{R}\boldsymbol{n}_1(i)\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)$$

$$+ \frac{1}{m^2}\sum_{i=1}^{L}\boldsymbol{n}_2^H(i)\mathbf{R}\boldsymbol{n}_2(i)\boldsymbol{n}_2(i)\boldsymbol{n}_2^H(i)$$

$$= \frac{2}{m^2}\sum_{i=1}^{L}\boldsymbol{n}_1^H(i)\mathbf{R}\boldsymbol{n}_1(i)\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i) \quad \text{(a.24)}$$

by using the orthogonality properties of the Time Multiplex sequence as defined by (2.31) - (2.34) it can be shown that (a.24) is equal to $2Diag(\mathbf{R})$

Derivation of (a.19)

Using (a.22) in the first term of (a.19) and expanding

$$\frac{1}{m^2}\sum_{i=1}^{2L}\boldsymbol{H}^H\delta(i)\delta^H(i)\boldsymbol{H}\delta(i)\delta^H(i) = \frac{1}{m^2}\sum_{i=1}^{L}\boldsymbol{H}^H\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)\boldsymbol{H}\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)$$

$$+ \frac{1}{m^2}\sum_{i=1}^{L}\boldsymbol{H}^H\boldsymbol{n}_2(i)\boldsymbol{n}_2^H(i)\boldsymbol{H}\boldsymbol{n}_2(i)\boldsymbol{n}_2^H(i)$$

$$= \frac{2}{m^2}\sum_{i=1}^{L}\boldsymbol{H}^H\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)\boldsymbol{H}\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i) \quad \text{(a.25)}$$

by using the orthogonality properties of the Time Multiplex sequence as defined by (2.31) - (2.34) it can be shown that (a.25) is equal to $2Diag(\boldsymbol{H}\boldsymbol{H}^H)$

Derivation of (a.20)

Using (a.22) in the first term of (a.20) and expanding

$$\frac{1}{m^2}\sum_{i=1}^{2L}[\boldsymbol{H}^H\delta(i)+\delta^H(i)\boldsymbol{H}]^2\delta(i)\delta^H(i)$$

$$=\frac{1}{m^2}\sum_{i=1}^{L}[\boldsymbol{H}^H\boldsymbol{n}_1(i)+\boldsymbol{n}_1^H(i)\boldsymbol{H}]^2\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)$$

$$+\frac{1}{m^2}\sum_{i=1}^{L}[\boldsymbol{H}^H\boldsymbol{n}_2(i)+\boldsymbol{n}_2^H(i)\boldsymbol{H}]^2\boldsymbol{n}_2(i)\boldsymbol{n}_2^H(i)$$

$$=\frac{1}{m^2}\sum_{i=1}^{L}[(\boldsymbol{H}^H\boldsymbol{n}_1(i)+\boldsymbol{n}_1^H(i)\boldsymbol{H})^2-(\boldsymbol{H}^H\boldsymbol{n}_1(i)-\boldsymbol{n}_1^H(i)\boldsymbol{H})^2]\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)$$

$$=\frac{4}{m^2}\sum_{i=1}^{L}[\boldsymbol{H}^H\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)\boldsymbol{H}]\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)$$

$$=\frac{4}{m^2}(2L)^2Diag(\boldsymbol{H}\boldsymbol{H}^H) \qquad \text{using (a.19)}$$

$$=4Diag(\boldsymbol{H}\boldsymbol{H}^H) \qquad \text{(a.26)}$$

Derivation of (a.21)

Using (a.22) in the first term of (a.21) and expanding

$$\frac{1}{m^2}\sum_{i=1}^{2L}[\delta^H(i)\mathbf{R}\delta(i)]^2\delta(i)\delta^H(i)=\frac{1}{m^2}\sum_{i=1}^{L}[\boldsymbol{n}_1^H(i)\mathbf{R}\boldsymbol{n}_1(i)]^2\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)$$

$$+\frac{1}{m^2}\sum_{i=1}^{L}[\boldsymbol{n}_2^H(i)\mathbf{R}\boldsymbol{n}_2(i)]^2\boldsymbol{n}_2(i)\boldsymbol{n}_2^H(i)$$

$$=\frac{2}{m^2}\sum_{i=1}^{L}[\boldsymbol{n}_1^H(i)\mathbf{R}\boldsymbol{n}_1(i)]^2\boldsymbol{n}_1(i)\boldsymbol{n}_1^H(i)$$

$$=4L(Diag(\mathbf{R}))^2 \qquad \text{(a.27)}$$

Lemma A.7

Let *H(Y)* be a bounded linear function which maps an L dimensional vector $\boldsymbol{Y}$ into a $L\times L$ hermitian matrix. If the sequence $\{Y(n)\}$ satisfies $\sum_{n=1}^{\infty}\|Y(n)\|\le\infty$ then $\sum_{n=1}^{\infty}\|H(Y(n))\|\le\infty$. Proof [39].

Lemma A.8

If a sequence S satisfies the conditions of orthogonality as defined by (2.31) - (2.34) for any $L \times 1$ column vector $\boldsymbol{H}$

$$\frac{1}{m}\sum_{i=1}^{m}\delta(i)\{\delta^H(i)\boldsymbol{H} + \boldsymbol{H}^H\delta(i)\} = 2\boldsymbol{H}$$

Lemma A.9

Let $\boldsymbol{X}$ be a random complex $L \times 1$ vector with $E[\boldsymbol{X}] = \mu$ and the covariance of $\boldsymbol{X}$ defined by $\boldsymbol{V}(\boldsymbol{X}) = \zeta$. Let $\mathbf{A}$ be an $L \times L$ matrix, then

$$E[\boldsymbol{X}^H\mathbf{A}\boldsymbol{X}] = Tr(\mathbf{A}\zeta) + \mu^H\mathbf{A}\mu$$

Proof

$$\begin{aligned} E[\boldsymbol{X}^H\mathbf{A}\boldsymbol{X}] = E[Tr(\boldsymbol{X}^H\mathbf{A}\boldsymbol{X})] &= E[Tr(\mathbf{A}\boldsymbol{X}\boldsymbol{X}^H)] \\ &= Tr(E[\mathbf{A}\boldsymbol{X}\boldsymbol{X}^H]) \\ &= Tr(\mathbf{A}(E[\boldsymbol{X}\boldsymbol{X}^H])) \\ &= Tr(\mathbf{A}(\zeta + \mu\mu^H)) \\ &= Tr(\mathbf{A}\zeta) + \mu^H\mathbf{A}\mu \end{aligned}$$

Theorem A.1

Let the set of difference equations

$$\boldsymbol{D}_i(n+1) = \alpha_i\boldsymbol{D}_i(n) + \beta_i\left[\xi(n) + \sum_{l=1}^{L}\boldsymbol{D}_l(n)\right] + \boldsymbol{C}_i(n) \quad i=1,2,...L \tag{a.28}$$

be such that

$$\lim_{n \to \infty}\xi(n) = \xi' \tag{a.29}$$

$$\lim_{n \to \infty}\boldsymbol{C}_i(n) = 0 \tag{a.30}$$

$$|\boldsymbol{D}_i(0)| < \infty \tag{a.31}$$

and all the eigenvalues of the system, (a.28), are real and positive. (a.32)

If $|\alpha_{max} + \beta_{max}| < 1$ (a.33)

and

$$\prod_{i=1}^{L}(1+\delta-\alpha_i) - \sum_{i=1}^{L}\beta_i\prod_{l \neq i}^{L}(1+\delta-\alpha_l) > 0 \;\forall\; (\delta \geq 0) \tag{a.34}$$

then $\lim_{n \to \infty} \sum_{i=1}^{L} \boldsymbol{D}_i(n)$ exists and is given by

$$\lim_{n \to \infty} \sum_{i=1}^{L} \boldsymbol{D}_i(n) = \frac{\xi' \sum_{i=1}^{L} \frac{\beta_i}{1-\alpha_i}}{1 - \sum_{i=1}^{L} \frac{\beta_i}{1-\alpha_i}} \qquad \text{(a.35)}$$

where α_{max} and β_{max} are maximum of α_i and β_i, i=1,2,...L respectively. Proof [40]. This theorem is used in the *Direct* misadjustment analysis.

Theorem A.2

Let $\{\boldsymbol{B}(n), n=0,1,2...\}$ be a sequence of matrices such that

$$\boldsymbol{B}(0) \in H^+, Tr[\boldsymbol{B}(0)] < \infty \qquad \text{(a.36)}$$

$$\boldsymbol{B}(n+1) = G(\boldsymbol{B}(n)) + \boldsymbol{Z}(n) + \boldsymbol{Q} \qquad \text{(a.37)}$$

where $G(\boldsymbol{A})$ is a linear function of $\boldsymbol{A}$. If

$i)\ G(\boldsymbol{A}) \in H^+ \ \forall \ \ \boldsymbol{A} \in H^+$ (a.38)

$ii)\ Tr(G(\boldsymbol{A})) \le \beta Tr(\boldsymbol{A}) \ \forall \ \ \boldsymbol{A} \in H^+, 0 \le \beta \le 1$ (a.39)

$iii)\ \boldsymbol{Z}(n) \in H \ \forall \ \ n$ (a.40)

$iv)\ \sum_{n=1}^{\infty} \|\boldsymbol{Z}(n)\| < \infty$ (a.41)

$v)\ \boldsymbol{Q} \in H^+$ (a.42)

Then $\exists \ \ q < \infty$ determined only by $\boldsymbol{Q}$ and $G(.)$ such that

$$\lim_{n \to \infty} Tr(\boldsymbol{B}(n)) = q \qquad \text{(a.43)}$$

Proof [39]. This Theorem is used in the *Bounds* misadjustment analysis.

Theorem A.3

If two hermitian matrices do not commute then they cannot be simultaneously transformed into the diagonal form by means of a unitary transformation.
Proof [39].

Theorem A.4

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ be complex gaussian random $L \times 1$ vectors, representing narrowband signals as defined in Chapter 2, which pairwise have the properties defined in [11].

Then

$$E[\boldsymbol{A}^H\boldsymbol{B}\boldsymbol{C}^H\boldsymbol{D}] = Tr(\mathbf{R}_{BC}\mathbf{R}_{DA}) + Tr(\mathbf{R}_{BA})Tr(\mathbf{R}_{DC})$$

where $\mathbf{R}_{XY} = E[XY^H]$

This can be proved by using the result in [9] which states that for any $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{X}, \boldsymbol{Y}$ complex gaussian random $L \times 1$ vectors with the properties as defined in [9] then

$$E[\boldsymbol{U}\boldsymbol{V}^H\boldsymbol{X}\boldsymbol{Y}^H] = E(\boldsymbol{U}\boldsymbol{V}^H)E(\boldsymbol{X}\boldsymbol{Y}^H) + E(\boldsymbol{U}\boldsymbol{Y}^H)E(\boldsymbol{V}^H\boldsymbol{X})$$

Appendix B

This Appendix contains the derivation of the gradient covariance results and intermediate results for the *Direct* misadjustment analysis contained in Chapter 3. Throughout this and other appendices to simplify the notation expressions such as *W(n)*, *X(l+i)* and $E[G(W)|W]$ may be abbreviated to *W*, *X* and $E[G(W)]$ respectively. Where this occurs it is assumed that the reduced notation is obvious.

The conditional covariance of the gradient estimate is defined as

$$\mathbf{V}_G(W) = E[G(W)G^H(W)|W] - E[G(W)|W]E[G^H(W)|W] \tag{b.1}$$

Derivation of Result 3.2.a.

By setting $d_i(W) = \dfrac{(f_1(W_p, i) - f_2(W_m, i))}{2}$ in (a.7) one obtains $T = G_1(W(n))$.

By assumption $\{X(i)\}$ is a sequence of independent random vectors and this implies that (a.9) is satisfied. Applying Lemma A.4 gives

$$\mathbf{V}_{G1}(W(n)) = \frac{1}{4\gamma^2 m^2}\sum_{i=1}^{m}\{E[(f_1 - f_2)(f_1 - f_2)^H|W(n)] - E[(f_1 - f_2)|W(n)]E[(f_1 - f_2)^H|W(n)]\}\delta_p(i)\delta_p^H(i) \tag{b.2}$$

We now evaluate the individual terms on the right hand side of (b.2)

Note that
$$(f_1 - f_2) = (f_1 - f_2)^H = W_p^H(W, i)X(l+i)X^H(l+i)W_p(W, i) - W_m^H(W, i)X(l+i)X^H(l+i)W_m(W, i) \tag{b.3}$$

substituting the expressions for the weights defined by (2.38) and (2.39) in (b.3) gives

$$\begin{aligned}(f_1 - f_2) &= W^H XX^H W + W^H XX^H \gamma\delta_p + \gamma\delta_p^H XX^H W + \gamma^2\delta_p^H XX^H \delta_p \\ &\quad -(W^H XX^H W - W^H XX^H \gamma\delta_p - \gamma\delta_p^H XX^H W + \gamma^2\delta_p^H XX^H \delta_p) \\ &= 2(W^H XX^H \gamma\delta_p + \gamma\delta_p^H XX^H W)\end{aligned} \tag{b.4}$$

using (b.4), the first term in (b.2) is given by

$$\begin{aligned}(f_1 - f_2)(f_1 - f_2)^H &= (f_1 - f_2)^2 \\ &= 4[(W^H XX^H \gamma\delta_p)^2 + (\gamma\delta_p^H XX^H W)^2 + 2\gamma^2(W^H XX^H \delta_p\delta_p^H XX^H W)]\end{aligned} \tag{b.5}$$

Taking the conditional expectation, with respect to *X*, of the individual terms on the right hand side of (b.5), given *W*, and using Lemma A.3 and Lemma A.2 (b), one obtains

$$4E[W^H XX^H \gamma\delta_p W^H XX^H \gamma\delta_p | W(n)] = 4W^H(\mathbf{R}\gamma\delta_p W^H \mathbf{R} + W^H \mathbf{R}\gamma\delta_p \mathbf{R})\gamma\delta_p$$
$$= 8\gamma^2 (W^H \mathbf{R}\delta_p)^2 \tag{b.6}$$

$$4E[\gamma\delta_p^H XX^H W\gamma\delta_p^H XX^H W | W(n)] = 8\gamma^2(\delta_p^H \mathbf{R} W)^2 \tag{b.7}$$

$$8\gamma^2 E[W^H XX^H \delta_p \delta_p^H XX^H W | W(n)] = 8\gamma^2 [W^H(\mathbf{R}\delta_p\delta_p^H \mathbf{R} + \delta_p^H \mathbf{R}\delta_p \mathbf{R})W]$$
$$= 8\gamma^2(W^H \mathbf{R}\delta_p\delta_p^H \mathbf{R}W + \delta_p^H \mathbf{R}\delta_p W^H \mathbf{R}W) \tag{b.8}$$

The second term of (b.2) can be evaluated as

$$E[(f_1 - f_2)|W(n)]E[(f_1 - f_2)^H | W(n)]$$
$$= 4[(W^H \mathbf{R}\gamma\delta_p + \gamma\delta_p^H \mathbf{R}W)(W^H \mathbf{R}\gamma\delta_p + \gamma\delta_p^H \mathbf{R}W)]$$
$$= 4\gamma^2[(W^H \mathbf{R}\delta_p)^2 + 2(W^H \mathbf{R}\delta_p)(\delta_p^H \mathbf{R}W) + (\delta_p^H \mathbf{R}W)^2] \tag{b.9}$$

substituting (b.6) to (b.9) into (b.2) gives

$$\mathbf{V}_{G1}(W(n)) = \frac{1}{m^2}\sum_{i=1}^{m} [(W^H \mathbf{R}\delta_p(i))^2 + (\delta_p^H(i)\mathbf{R}W)^2$$
$$+ 2(W^H \mathbf{R}W)(\delta_p^H(i)\mathbf{R}\delta_p(i))]\delta_p(i)\delta_p^H(i) \tag{b.10}$$

This establishes the result.

Derivation of Result 3.2.b.

By setting $d_i(W) = (f_1(W_p, i) - f_2(W, i))$ in (a.7) one obtains $T = G_2(W(n))$. By assumption $\{X(i)\}$ is a sequence of independent random vectors and this implies that (a.9) is satisfied. Applying Lemma A.4 gives

$$\mathbf{V}_{G2}(W(n)) = \frac{1}{\gamma^2 m^2}\sum_{i=1}^{m} \{E[(f_1 - f_2)(f_1 - f_2)^H | W(n)] -$$
$$E[(f_1 - f_2)|W(n)]E[(f_1 - f_2)^H | W(n)]\}\delta_p(i)\delta^H{}_p(i) \tag{b.11}$$

we now evaluate the individual terms on the right hand side of (b.11). Considering the first term

$$E[(f_1 - f_2)(f_1 - f_2)^H | W(n)]$$
$$= E[(W_p^H XX^H W_P - W(n)XX^H W(n))^2]$$
$$= E[(W_p^H XX^H W_p)^2 + (W^H(n)XX^H W(n))^2 - 2W_p^H XX^H W_p W^H(n)XX^H W(n)] \tag{b.12}$$

using Lemmas A.3 and A.2 (b) the expectant values of the right hand side of (b.12) with respect to X can be evaluated separately as follows

$$E[W_p^H X X^H W_p W_p^H X X^H W_p | W(n)] = W_p^H[\mathbf{R} W_p W_p^H \mathbf{R} + W_p^H \mathbf{R} W_p \mathbf{R}] W_p$$
$$= 2(W_p^H \mathbf{R} W_p)^2 \quad \text{(b.13)}$$

$$E[(W^H(n) X X^H W(n))^2 | W(n)] = 2(W^H(n) \mathbf{R} W(n))^2 \quad \text{(b.14)}$$

$$-2E[W_p^H X X^H W_p W^H(n) X X^H W(n) | W(n)]$$
$$= -2W_p^H[\mathbf{R} W_p W^H(n) \mathbf{R} + W^H(n) \mathbf{R} W_p \mathbf{R}] W(n)$$
$$= -2[(W_p^H \mathbf{R} W_p)(W^H(n) \mathbf{R} W(n)) + (W^H(n) \mathbf{R} W_p)(W_p^H \mathbf{R} W(n))] \quad \text{(b.15)}$$

The second term on the right hand side of (b.11) can be evaluated as

$$E[(f_1 - f_2) | W(n)] E[(f_1 - f_2)^H | W(n)]$$
$$= (W_p^H \mathbf{R} W_p - W^H(n) \mathbf{R} W(n))(W_p^H \mathbf{R} W_p - W^H(n) \mathbf{R} W(n))$$
$$= (W_p^H \mathbf{R} W_p)^2 - 2W_p^H \mathbf{R} W_p W^H(n) \mathbf{R} W(n) + (W^H(n) \mathbf{R} W(n))^2 \quad \text{(b.16)}$$

summing (b.13) to (b.15) and subtracting (b.16) gives

$$(W_p^H \mathbf{R} W_p)^2 + (W^H(n) \mathbf{R} W(n))^2 - 2(W_p^H \mathbf{R} W_p)(W^H(n) \mathbf{R} W(n))$$
$$+ 2(W_p^H \mathbf{R} W_p)(W^H(n) \mathbf{R} W(n)) - 2(W^H(n) \mathbf{R} W_p)(W_p^H \mathbf{R} W(n)) \quad \text{(b.17)}$$

substituting the expressions for the weights W_p, (2.43), and re-arranging

$$(W^H(n) \mathbf{R} \gamma \delta_p)^2 + (\delta_p^H \gamma \mathbf{R} W(n))^2 + (\delta_p^H \gamma^2 \mathbf{R} \delta_p)^2 + 2\gamma^2 (W^H(n) \mathbf{R} W(n))(\delta_p^H \mathbf{R} \delta_p)$$
$$+ 2[(W^H(n) \mathbf{R} \gamma \delta_p)(\gamma^2 \delta_p^H \mathbf{R} \delta_p) + (\gamma \delta_p^H \mathbf{R} W(n))(\gamma^2 \delta_p^H \mathbf{R} \delta_p)] \quad \text{(b.18)}$$

substituting (b.18) into (b.11) gives

$$\mathbf{V}_{G2}(W(n)) = \frac{\gamma^2}{m^2} \sum_{i=1}^{m} (\delta_p^H(i) \mathbf{R} \delta_p(i))^2 \delta_p(i) \delta_p^H(i)$$
$$+ \frac{2W^H(n) \mathbf{R} W(n)}{m^2} \sum_{i=1}^{m} (\delta_p^H(i) \mathbf{R} \delta_p(i)) \delta_p(i) \delta_p^H(i)$$
$$+ \frac{1}{m^2} \sum_{i=1}^{m} [(W^H(n) \mathbf{R} \delta_p(i))^2 + (\delta_p^H(i) \mathbf{R} W(n))^2] \delta_p(i) \delta_p^H(i)$$
$$+ \frac{\gamma}{m^2} \sum_{i=1}^{m} [(W^H(n) \mathbf{R} \delta_p(i))(\delta_p^H(i) \mathbf{R} \delta_p(i))$$
$$(\delta_p^H(i) \mathbf{R} W(n))(\delta_p^H(i) \mathbf{R} \delta_p(i))] \delta_p(i) \delta_p^H(i) \quad \text{(b.19)}$$

Assuming the sequence has odd symmetry, using Lemma A.5, the last term in (b.19) sums to zero.

The gradient covariance is thus defined as

$$\mathbf{V}_{G2}(W(n)) = \gamma^2\mathbf{D} + \mathbf{E}$$

where $\mathbf{D} = \frac{1}{m^2}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))^2\delta_p(i)\delta_p^H(i)$

and

$$\mathbf{E} = \frac{2W^H(n)\mathbf{R}W(n)}{m^2}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))\delta_p(i)\delta_p^H(i)$$
$$+\frac{1}{m^2}\sum_{i=1}^{m}[(W^H(n)\mathbf{R}\delta_p(i))^2 + (\delta_p^H(i)\mathbf{R}W(n))^2]\delta_p(i)\delta_p^H(i)$$

This establishes the result.

Derivation of Result 3.2.c.

By setting $d_i(W) = f_1(W_p, i)$ in (a.7) one obtains $\boldsymbol{T} = G_3(W(n))$.

By assumption $\{X(i)\}$ is a sequence of independent random vectors which implies that (a.9) is satisfied. Applying Lemma A.4 gives

$$\mathbf{V}_{G3}(W(n)) = \frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}\{E[f_1 f_1^H|W(n)] - E[f_1^H|W(n)]E[f_1|W(n)]\}\delta_p(i)\delta_p^H(i) \quad \text{(b.20)}$$

considering the first term of (b.20)

$$E[f_1 f_1^H|W(n)] = E[((W(n)+\gamma\delta_p(i))^H X(l+i)X^H(l+i)(W(n)+\gamma\delta_p(i)))^2] \quad \text{(b.21)}$$

taking expectation with respect to X and applying Lemma A.3 this becomes

$$E[f_1 f_1^H|W(n)] = 2[(W(n)+\gamma\delta_p(i))^H\mathbf{R}(W(n)+\gamma\delta_p(i))]^2 \quad \text{(b.22)}$$

The second term of (b.20) can be evaluated as

$$E[f_1^H|W(n)]E[f_1|W(n)] = [(W(n)+\gamma\delta_p(i))^H\mathbf{R}(W(n)+\gamma\delta_p(i))]^2 \quad \text{(b.23)}$$

substituting (b.22) and (b.23) into (b.20) gives

$$\mathbf{V}_{G3}(W(n)) = \frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}[(W(n)+\gamma\delta_p(i))^H\mathbf{R}(W(n)+\gamma\delta_p(i))]^2\delta_p(i)\delta_p^H(i)$$
$$= \frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}(W^H(n)\mathbf{R}W(n))^2\delta_p(i)\delta_p^H(i)$$

$$+\frac{1}{m^2}\sum_{i=1}^{m}(W^H(n)\mathbf{R}\delta_p(i)+\delta_p^H(i)\mathbf{R}W(n))^2\delta(i)\delta_p^H(i)$$

$$+\frac{\gamma^2}{m^2}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))^2\delta_p(i)\delta_p^H(i)$$

$$+\frac{2}{\gamma m^2}\sum_{i=1}^{m}(W^H(n)\mathbf{R}W(n))(W^H(n)\mathbf{R}\delta_p(i)+\delta_p^H(i)\mathbf{R}W(n))\delta_p(i)\delta_p^H(i)$$

$$+\frac{2}{m^2}\sum_{i=1}^{m}W^H(n)\mathbf{R}W(n)\delta_p^H(i)\mathbf{R}\delta_p(i)\delta_p(i)\delta_p^H(i)$$

$$+\frac{2\gamma}{m^2}\sum_{i=1}^{m}(W^H(n)\mathbf{R}\delta_p(i)+\delta_p^H(i)\mathbf{R}W(n))\delta_p^H(i)\mathbf{R}\delta_p(i)\delta_p(i)\delta_p^H(i) \tag{b.24}$$

Assuming the sequence has odd symmetry, using Lemma A.5, the *4th* and the *6th* terms of (b.24) sum to zero. The gradient covariance can thus be defined as

$$\mathbf{V}_{G3}(W(n)) = \gamma^2\mathbf{A}+\frac{1}{\gamma^2}\mathbf{B}+\mathbf{C}$$

where

$$\mathbf{A} = \frac{1}{m^2}\sum_{i=1}^{m}[\delta_p^H(i)\mathbf{R}\delta_p(i)]^2\delta_p(i)\delta_p^H(i)$$

$$\mathbf{B} = \frac{(W^H(n)\mathbf{R}W(n))^2}{m^2}\sum_{i=1}^{m}\delta_p(i)\delta_p^H(i)$$

$$\mathbf{C} = \frac{1}{m^2}\sum_{i=1}^{m}[W^H(n)\mathbf{R}\delta_p(i)+\delta_p^H(i)\mathbf{R}W(n)]^2\delta_p(i)\delta_p^H(i)$$

$$+\frac{2W^H(n)\mathbf{R}W(n)}{m^2}\sum_{i=1}^{m}\delta_p^H(i)\mathbf{R}\delta_p(i)\delta_p(i)\delta_p^H(i)$$

This establishes the result.

Derivation of Result 3.2.g.

From the definition of the covariance in Lemma A.4 it can be observed that the power term, $d_i(W)$, in (a.7) and the expectant terms in (a.10) are independent of the complex vector sequence S. This indicates, that in the gradient estimation process, when correlation of the output power sequence is performed with another vector sequence, the gradient covariance results established for the projected perturbation approach can be easily modified for the hybrid perturbation analysis. This is done by replacing the last vector product in the generic gradient covariance expressions, $\delta_p(i)\delta_p^H(i)$, with the new correlation sequence $\delta_{new}(i)\delta_{new}^H(i)$.

For example, for the dual receiver dual perturbation system with the hybrid perturbation approach, the gradient covariance is obtained by replacing the last vector product $\delta_p(i)\delta_p^H(i)$ in (3.1) with $\delta(i)\delta^H(i)$. (3.12) is thus obtained.

In this case, and for all receiver structures under study, it can be observed that the gradient covariance expression for the projected perturbation approach is equivalent to pre and post multiplying the gradient covariance expression, for the hybrid perturbation approach, by the projection matrix.

Similarly for the dual receiver with reference receiver and single receiver system when an odd symmetry sequence is used the following are obtained:

Gradient Covariance of Dual Receiver Reference Receiver System using the Hybrid Perturbation Approach

$$\mathbf{V}_{G2}(W(n)) = \gamma^2\mathbf{D} + \mathbf{E} \tag{b.25}$$

where

$$\mathbf{D} = \frac{1}{m^2}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))^2\delta(i)\delta^H(i) \tag{b.26}$$

and

$$\mathbf{E} = \frac{2W^H(n)\mathbf{R}W(n)}{m^2}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))\delta(i)\delta^H(i) + \frac{1}{m^2}\sum_{i=1}^{m}[(W^H(n)\mathbf{R}\delta_p(i))^2 + (\delta_p^H(i)\mathbf{R}W(n))^2]\delta(i)\delta^H(i) \tag{b.27}$$

Gradient Covariance of Single Receiver System using the Hybrid Perturbation Approach

$$\mathbf{V}_{G3}(W(n)) = \gamma^2\mathbf{A} + \frac{1}{\gamma^2}\mathbf{B} + \mathbf{C} \tag{b.28}$$

where

$$\mathbf{A} = \frac{1}{m^2}\sum_{i=1}^{m}[\delta_p^H(i)\mathbf{R}\delta_p(i)]^2\delta(i)\delta^H(i)$$

$$\mathbf{B} = \frac{(W^H(n)\mathbf{R}W(n))^2}{m^2}\sum_{i=1}^{m}\delta(i)\delta^H(i)$$

$$C = \frac{1}{m^2}\sum_{i=1}^{m}[W^H(n)\mathbf{R}\delta_p(i) + \delta_p^H(i)\mathbf{R}W(n)]^2\delta(i)\delta^H(i)$$
$$+\frac{2W^H(n)\mathbf{R}W(n)}{m^2}\sum_{i=1}^{m}\delta_p^H(i)\mathbf{R}\delta_p(i)\delta(i)\delta^H(i) \quad \text{(b.29)}$$

Derivation of Result 3.2.d.

Examining the first two terms of (b.10) and substituting the expression for the projected sequence, (2.67), gives

$$\sum_{i=1}^{m}[(W^H\mathbf{R}\delta_p(i))^2 + (\delta_p^H(i)\mathbf{R}W)^2]$$
$$= \mathbf{P}\sum_{i=1}^{m}[W^H\mathbf{RP}\delta(i) + \delta^H(i)\mathbf{PR}W]^2\delta(i)\delta^H(i)\mathbf{P}$$
$$-2\mathbf{P}\sum_{i=1}^{m}W^H\mathbf{RP}\delta(i)\delta^H(i)\mathbf{PR}W\delta(i)\delta^H(i)\mathbf{P} \quad \text{(b.30)}$$

using (a.14) and (a.15) shows (b.30) is equal to zero.

Considering the last term of (b.10) and making the substituting for the projected perturbation sequence, (2.67) gives

$$\mathbf{V}_{G1}(W(n)) = \frac{\mathbf{P}}{m^2}\sum_{i=1}^{m}2(W^H\mathbf{R}W)(\delta^H(i)\mathbf{PRP}\delta(i))\delta(i)\delta^H(i)\mathbf{P}$$

using (a.13)

$$\mathbf{V}_{G1}(W(n)) = 2W^H(n)\mathbf{R}W(n)\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}$$

This establishes (3.8). By similar application of the Lemmas for the $2L$ length Time Multiplex sequence, (3.9) can be proved.

Derivation of Result 3.2.e.

The expressions for matrices **D** and **E** are evaluated separately for a $4L$ length Time Multiplex sequence. Substituting for the projected perturbation sequence, (2.67), into (3.3) matrix **D** is given by

$$\mathbf{D} = \frac{\mathbf{P}}{m^2}\sum_{i=1}^{m}(\delta^H(i)\mathbf{PRP}\delta(i))^2\delta(i)\delta^H(i)\mathbf{P} \quad \text{(b.31)}$$

using (a.16) gives

$$\mathbf{D} = 2L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P} \tag{b.32}$$

Examining the expression for matrix **E**, (3.4) it can be observed that as in the derivation of *Result 3.2.d.* the last term sums to zero. Thus matrix **E** is equivalent to

$$\mathbf{E} = \frac{2W^H(n)\mathbf{R}W(n)}{m^2}\mathbf{P}\sum_{i=1}^{m}(\delta^H(i)\mathbf{PRP}\delta(i))\delta(i)\delta^H(i)\mathbf{P} \tag{b.33}$$

applying (a.13)

$$\mathbf{E} = 2W^H(n)\mathbf{R}W(n)\mathbf{P}(Diag(\mathbf{PRP}))\mathbf{P} \tag{b.34}$$

using (b.34) and (b.32) the gradient covariance is equivalent to

$$\mathbf{V}_{G2}(W(n)) = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P} + 2W^H(n)\mathbf{R}W(n)\mathbf{P}(Diag(\mathbf{PRP}))\mathbf{P} \tag{b.35}$$

This establishes the result.

Derivation of Result 3.2.f.

The expressions for matrices **A, B** and **C** (3.6) are evaluated separately for a $4L$ length Time Multiplex sequence.

Substituting for the projected perturbation sequence, (2.67), in **A** and applying (a.16)

$$\mathbf{A} = \frac{1}{m^2}\mathbf{P}\sum_{i=1}^{m}[\delta^H(i)\mathbf{PRP}\delta(i)]^2\delta(i)\delta^H(i)\mathbf{P} = 2L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P} \tag{b.36}$$

substituting for the projected perturbation sequence, (2.67), in **B** and applying (a.12)

$$\mathbf{B} = \frac{(W^H(n)\mathbf{R}W(n))^2}{m^2}\mathbf{P}\sum_{i=1}^{m}\delta(i)\delta^H(i) = \frac{(W^H(n)\mathbf{R}W(n))^2}{2L}\mathbf{P} \tag{b.37}$$

Similarly for matrix **C**

$$\mathbf{C} = \frac{1}{m^2}\mathbf{P}\sum_{i=1}^{m}[W^H(n)\mathbf{RP}\delta(i) + \delta^H(i)\mathbf{PR}W(n)]^2\delta(i)\delta^H(i)\mathbf{P}$$

$$+\frac{2W^H(n)\mathbf{R}W(n)}{m^2}\mathbf{P}\sum_{i=1}^{m}\delta^H(i)\mathbf{PRP}\delta(i)\delta(i)\delta^H(i)\mathbf{P}$$

applying (a.15) and (a.13)

$$\mathbf{C} = 2\mathbf{P}Diag(\mathbf{PR}W(n)W^H(n)\mathbf{RP})\mathbf{P} + 2W^H(n)\mathbf{R}W(n)\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} \tag{b.38}$$

substituting (b.36), (b.37) and (b.38) into (3.5) gives (3.11).

Derivation of (3.15), (3.16) and (3.18)

Applying *Assumption 3.3* in (3.11) gives

$$\mathbf{V}_{G3}(W(n)) = \mathbf{P}\Big(\frac{\gamma^2 2(Tr(\mathbf{PRP}))^2}{L} + \frac{(W^H(n)\mathbf{R}W(n))^2}{\gamma^2 2L}$$

$$+ 2Diag(\mathbf{PR}W(n)W^H(n)\mathbf{RP}) + 2\frac{W^H(n)\mathbf{R}W(n)Tr(\mathbf{PRP})}{L}\Big)\mathbf{P} \quad \text{(b.39)}$$

Taking the derivative of (b.39) with respect to γ and setting it equal to zero and re-arranging

$$\frac{1}{\hat{\gamma}^4}\mathbf{I}_{LL} = \frac{4(Tr(\mathbf{PRP}))^2}{(W^H(n)\mathbf{R}W(n))^2}\mathbf{I}_{LL}$$

which implies

$$\hat{\gamma}(W(n)) = \left[\frac{W^H(n)\mathbf{R}W(n)}{2Tr(\mathbf{PRP})}\right]^{1/2} \quad \text{this is (3.15).}$$

Substituting (3.17) into (3.11) and using *Assumption 3.3* gives

$$\hat{\mathbf{V}}_{G3}(W(n)) = \mathbf{P}\Big(c^2\frac{Tr(\mathbf{PRP})W^H(n)\mathbf{R}W(n)}{L} + \frac{W^H(n)\mathbf{R}W(n)Tr(\mathbf{PRP})}{c^2 L}$$

$$+ 2Diag(\mathbf{PR}W(n)W^H(n)\mathbf{RP}) + 2\frac{W^H(n)\mathbf{R}W(n)Tr(\mathbf{PRP})}{L}\Big)\mathbf{P} \quad \text{(b.40)}$$

re-arranging

$$\hat{\mathbf{V}}_{G3}(W(n)) = 2\mathbf{P}[(Diag(\mathbf{PR}W(n)W^H(n)\mathbf{RP}))$$

$$+ aW^H(n)\mathbf{R}W(n)Diag(\mathbf{PRP})]\mathbf{P} \quad \text{(b.41)}$$

(b.41) is (3.18), and $a = \left(c + \frac{1}{c}\right)^2$. (3.16) can be obtained by substituting $c = 1$ in (b.41).

DIRECT MISADJUSTMENT ANALYSIS INTERMEDIATE RESULTS

Derivation of Result 3.4.c.

Substituting (3.30) into (3.33) and using

$$E[\boldsymbol{W}^H\mathbf{R}\boldsymbol{W}] = Tr[\mathbf{R}\mathbf{R}_{WW}(n)]$$
$$= Tr[\mathbf{R}(\mathbf{K}_{WW} + \overline{\boldsymbol{W}}\overline{\boldsymbol{W}}^H)]$$
$$= Tr(\Lambda\Sigma(n)) + \overline{\boldsymbol{W}}^H\mathbf{R}\overline{\boldsymbol{W}}$$

the following matrix difference equation is obtained

$$\Sigma(n+1) = \Sigma(n) - 2\mu\Lambda\Sigma(n) - 2\mu\Sigma(n)\Lambda + 4\mu^2\Lambda\Sigma(n)\Lambda + \frac{2}{L}\mu^2 Tr(\mathbf{PRP})Tr(\mathbf{R}\mathbf{R}_{WW}(n))\Gamma$$

$$= (\mathbf{I}_{LL} - 4\mu\Lambda + 4\mu^2\Lambda^2)\Sigma(n) + \mu^2 k_2(n)\Gamma \qquad \text{(b.42)}$$

where

$$k_2(n) = \frac{2}{L}Tr(\mathbf{PRP})Tr(\Lambda\Sigma(n)) + \frac{2}{L}Tr(\mathbf{PRP})\overline{\boldsymbol{W}}^H(n)\mathbf{R}\overline{\boldsymbol{W}}(n) \qquad \text{(b.43)}$$

letting δ_1 denote the $L \times 1$ dimensional vector of eigenvalues of **P**

$\boldsymbol{n}_l(n)$ denote the $L \times 1$ dimensional vector of eigenvalues of $\mathbf{K}_{WW}(n)$

λ denote the $L \times 1$ dimensional vector of eigenvalues of **PRP**

then $Tr(\Lambda\Sigma(n)) = \lambda^T\boldsymbol{n}_1(n)$ (b.44)

thus (b.42) reduces to

$$\boldsymbol{n}_1(n+1) = \left(\mathbf{I}_{LL} - 4\mu\Lambda + 4\mu^2\Lambda^2 + \mu^2\frac{2}{L}Tr(\mathbf{PRP})\delta_1\lambda^T\right)\boldsymbol{n}_1(n) + \mu^2\frac{2}{L}Tr(\mathbf{PRP})k_o(n)\delta_1 \qquad \text{(b.45)}$$

where

$$k_o(n) = \overline{\boldsymbol{W}}^H(n)\mathbf{R}\overline{\boldsymbol{W}}(n) \qquad \text{(b.46)}$$

In all cases considered in the thesis, under suitable conditions, $\lim_{n \to \infty} \overline{\boldsymbol{W}}(n) = \hat{\boldsymbol{W}}$. Thus

$$\lim_{n \to \infty} k_o(n) = \hat{\boldsymbol{W}}^H\mathbf{R}\hat{\boldsymbol{W}} \qquad \text{(b.47)}$$

defining

$$\boldsymbol{H}_1 = 4\mu\Lambda - 4\mu^2\Lambda^2 - \frac{2}{L}Tr(\mathbf{PRP})\mu^2\delta_1\lambda^T \qquad \text{(b.48)}$$

(b.45) has a solution defined by

$$n_1(n) = (\mathbf{I}_{LL} - H_1)^n n_1(0) + \mu^2 \frac{2}{L} Tr(\mathbf{PRP}) \sum_{i=1}^{L} k_o(n-i)(\mathbf{I}_{LL} - H_1)^{i-1} \delta_1 \quad (b.49)$$

where $n_1(0)$ denotes the eigenvalues of $\mathbf{K}_{WW}(0)$. The weight covariance matrix is thus defined by

$$\mathbf{K}_{WW}(n) = \sum_{l=1}^{L} n_1(n) Q_l Q_l^H \quad (b.50)$$

where Q_l l=1,2,...L are the eigenvectors of **PRP.** This completes the derivation. A similar analysis can also be performed for the $2L$ length sequence.

Derivation of Result 3.4.d.

Substituting (3.31) into (3.33) the following matrix difference equation is obtained

$$\Sigma(n+1) = (\mathbf{I}_{LL} - 4\mu\Lambda + 4\mu^2\Lambda^2)\Sigma(n) + \mu k_2(n)\Gamma \quad (b.51)$$

where

$$k_2(n) = \frac{2}{L}\gamma^2(Tr(\mathbf{PRP}))^2 + \frac{2}{L}Tr(\mathbf{PRP})Tr(\Lambda\Sigma(n)) + \frac{2}{L}Tr(\mathbf{PRP})\overline{W}^H(n)\mathbf{R}\overline{W}(n) \quad (b.52)$$

letting $n_2(n)$ represent the $L \times 1$ dimensional vector of eigenvalues for $\mathbf{K}_{WW}(n)$, the following vector difference equation can be derived

$$n_2(n+1) = \left(\mathbf{I}_{LL} - 4\mu\Lambda + 4\mu^2\Lambda + \mu^2\frac{2}{L}Tr(\mathbf{PRP})\delta_1\lambda^T\right)n_2(n) + \mu^2\frac{2}{L}Tr(\mathbf{PRP})k_o(n)\delta_1 + \mu^2\gamma^2\frac{2}{L}(Tr(\mathbf{PRP}))^2\delta_1 \quad (b.53)$$

(b.53) has a solution defined by

$$n_2(n) = (\mathbf{I}_{LL} - H_1)^n n_2(0) + \mu^2\frac{2}{L}Tr(\mathbf{PRP})\sum_{i=1}^{L}(k_o(n-i) + \gamma^2 Tr(\mathbf{PRP}))(\mathbf{I}_{LL} - H_1)^{i-1}\delta_1 \quad (b.54)$$

It follows that the weight covariance matrix is given by

$$\mathbf{K}_{WW}(n) = \sum_{l=1}^{L} n_2(n) Q_l Q_l^H \quad (b.55)$$

where $n_2(0)$ is the vector of eigenvalues of $\mathbf{K}_{WW}(0)$ and Q_l l=1,2,...L are the eigenvectors of **PRP**. This completes the derivation.

Generating a Reduced Set of Difference Equations

The difference equations (b.45) and (b.53) each represent a set of L difference equations. These equations can be reduced to a set of L-1 difference equations since one of the components in each of the vectors is identical to zero. A component corresponds to the zero due to the rank deficiency of **P**. As there is a zero eigenvalue in **P** there are corresponding zero eigenvalues in **PRP** and $\mathbf{K}_{WW}(n)$, [40].

Letting the l^{th} component of each of the difference equations correspond to the zero eigenvalue and $\boldsymbol{Q}_l$ the corresponding eigenvector then, $\boldsymbol{n}_{1l}(n) = \boldsymbol{n}_{2l}(n) = 0$.

Thus each of the difference equations can be reduced to a set of L-1 difference equations. Defining the L-1 dimensional vectors $\lambda', \delta', \boldsymbol{n}'_1(n)$ and $\boldsymbol{n}'_2(n)$ such that the i^{th} component is defined by

$$(.)'_i = \begin{cases} (.)_i & i=1,2,.......l\text{-}1 \\ (.)_{i+1} & i=l,l+1,...L\text{-}1 \end{cases} \tag{b.56}$$

where $(.)'$ denotes the $(L-1)\times 1$ dimensional vectors $\lambda', \delta', \boldsymbol{n}'_1(n)$ and $\boldsymbol{n}'_2(n)$ and $(.)$ denotes the corresponding $L \times 1$ dimensional vector. The $(L-1)\times(L-1)$ dimensional matrix $\boldsymbol{H}'_1$ can similarly be defined by dropping the zero column and zero row from $\boldsymbol{H}_1$.

Note that these set of equations correspond to a constraint matrix having a single linear constraint.

Applying (b.56) to (b.45), (b.48) and (b.53) the reduced set of equations are

$$\boldsymbol{H}'_1 = 4\mu\Lambda' - 4\mu^2\Lambda'^2 - \frac{2}{L}Tr(\mathbf{PRP})\mu^2\delta'_1\lambda'^T \tag{b.57}$$

$$\boldsymbol{n}'_1(n+1) = (\mathbf{I} - \boldsymbol{H}'_1)\boldsymbol{n}'_1(n) + \mu^2\frac{2}{L}Tr(\mathbf{PRP})k_o(n)\delta'_1 \tag{b.58}$$

$$\boldsymbol{n}'_2(n+1) = (\mathbf{I} - \boldsymbol{H}'_1)\boldsymbol{n}'_2(n) + \mu^2\frac{2}{L}Tr(\mathbf{PRP})k_o(n)\delta'_1 + \mu^2\gamma^2\frac{2}{L}(Tr(\mathbf{PRP}))^2\delta'_1 \tag{b.59}$$

Derivation of Result 3.4.g.

Let an $(L-1)\times 1$ dimensional vector be defined as

$$D_1(n) = \Lambda' n'_1(n) \quad \text{(b.60)}$$

where $n'_1(n)$ is defined by (b.58) and Λ' is the diagonal matrix of the L-1 non zero eigenvalues of **PRP**. It follows from (b.57) and (b.58) that

$$D_1(n+1) = \left(\mathbf{I} - 4\mu\Lambda' + 4\mu^2\Lambda'^2 + \mu^2\frac{2}{L}Tr(\mathbf{PRP})\Lambda'\grave{e}\grave{e}^T\right)D_1(n) + \mu^2\frac{2}{L}Tr(\mathbf{PRP})\Lambda'\grave{e}k_o(n) \quad \text{(b.61)}$$

where $\grave{e}$ is a column vectors of 1's and $k_o(n)$ is defined by (b.46)

The difference equation of the i^{th} component of D_1 can be obtained from (b.61) as

$$D_{1i}(n+1) = (1 - 4\mu\lambda_i + 4\mu^2\lambda_i^2)D_{1i}(n) + \frac{2}{L}Tr(\mathbf{PRP})\mu^2\lambda_i\left[k_o(n) + \sum_{l=1}^{L-1} D_{1l}(n)\right] \quad \text{(b.62)}$$

where λ_i represents the i^{th} vector component of Λ.

With

$$\alpha_i = 1 - 4\mu\lambda_i + 4\mu^2\lambda_i^2 \quad \text{(b.63)}$$

$$\beta_i = \frac{2}{L}Tr(\mathbf{PRP})\mu^2\lambda_i \quad \text{(b.64)}$$

$$\xi(n) = k_o(n) = \overline{W}^H(n)\mathbf{R}\overline{W}(n) \quad \text{(b.65)}$$

and

$$C_i(n) = 0 \quad \text{(b.66)}$$

(b.62) is similar to (a.28) and Theorem A.1 can be applied. The convergence criteria (a.29)-(a.34) as established in this theorem are now examined.

Since $\lim_{n\to\infty} \overline{W}(n) = \hat{W}$ (b.67)

it follows from (b.65), $\lim_{n\to\infty} \xi(n) = \xi_o = \hat{W}^H\mathbf{R}\hat{W}$ (b.68)

which is (a.29). (b.66) implies (a.30).

Since $\mathbf{K}_{WW}(0) \equiv 0$, it implies that $D_{1i}(0)=0$ which satisfies (a.31).

All the eigenvalues of this system are positive since (b.61) is propagated by a symmetric, positive definite transition matrix for all values of μ. This satisfies (a.32).

Condition (a.33) requires

$$0 < \alpha_{max} + \beta_{max} < 1 \quad \text{(b.69)}$$

$$\Rightarrow 0 < 1 - 4\mu\lambda_{max} + 4\mu^2\lambda^2_{max} + \frac{2}{L}Tr(\mathbf{PRP})\mu^2\lambda_{max} < 1 \quad \text{(b.70)}$$

$$\Rightarrow 0 < \mu < \frac{1}{\lambda_{max} + \frac{Tr(\mathbf{PRP})}{2L}} \tag{b.71}$$

(b.71) is the first term of (3.46), and satisfies (a.33).

Substituting for α_i and β_i, condition (a.34) requires

$$\prod_{i=1}^{L-1} [\delta + 4\mu\lambda_i(1-\mu\lambda_i)] - \sum_{i=1}^{L-1} \frac{2}{L} Tr(\mathbf{PRP})\lambda_i\mu^2 \prod_{l \neq i}^{L-1} [\delta + 4\mu\lambda_l(1-\mu\lambda_l)] > 0 \forall(\delta \geq 0) \tag{b.72}$$

any μ which satisfies (b.71) will also satisfy

$$\prod_{i=1}^{L-1} [\delta + 4\mu\lambda_i(1-\mu\lambda_i)] > 0 \forall(\delta \geq 0) \tag{b.73}$$

substituting (b.73) into (b.72)

$$\Rightarrow \sum_{i=1}^{L-1} \frac{2}{L} Tr(\mathbf{PRP})\lambda_i\mu^2 \prod_{l \neq i}^{L-1} [\delta + 4\mu\lambda_l(1-\mu\lambda_l)] < \prod_{i=1}^{L-1} [\delta + 4\mu\lambda_i(1-\mu\lambda_i)]$$

$$\Rightarrow \sum_{i=1}^{L-1} \frac{\frac{2}{L} Tr(\mathbf{PRP})\lambda_i\mu^2}{\delta + 4\mu\lambda_i(1-\mu\lambda_i)} < 1 \tag{b.74}$$

$$\Rightarrow \sum_{i=1}^{L-1} \frac{\frac{\mu Tr(\mathbf{PRP})}{2L}}{\frac{\delta}{4\mu\lambda_i} + (1-\mu\lambda_i)} < 1 \qquad \forall(\delta \geq 0) \tag{b.75}$$

which is satisfied if

$$\frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i} < 1 \tag{b.76}$$

(b.76) is the second term of (3.46) and satisfies (a.34).

Thus the $\lim_{n \to \infty} \sum_{i=1}^{L-1} \boldsymbol{D}_{1i}(n)$ exists and is given by (a.35). From (b.63) - (b.65), if (b.71) and (b.76) are satisfied

$$\lim_{n \to \infty} \sum_{i=1}^{L-1} \boldsymbol{D}_{1i}(n) = \frac{\xi' \sum_{i=1}^{L-1} \frac{\beta_i}{1-\alpha_i}}{1 - \sum_{i=1}^{L-1} \frac{\beta_i}{1-\alpha_i}} = \frac{\frac{\mu \hat{\boldsymbol{W}}^H \mathbf{R} \hat{\boldsymbol{W}} Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i}}{1 - \frac{\mu Tr(\mathbf{PRP})}{2L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i}} \tag{b.77}$$

It also follows from (3.45) and (b.60) and the fact that $\lambda_L \equiv \lambda_{min} = 0$ that

$$\lim_{n \to \infty} \sum_{i=1}^{L-1} \boldsymbol{D}_{1i}(n) \equiv \lambda^T \boldsymbol{d}(n) \tag{b.78}$$

Hence $\lim_{n \to \infty} \lambda^T \boldsymbol{d}(n)$ converges and is given by (b.77).

Using (3.44) and (b.77) gives (3.47). This establishes the result.

Derivation of (3.39)

the existence of $\lim_{n \to \infty} \sum_{i=1}^{L-1} \boldsymbol{D}_{1i}(n), 0 < \lambda_i < \infty \forall i$ and (b.60) imply that

$\lim_{n \to \infty} \boldsymbol{n}'_1(n)$ exists. Assuming the conditions for convergence hold it follows that

$$\lim_{n \to \infty} \boldsymbol{n}'_1(n) = \mu^2 \frac{2}{L} Tr(\mathbf{PRP}) \hat{\boldsymbol{W}}^H \mathbf{R} \hat{\boldsymbol{W}} \boldsymbol{H}'^{-1}_1 \delta'_1 \tag{b.79}$$

Substituting for the inverse of $\boldsymbol{H}'^{-1}_1$ in the above gives the converged values for the eigenvalues of $\mathbf{K}_{WW}(n)$ which is (3.39)

Derivation of Result 3.4.h.

Let an $(L-1) \times 1$ dimensional vector be defined as

$$\boldsymbol{D}_2(n) = \Lambda' \boldsymbol{n}'_2(n) \tag{b.80}$$

where $\boldsymbol{n}'_2(n)$ is defined by (b.59) and Λ' is the diagonal matrix of the L-1 non zero eigenvalues of **PRP**. It follows from (b.57) and (b.59) that

$$\boldsymbol{D}_2(n+1) = \left(\mathbf{I} - 4\mu\Lambda' + 4\mu^2\Lambda'^2 + \mu^2 \frac{2}{L} Tr(\mathbf{PRP}) \Lambda' \grave{e} \grave{e}^T \right) \boldsymbol{D}_2(n)$$
$$+ \mu^2 \frac{2}{L} Tr(\mathbf{PRP})(k_o(n) + \gamma^2 Tr(\mathbf{PRP})) \Lambda' \grave{e} \tag{b.81}$$

where $\grave{e}$ is a column vectors of 1's and $k_o(n)$ is defined by (b.46)

The difference equation of the i^{th} component of $\boldsymbol{D}_2$ can be obtained from (b.81) as

$$\boldsymbol{D}_{2i}(n+1) = (1 - 4\mu\lambda_i + 4\mu^2\lambda_i^2) \boldsymbol{D}_{2i}(n)$$
$$+ \frac{2}{L} Tr(\mathbf{PRP}) \mu^2 \lambda_i \left[k_o(n) + \gamma^2 Tr(\mathbf{PRP}) + \sum_{l=1}^{L-1} \boldsymbol{D}_{2l}(n) \right] \tag{b.82}$$

With

$$\alpha_i = 1 - 4\mu\lambda_i + 4\mu^2\lambda_i^2 \tag{b.83}$$

$$\beta_i = \frac{2}{L} Tr(\mathbf{PRP}) \mu^2 \lambda_i \tag{b.84}$$

$$\xi(n) = k_o(n) = \overline{W}^H(n)\mathbf{R}\overline{W}(n) + \gamma^2 Tr(\mathbf{PRP}) \tag{b.85}$$

$$C_i(n) = 0 \tag{b.86}$$

(b.82) is similar to (a.28) and Theorem A.1 can be applied. The convergence criteria (a.29)-(a.34) as established in this theorem are now examined.

Since $\lim_{n \to \infty} \overline{W}(n) = \hat{W}$ (b.87)

it follows from (b.85) that $\lim_{n \to \infty} \xi(n) = \xi_o = \hat{W}^H\mathbf{R}\hat{W} + \gamma^2 Tr(\mathbf{PRP})$ (b.88)

which is (a.29). (b.86) implies (a.30).

Since $\mathbf{K}_{WW}(0) \equiv 0$, it implies that $D_{2i}(0)=0$ which satisfies (a.31).

All the eigenvalues of this system are positive since (b.81) is propagated by a symmetric, positive definite transition matrix for all values of μ. This satisfies (a.32).

Condition (a.33) and (a.34) are the same as for the dual receiver dual perturbation case shown previously and are not repeated here.

Since all the conditions of Theorem A.1 are satisfied it follows that

$\lim_{n \to \infty} \sum_{i=1}^{L-1} D_{2i}(n)$ exists and is given by (a.35). Hence from (b.83) to (b.88), if (b.71) and (b.76) are satisfied

$$\lim_{n \to \infty} \sum_{i=1}^{L-1} D_{2i}(n) = \frac{[\hat{W}^H\mathbf{R}\hat{W} + \gamma^2 Tr(\mathbf{PRP})]\frac{\mu Tr(\mathbf{PRP})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}}{1 - \frac{\mu Tr(\mathbf{PRP})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}} \tag{b.89}$$

It follows from (3.45) and (b.80) and the fact that $\lambda_L \equiv \lambda_{min} = 0$ that

$$\lim_{n \to \infty} \sum_{i=1}^{L-1} D_{2i}(n) \equiv \lambda^T d(n) \tag{b.90}$$

Hence $\lim_{n \to \infty} \lambda^T d(n)$ converges and is given by (b.89). (3.50) follows from (b.89) and (3.44). This establishes the result.

Derivation of (3.40)

The existence of $\lim_{n \to \infty} \sum_{i=1}^{L-1} D_{2i}(n)$, $0 < \lambda_i < \infty \forall i$ and (b.80) imply that $\lim_{n \to \infty} n'_2(n)$ exists.

Assuming that these conditions hold it follows that:

$$\lim_{n\to\infty} \boldsymbol{n}'_2(n) = \mu^2\frac{2}{L}Tr(\mathbf{PRP})(\hat{W}^H\mathbf{R}\hat{W} + \gamma^2 Tr(\mathbf{PRP}))H'^{-1}_1\delta'_1 \qquad \text{(b.91)}$$

Substituting for the inverse of H'^{-1}_1 in the above gives the converged values for the eigenvalues of $\mathbf{K}_{WW}(n)$ which is (3.40).

Derivation of Result 3.4.i.

In the first instance the single receiver analysis is handled differently from the dual receiver analysis' since it is not possible to diagonalise the weight covariance matrix and **PRP** by the same unitary transformation. To overcome this we consider the system to be in steady state and determine the gradient covariance.

Using $\lim_{n\to\infty} \overline{W}(n) = \hat{W}$ and the property $\mathbf{PR}\hat{W} = 0$ the gradient covariance expression, (3.18), reduces to

$$\lim_{n\to\infty} \hat{\mathbf{V}}_{G3}(W(n)) = \frac{aTr(\mathbf{PRP})W^H(n)\mathbf{R}W(n)}{L}\mathbf{P} \qquad \text{(b.92)}$$

Note that (b.92) can be diagonalised by the same unitary transformation that diagonalises $\mathbf{K}_{WW}(n)$ and **PRP**. We can now use the previous approach used for the dual receiver cases. Substituting (b.92) into (3.24) gives

$$\begin{aligned}\mathbf{K}_{WW}(n+1) = {} & \mathbf{K}_{WW}(n) - 2\mu[\mathbf{PRPK}_{WW}(n) + \mathbf{K}_{WW}(n)\mathbf{PRP}] \\ & + \frac{\mu^2 aTr(\mathbf{PRP})}{L}Tr(\mathbf{PRPK}_{WW}(n))\mathbf{P} \\ & + \frac{\mu^2 aTr(\mathbf{PRP})}{L}\overline{W}^H(n)\mathbf{R}\overline{W}(n)\mathbf{P} \\ & + 4\mu^2\mathbf{PRPK}_{WW}(n)\mathbf{PRP}\end{aligned} \qquad \text{(b.93)}$$

Pre and Post multiplying (b.93) by $\mathbf{Q}^H$ and $\mathbf{Q}$ respectively and using (3.29) and (3.45) one obtains

$$\begin{aligned}\boldsymbol{d}(n+1) = {} & \left(\mathbf{I}_{LL} - 4\mu\Lambda + 4\mu^2\Lambda^2 + a\mu^2\frac{Tr(\mathbf{PRP})}{L}\delta_1\lambda^T\right)\boldsymbol{d}(n) \\ & + \frac{a\mu^2 Tr(\mathbf{PRP})}{L}k_o(n)\delta_1\end{aligned} \qquad \text{(b.94)}$$

Note that $\boldsymbol{d}(0) \equiv 0$ and due to the rank deficiency of the projection matrix equation (b.94) can be reduced to a set of L-1 difference equations. As performed in the previous derivations of *Results 3.4.g* and *3.4.h*, using (b.56)

$$d'(n+1) = \left(\mathbf{I} - 4\mu\Lambda' + 4\mu^2\Lambda'^2 + a\mu^2\frac{Tr(\mathbf{PRP})}{L}\delta'_1\lambda'^T\right)d'(n)$$
$$+ \frac{a\mu^2 Tr(\mathbf{PRP})}{L}k_o(n)\delta'_1 \qquad \text{(b.95)}$$

where δ'_1 and λ' represent the L-1 non-zero eigenvalues of **P** and **PRP** respectively. $d'(n)$ represents the L-*1*diagonal elements of $\mathbf{Q}^H\mathbf{K}_{WW}(n)\mathbf{Q}$.

Letting

$$D_3(n) = \Lambda' d'(n) \qquad \text{(b.96)}$$

it then follows from (b.95) and (b.96) that

$$D_3(n+1) = \left(\mathbf{I} - 4\mu\Lambda' + 4\mu^2\Lambda'^2 + a\mu^2\frac{Tr(\mathbf{PRP})}{L}\delta'_1\lambda'^T\right)D_3(n)$$
$$+ \frac{a\mu^2 Tr(\mathbf{PRP})}{L}k_o(n)\Lambda'\grave{e} \qquad \text{(b.97)}$$

From (b.97) a difference equation of the i^{th} component of D_3 can be written as

$$D_{3i}(n+1) = (1 - 4\mu\lambda_i + 4\mu^2\lambda_i^2)D_{3i}(n)$$
$$+ \frac{a\mu^2 Tr(\mathbf{PRP})}{L}\lambda_i\left[k_o(n) + \sum_{l=1}^{L-1} D_{3l}(n)\right] \qquad \text{(b.98)}$$

With

$$\alpha_i = 1 - 4\mu\lambda_i + 4\mu^2\lambda^2{}_i \qquad \text{(b.99)}$$

$$\beta_i = \frac{a}{L}Tr(\mathbf{PRP})\mu^2\lambda_i \qquad \text{(b.100)}$$

$$\xi(n) = k_o(n) = \overline{W}^H(n)\mathbf{R}\overline{W}(n) \qquad \text{(b.101)}$$

$$C_i(n) = 0 \qquad \text{(b.102)}$$

(b.98) is similar to (a.28) and Theorem A.1 can be applied. The convergence criteria (a.29)-(a.34) as established in this theorem are now examined.

Since $\lim_{n\to\infty} \overline{W}(n) = \hat{W}$ (b.103)

it follows from (b.101) that $\lim_{n\to\infty} \xi(n) = \xi_o = \hat{W}^H\mathbf{R}\hat{W}$ (b.104)

which is (a.29). (b.102) implies (a.30).

Since $\mathbf{K}_{WW}(0) \equiv 0$, it implies that $D_{3i}(0)=0$ which satisfies (a.31).

All the eigenvalues of this system are positive since (b.97) is propagated by a symmetric, positive definite transition matrix for all values of μ. This satisfies (a.32).

Condition (a.33) requires

$$0 < \alpha_{max} + \beta_{max} < 1 \quad \text{(b.105)}$$

$$\Rightarrow 0 < 1 - 4\mu\lambda_{max} + 4\mu^2\lambda^2{}_{max} + \frac{a}{L}Tr(\mathbf{PRP})\mu^2\lambda_{max} < 1 \quad \text{(b.106)}$$

$$\Rightarrow 0 < \mu < \frac{1}{\lambda_{max} + \frac{aTr(\mathbf{PRP})}{4L}} \quad \text{(b.107)}$$

(b.107) is the first term of (3.51) and satisfies (a.33).

Condition (a.34) requires

$$\prod_{i=1}^{L-1}[\delta + 4\mu\lambda_i(1-\mu\lambda_i)] - \sum_{i=1}^{L-1}\frac{a}{L}Tr(\mathbf{PRP})\lambda_i\mu^2\prod_{l \neq i}^{L-1}[\delta + 4\mu\lambda_l(1-\mu\lambda_l)] > 0\forall(\delta \geq 0) \quad \text{(b.108)}$$

any μ which satisfies (b.107) will also satisfy

$$\prod_{i=1}^{L-1}[\delta + 4\mu\lambda_i(1-\mu\lambda_i)] > 0\forall(\delta \geq 0) \quad \text{(b.109)}$$

substituting (b.109) into (b.108)

$$\Rightarrow \sum_{i=1}^{L-1}\frac{a}{L}Tr(\mathbf{PRP})\lambda_i\mu^2\prod_{l=i}^{L-1}[\delta + 4\mu\lambda_l(1-\mu\lambda_l)] < \prod_{i=1}^{L-1}[\delta + 4\mu\lambda_i(1-\mu\lambda_i)]$$

$$\Rightarrow \sum_{i=1}^{L-1}\frac{\frac{a}{L}Tr(\mathbf{PRP})\lambda_i\mu^2}{\delta + 4\mu\lambda_i(1-\mu\lambda_i)} < 1 \quad \text{(b.110)}$$

$$\Rightarrow \sum_{i=1}^{L-1}\frac{\frac{a\mu Tr(\mathbf{PRP})}{L}}{\frac{\delta}{\mu\lambda_i} + 4(1-\mu\lambda_i)} < 1 \qquad \forall(\delta \geq 0) \quad \text{(b.111)}$$

which is satisfied if

$$\frac{a\mu Tr(\mathbf{PRP})}{4L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i} < 1 \quad \text{(b.112)}$$

(b.112) is the second term of (3.51) and satisfies (a.34). Thus the $\lim_{n \to \infty}\sum_{i=1}^{L-1}\boldsymbol{D}_{3i}(n)$ exists and is given by (a.35). Hence from (b.99) to (b.102), if (b.107) and (b.112) are satisfied

$$\lim_{n \to \infty} \sum_{i=1}^{L-1} D_{3i}(n) = \frac{\xi' \sum_{i=1}^{L-1} \frac{\beta_i}{1-\alpha_i}}{1 - \sum_{i=1}^{L-1} \frac{\beta_i}{1-\alpha_i}} = \frac{[\hat{W}^H \mathbf{R} \hat{W}] \frac{a\mu Tr(\mathbf{PRP})}{4L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i}}{1 - \frac{a\mu Tr(\mathbf{PRP})}{4L} \sum_{i=1}^{L-1} \frac{1}{1-\mu\lambda_i}} \quad \text{(b.113)}$$

From (3.45) and (b.96) and the fact that $\lambda_L \equiv \lambda_{min} = 0$ then

$\lim_{n \to \infty} \sum_{i=1}^{L-1} D_{3i}(n) \equiv \lambda^T \boldsymbol{d}(n)$. Hence $\lim_{n \to \infty} \lambda^T \boldsymbol{d}(n)$ converges and is given by (b.113).

(3.52) follows from (3.44) and (b.113). This establishes the result.

Extension to a Multiple Linear Constraint Optimisation Problem

Even though we have only considered a single linear constraint system, the expressions derived for the converged weight covariance matrix and the misadjustment can be extended to encompass the multi-linear constraint problem when the constraints are linearly independent. When there are *k* linearly independent constraints placed on the system the rank of the projection matrix is reduced to *L-k*. In this instance the projection matrix can be expressed as [9],[12]

$$\mathbf{P} = \mathbf{P}_1 \mathbf{P}_2 ... \mathbf{P}_k \text{ where } \mathbf{P}_i = \left[\mathbf{I}_{LL} - \frac{C_i C_i^H}{C_i^H C_i} \right] \text{ i=1,2...k} \quad \text{(b.114)}$$

Due to the rank property of the projection matrix the eigenvalue expression for the weight covariance matrix, $\boldsymbol{n}_1(n+1)$ and $\boldsymbol{n}_2(n+1)$ and $\boldsymbol{n}_3(n+1)$ can be reduced to a set of $(L-k)$ difference equations. The results for the multi-linear constraint optimisation problem can be derived by inspection by performing the summation over the field $\sum_{i=1}^{L-k}$ in equations (3.41) (3.42) (3.47) (3.50) and (3.52).

Appendix C

This Appendix contains the derivation of intermediate results for the Bounds analysis contained in Chapter 3. In this appendix to simplify the notation expressions such as $E[G(W)|W]$ may be abbreviated to $E[G(W)]$. Where this occurs it is assumed that the reduced notation is obvious.

Derivation of (3.58), Norm of the Weight Error Vector

Considering a system with a single look direction constraint and substituting the expression for the weight error vector, (3.54), into the weight update algorithm defined by (2.27) gives

$$V(n+1) = \mathbf{P}[V(n) - \mu G(W(n))] + \frac{C}{C^H C} - \hat{W} + \mathbf{P}\hat{W} \quad \text{(c.1)}$$

using $\mathbf{P}\hat{W} = \hat{W} - \frac{C}{C^H C}$ in (c.1) gives

$$V(n+1) = \mathbf{P}V(n) - \mu\mathbf{P}G(W(n)) \quad \text{(c.2)}$$

now it can be observed that

$$\mathbf{P}V(n) = V(n) \quad \text{(c.3)}$$

Forming the outer product of (c.2) and taking the conditional expectation of the product with respect to X given $W(n)$ gives

$$\begin{aligned} E[V(n+1)V^H(n+1)|W(n)] = {} & E[\mathbf{P}V(n)V^H(n)\mathbf{P}] \\ & -\mu\mathbf{P}E[G(W(n))V^H(n) + V(n)G^H(W(n))] \\ & + \mu^2\mathbf{P}E[G(W(n))G^H(W(n))]\mathbf{P} \end{aligned} \quad \text{(c.4)}$$

Using (c.3), (b.1) and (2.69) in (c.4)

$$\begin{aligned} E[V(n+1)V^H(n+1)] = {} & E[V(n)V^H(n)] \\ & -2\mu\mathbf{P}E[\mathbf{R}W(n)V^H(n) + V(n)W^H(n)\mathbf{R}]\mathbf{P} \\ & + \mu^2\mathbf{P}E[\mathbf{V}_G(W(n)) + 4\mathbf{P}\mathbf{R}W(n)W(n)\mathbf{R}\mathbf{P}]\mathbf{P} \end{aligned} \quad \text{(c.5)}$$

substituting the expression for the weight error vector in (c.5)

$$\begin{aligned} E[V(n+1)V^H(n+1)] = {} & E[V(n)V^H(n)] \\ & -2\mu\mathbf{P}E[\mathbf{R}(V(n) + \hat{W})V^H(n) + V(n)(V(n) + \hat{W})^H\mathbf{R}]\mathbf{P} \\ & + \mu^2\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P} + \mu^2\mathbf{P}E[4\mathbf{P}\mathbf{R}(V(n) + \hat{W})(V(n) + \hat{W})^H\mathbf{R}\mathbf{P}]\mathbf{P} \end{aligned} \quad \text{(c.6)}$$

using the property

$$\mathbf{PR}\hat{W} = 0 \qquad \text{(c.7)}$$

in (c.6) gives

$$E[V(n+1)V^H(n+1)] = E[V(n)V^H(n)]$$
$$-2\mu\mathbf{P}(\mathbf{R}E[V(n)V^H(n)] + E[V(n)V^H(n)]\mathbf{R})\mathbf{P}$$
$$+\mu^2\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P} + 4\mu^2\mathbf{PR}E[V(n)V^H(n)]\mathbf{RP} \qquad \text{(c.8)}$$

Substituting $B(n) = E[V(n)V^H(n)]$ in (c.8) gives (3.58). This completes the derivation.

Gradient Covariance With Respect to the Weight Error Vector

To determine the misadjustment we require the gradient covariance to be expressed in terms of the weight error vector. Here we evaluate these expressions.

Substituting the expression for the weight error vector, (3.54), into (3.8), (3.10) and (3.18) respectively one obtains the following

$$\mathbf{V}_{G1}(W(n)) = 2(V(n)+\hat{W})^H\mathbf{R}(V(n)+\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} \qquad \text{(c.9)}$$

$$\mathbf{V}_{G2}(W(n)) = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P}$$
$$+2(V(n)+\hat{W})^H\mathbf{R}(V(n)+\hat{W})\mathbf{P}(Diag(\mathbf{PRP}))\mathbf{P} \qquad \text{(c.10)}$$

$$\hat{\mathbf{V}}_{G3}(W(n)) = 2\mathbf{P}[Diag(\mathbf{PR}(V(n)+\hat{W})(V(n)+\hat{W})^H\mathbf{RP})]\mathbf{P}$$
$$+\mathbf{P}[a(V(n)+\hat{W})^H\mathbf{R}(V(n)+\hat{W})Diag(\mathbf{PRP})]\mathbf{P} \qquad \text{(c.11)}$$

substituting (c.3) and (c.7) into the above three equations and rearranging gives

$$\mathbf{V}_{G1}(W(n)) = 2(V^H(n)\mathbf{R}V(n)+\hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} \qquad \text{(c.12)}$$

$$\mathbf{V}_{G2}(W(n)) = 2(V^H(n)\mathbf{R}V(n)+\hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}$$
$$+2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P} \qquad \text{(c.13)}$$

$$\hat{\mathbf{V}}_{G3}(W(n)) = a(V^H(n)\mathbf{R}V(n)+\hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}$$
$$+2\mathbf{P}Diag(\mathbf{PR}V(n)V^H(n)\mathbf{RP})\mathbf{P} \qquad \text{(c.14)}$$

The previous three equations can be expressed in a generic form given by

$$\mathbf{V}_G(W(n)) = (dV^H(n)\mathbf{R}V(n)+e)\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}$$
$$+f\mathbf{P}Diag(\mathbf{PR}V(n)V^H(n)\mathbf{RP})+\mathbf{E} \qquad \text{(c.15)}$$

where

1. Dual Receiver Dual Perturbation System

$$d=2,\ e = 2\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}},\ f=0 \text{ and } \mathbf{E}=0 \tag{c.16}$$

2. Dual Receiver Reference Receiver System

$$d=2,\ e = 2\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}},\ f=0 \text{ and } \mathbf{E} = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P} \tag{c.17}$$

3. Single Receiver System

$$d=a,\ e = a\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}},\ f=2 \text{ and } \mathbf{E}=0 \tag{c.18}$$

Pre and post multiplying (c.15) by **P** taking the expectant value with respect to ***W*** and using the substitution

$$\boldsymbol{B}(n) = E[\boldsymbol{V}(n)\boldsymbol{V}^H(n)] \tag{c.19}$$

the following is obtained.

$$\mathbf{P}E[\mathbf{V}_G(\boldsymbol{W}(n))]\mathbf{P} = (dTr(\mathbf{R}\boldsymbol{B}(n))+e)\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + f\mathbf{P}Diag(\mathbf{PR}\boldsymbol{B}(n)\mathbf{RP})+\mathbf{E} \tag{c.20}$$

Bounds Misadjustment Analysis

Here we evaluate the recursive equation for the norm of the weight error vector to show that it is in a form similar to Theorem A.2

Substituting (c.20) into (3.58) and re-arranging

$$\boldsymbol{B}(n+1) = \boldsymbol{B}(n) - 2\mu\mathbf{P}[\mathbf{R}\boldsymbol{B}(n)+\boldsymbol{B}(n)\mathbf{R}]\mathbf{P} + 4\mu^2\mathbf{PR}\boldsymbol{B}(n)\mathbf{RP} + \mu^2(dTr(\mathbf{R}\boldsymbol{B}(n))\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + f\mathbf{P}Diag(\mathbf{PR}\boldsymbol{B}(n)\mathbf{RP})) + e\mu^2\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \mu^2\mathbf{E} \tag{c.21}$$

using (3.59), the previous equation can be expressed as

$$\boldsymbol{B}(n+1) = G(\boldsymbol{B}(n)) + \mathbf{Q} \tag{c.22}$$

where

$$G(\mathbf{A}) = G_1(\mathbf{A}) + G_2(\mathbf{A}) \tag{c.23}$$

$$G_1(\mathbf{A}) = \mathbf{P}[\mathbf{A} - 2\mu(\mathbf{RA}+\mathbf{AR}) + 4\mu^2\mathbf{RAR}]\mathbf{P} \tag{c.24}$$

$$G_2(\mathbf{A}) = \mu^2\mathbf{P}[dTr(\mathbf{RA})Diag(\mathbf{PRP}) + fDiag(\mathbf{PRARP})]\mathbf{P} \tag{c.25}$$

and

$$\mathbf{Q} = e\mu^2\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \mu^2\mathbf{E} \tag{c.26}$$

The variables d,e,f and **E** are defined by (c.16), (c.17) and (c.18). Now (c.22) is in a form similar form to (a.37) hence Theorem A.2 can be applied. The tests for conditions of convergence are shown next. Note that $\boldsymbol{Z}(n) = 0$ for all the receiver structures.

Common Conditions of convergence for $Tr(\boldsymbol{B}(n))$

Tests on the convergence of $Tr(\boldsymbol{B}(n))$ for the three receiver structures are now presented. It is only necessary to test conditions (a.36), (a.38), (a.39) and (a.42). Condition (a.40) is always satisfied since $\boldsymbol{Z}(n) = 0$. Condition (a.36) and (a.38) are common to all receiver structures.

•Condition (a.36)

$\boldsymbol{B}(0) = E[\boldsymbol{V}(0)\boldsymbol{V}^H(0)]$, $\boldsymbol{B}(0) \in H^+$ since it is a covariance matrix and $Tr[\boldsymbol{B}(0)] = E[\|\boldsymbol{V}(0)\|^2] < \infty$ by assumption. This verifies (a.36).

•Condition (a.38)

Let $\boldsymbol{Y}$ be an arbitrary L dimensional vector and define the vector $\boldsymbol{Y}'$ by

$$\boldsymbol{Y}' = (\mathbf{A}^{1/2} - 2\mu\mathbf{A}^{1/2}\mathbf{R})\mathbf{P}\boldsymbol{Y} \text{ then} \tag{c.27}$$

$$\begin{aligned} \|\boldsymbol{Y}'\| &= \boldsymbol{Y}^H\mathbf{P}[\mathbf{A}^{1/2} - 2\mu\mathbf{R}\mathbf{A}^{1/2}][\mathbf{A}^{1/2} - 2\mu\mathbf{R}\mathbf{A}^{1/2}]\mathbf{P}\boldsymbol{Y} \\ &= \boldsymbol{Y}^H\mathbf{P}[\mathbf{A} - 2\mu[\mathbf{RA} + \mathbf{AR}] + 4\mu^2\mathbf{RAR}]\mathbf{P}\boldsymbol{Y} \\ &= \boldsymbol{Y}^H G_1(\mathbf{A})\boldsymbol{Y} \end{aligned} \tag{c.28}$$

since $\|\boldsymbol{Y}'\| \geq 0$ and $\boldsymbol{Y}$ is an arbitrary vector this establishes $G_1(\mathbf{A}) \in H^+$. $G_2(\mathbf{A})$ clearly is in H^+ for all **A**. Since $G_1(\mathbf{A})$ and $G_2(\mathbf{A}) \in H^+$, $G(\mathbf{A}) \in H^+$. This satisfies condition (a.38).

Conditions for Convergence of $Tr(B(n))$, Dual Receiver Dual Perturbation

To verify the condition (a.39) the trace of (c.23) is taken with the appropriate substitution from (c.16)

$$Tr[G(\mathbf{A})] = Tr[G_1(\mathbf{A})] + Tr[G_2(\mathbf{A})] \quad (c.29)$$

now $Tr[G_1(\mathbf{A})] = Tr[\mathbf{PAP} - 2\mu\mathbf{P}(\mathbf{RA}+\mathbf{AR})\mathbf{P} + 4\mu^2\mathbf{PRARP}]$ (c.30)

Following the same steps of (c.27) and (c.28) it can be shown that the quantity in the square brackets of (c.30) is $\in H^+$. Applying Lemma A.1 and using the fact that the non-zero eigenvalues of **P** are 1

$$Tr[G_1(\mathbf{A})] \le Tr(\mathbf{A}) - 4\mu Tr(\mathbf{RA}) + 4\mu^2 Tr(\mathbf{RAR}) \quad (c.31)$$

applying Lemma A.2 (c)

$$Tr[G_1(\mathbf{A})] \le Tr(\mathbf{A}) + [-4\mu + 4\mu^2\lambda_{max}]Tr(\mathbf{RA}) \quad (c.32)$$

where λ represents the eigenvalues of **R** and λ_{max} is the maximum eigenvalue of **R**. Also

$$Tr[G_2(\mathbf{A})] = Tr[2\mu^2 Tr(\mathbf{RA})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}] \quad (c.33)$$

Applying Lemma A.2 (a) to (c.33)

$$Tr[G_2(\mathbf{A})] = 2\mu^2 Tr(\mathbf{RA})Tr(\mathbf{P}Diag(\mathbf{PRP})) \quad (c.34)$$

combining equations (c.34) and (c.32)

$$Tr[G(\mathbf{A})] \le Tr(\mathbf{A}) + [-4\mu + 4\mu^2\lambda_{max} + 2\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP}))]Tr(\mathbf{RA}) \quad (c.35)$$

for the factor in parenthesis of (c.35) to be negative

$$-4\mu + 4\mu^2\lambda_{max} + 2\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP})) < 0$$

$$\Rightarrow 0 < \mu < \frac{4}{2Tr(\mathbf{P}Diag(\mathbf{PRP})) + 4\lambda_{max}}$$

$$\Rightarrow 0 < \mu < \frac{1}{\frac{Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2} + \lambda_{max}} \quad (c.36)$$

If *Assumption 3.3* is used in (c.36) it reduces to

$$\Rightarrow 0 < \mu < \frac{1}{\frac{Tr(\mathbf{PRP})}{2} + \lambda_{max}}$$

Using Lemma A.1 in (c.35) yields

$$Tr[G(\mathbf{A})] \le Tr(\mathbf{A}) + \lambda_{min}[-4\mu + 4\mu^2\lambda_{max} + 2\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP}))]Tr(\mathbf{A}) \quad (c.37)$$

This can be written as

$$Tr[G(\mathbf{A})] \le \beta Tr(\mathbf{A}) \tag{c.38}$$

where

$$\beta = 1 + 2\mu\lambda_{min}[-2 + \mu Tr(\mathbf{P}Diag(\mathbf{PRP})) + 2\mu\lambda_{max}] \tag{c.39}$$

Equation (c.36) guarantees that β<1 and the fact that $G(\mathbf{A})$ is a positive semi-definite hermitian matrix shows that β>0. This verifies condition (a.39).

The final condition (a.42) requires **Q** to be a positive semi-definite matrix

$$\mathbf{Q} = \mu^2 2\hat{W}^H\mathbf{R}\hat{W}\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} \tag{c.40}$$

It is obvious this is true.

Therefore it has been shown that all the conditions for the convergence of the norm of the weight error vector are satisfied. Thus there exists a $q < \infty$ such that

$$\lim_{n \to \infty} Tr(\boldsymbol{B}(n)) = \lim_{n \to \infty} E[\|V(n)\|^2] = q \tag{c.41}$$

This establishes the convergence of the norm of the weight error vector for the dual receiver dual perturbation system.

Conditions for Convergence of $Tr(\boldsymbol{B}(n))$***, <u>Dual Receiver Reference Receiver</u>***

For the dual receiver reference receiver the conditions for the convergence of $\boldsymbol{B}(n)$, (a.37)- (a.41) as established in Theorem A.2 are equivalent to the dual receiver dual perturbation system using a $4L$ length sequence, hence they will not be repeated here.

The final condition, (a.42), requires **Q** to be a positive semi-definite matrix

$$\mathbf{Q} = \mu^2\mathbf{P}[2\hat{W}^H\mathbf{R}\hat{W}Diag(\mathbf{PRP}) + 2\gamma^2 L(Diag(\mathbf{PRP}))^2]\mathbf{P} \tag{c.42}$$

It is obvious this is true.

Therefore all the conditions for the convergence of the norm of the weight error vector are satisfied. Thus there exists a $q < \infty$ such that

$$\lim_{n \to \infty} Tr(\boldsymbol{B}(n)) = \lim_{n \to \infty} E[\|V(n)\|^2] = q \tag{c.43}$$

This establishes the convergence of the norm of the weight error vector for the dual receiver reference receiver system.

•Condition (a.39)

The trace of (c.23) is taken with the appropriate substitution from (c.18)

$$Tr[G(\mathbf{A})] = Tr[G_1(\mathbf{A})] + Tr[G_2(\mathbf{A})] \tag{c.44}$$

now $Tr[G_1(\mathbf{A})] = Tr[\mathbf{PAP} - 2\mu\mathbf{P}(\mathbf{RA}+\mathbf{AR})\mathbf{P} + 4\mu^2\mathbf{PRARP}]$ (c.45)

applying Lemma's A.1 and using the fact that the non-zero eigenvalues of **P** are 1

$$Tr[G_1(\mathbf{A})] \le Tr(\mathbf{A}) - 4\mu Tr(\mathbf{RA}) + 4\mu^2 Tr(\mathbf{RAR}) \tag{c.46}$$

applying Lemma A.2 (c)

$$Tr[G_1(\mathbf{A})] \le Tr(\mathbf{A}) + [-4\mu + 4\mu^2\lambda_{max}]Tr(\mathbf{RA}) \tag{c.47}$$

also

$$Tr[G_2(\mathbf{A})] = Tr[a\mu^2 Tr(\mathbf{RA})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + 2\mu^2\mathbf{P}Diag(\mathbf{PRARP})\mathbf{P}] \tag{c.48}$$

$$= Tr[a\mu^2 Tr(\mathbf{RA})\mathbf{P}Diag(\mathbf{PRP})] + 2\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRARP})) \tag{c.49}$$

now applying Lemma A.2 to the second component of (c.49) yields

$$Tr(\mathbf{P}Diag(\mathbf{PRARP})) \le Tr(Diag(\mathbf{PRARP})) = Tr(\mathbf{PRARP}) \le Tr(\mathbf{RAR}) \le \lambda_{max} Tr(\mathbf{RA}) \tag{c.50}$$

combining equations (c.49), (c.47) and using (c.50)

$$Tr[G(\mathbf{A})] \le Tr(\mathbf{A}) + [-4\mu + 4\mu^2\lambda_{max} + 2\mu^2\lambda_{max} + a\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP}))]Tr(\mathbf{RA}) \tag{c.51}$$

for the factor in parenthesis to be negative

$$-4\mu + 6\mu^2\lambda_{max} + a\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP})) < 0$$

$$\Rightarrow 0 < \mu < \frac{4}{aTr(\mathbf{P}Diag(\mathbf{PRP})) + 6\lambda_{max}}$$

$$\Rightarrow 0 < \mu < \frac{1}{\frac{aTr(\mathbf{P}Diag(\mathbf{PRP}))}{4} + 1.5\lambda_{max}} \tag{c.52}$$

using Lemma A.1, (c.51) yields

$$Tr[G(\mathbf{A})] \le Tr(\mathbf{A}) + \lambda_{min}[-4\mu + 6\mu^2\lambda_{max} + a\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP}))]Tr(\mathbf{A}) \tag{c.53}$$

This can be written as

$$Tr[G(\mathbf{A})] \le \beta Tr(\mathbf{A}) \tag{c.54}$$

where

$$\beta = 1 + 2\mu\lambda_{min}\left[-2 + \frac{a\mu}{2}Tr(\mathbf{P}Diag(\mathbf{PRP})) + 3\mu\lambda_{max}\right] \tag{c.55}$$

Equation (c.52) guarantees that $\beta<1$ and the fact that $G(\mathbf{A})$ is a positive semi-definite hermitian matrix shows that $\beta>0$. This verifies condition (a.39).

The final condition requires $\mathbf{Q}$ to be a positive semi-definite matrix

$$\mathbf{Q} = \mu^2\mathbf{P}[a\hat{W}^H\mathbf{R}\hat{W}Diag(\mathbf{PRP})]\mathbf{P} \tag{c.56}$$

It is obvious this is true.

Therefore it has been shown that all the conditions for the convergence of the norm of the weight error vector are satisfied. Thus there exists a $q < \infty$ such that

$$\lim_{n\to\infty} Tr(\boldsymbol{B}(n)) = \lim_{n\to\infty} E[\|\boldsymbol{V}(n)\|^2] = q \tag{c.57}$$

This establishes the convergence of the norm of the weight error vector for the single receiver system.

Derivation of Result 3.5.a.

Taking the trace of (c.22), substituting (c.16) and using Lemma A.2

$$Tr[\boldsymbol{B}(n+1)] = Tr[\boldsymbol{B}(n)] - 4\mu Tr[\mathbf{R}\boldsymbol{B}(n)] + 4\mu^2 Tr(\mathbf{PR}\boldsymbol{B}(n)\mathbf{RP})$$
$$+ 2\mu^2 Tr(\mathbf{R}\boldsymbol{B}(n))Tr(\mathbf{P}Diag(\mathbf{PRP})) + 2\mu^2\hat{W}^H\mathbf{R}\hat{W}Tr(\mathbf{P}Diag(\mathbf{PRP})) \tag{c.58}$$

using Lemma A.1 and A.2(c), (c.58) becomes

$$Tr[\boldsymbol{B}(n+1)] \le Tr[\boldsymbol{B}(n)]$$
$$+ [-4\mu + 2\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP})) + 4\mu^2\lambda_{max}]Tr(\mathbf{R}\boldsymbol{B}(n))$$
$$+ 2\mu^2\hat{W}^H\mathbf{R}\hat{W}Tr(\mathbf{P}Diag(\mathbf{PRP})) \tag{c.59}$$

Taking the limit as $n \to \infty$ if (c.36) is satisfied then

$$\lim_{n\to\infty} Tr(\boldsymbol{B}(n+1)) = \lim_{n\to\infty} Tr(\boldsymbol{B}(n)) = q < \infty \tag{c.60}$$

By applying Lemmas A.1 and A.2 (c.59) yields

$$\lim_{n\to\infty} E[V^H(n)\mathbf{R}V(n)] \le \frac{\mu\hat{W}^H\mathbf{R}\hat{W}Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu[Tr(\mathbf{P}Diag(\mathbf{PRP})) + 2\lambda_{max}]} \tag{c.61}$$

and similarly by applying Lemmas A.1 and A.2

$$\lim_{n\to\infty} E[V^H(n)\mathbf{R}V(n)] \ge \frac{\mu\hat{W}^H\mathbf{R}\hat{W}Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu[Tr(\mathbf{P}Diag(\mathbf{PRP}))]} \tag{c.62}$$

Since the norm of the weight error vector exists *Result 3.5.a* can be derived using (c.61), (c.62), (3.56) and (3.57).

Derivation of Result 3.5.b.

Taking the trace of (c.22), substituting (c.17) and using Lemma A.2

$$\begin{aligned} Tr[\boldsymbol{B}(n+1)] &= Tr[\boldsymbol{B}(n)] - 4\mu Tr[\mathbf{R}\boldsymbol{B}(n)] + 4\mu^2 Tr(\mathbf{P}\mathbf{R}\boldsymbol{B}(n)\mathbf{R}\mathbf{P}) \\ &+ 2\mu^2 Tr(\mathbf{R}\boldsymbol{B}(n))Tr(\mathbf{P}Diag(\mathbf{PRP})) \\ &+ 2\mu^2\gamma^2 L Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2) + 2\mu^2\hat{\boldsymbol{W}}^H\mathbf{R}\hat{\boldsymbol{W}}Tr(\mathbf{P}Diag(\mathbf{PRP})) \end{aligned} \quad \text{(c.63)}$$

using Lemma A.2(c), (c.63) becomes

$$\begin{aligned} Tr[\boldsymbol{B}(n+1)] &\le Tr[\boldsymbol{B}(n)] \\ &+ [-4\mu + 2\mu^2 Tr(\mathbf{P}Diag(\mathbf{PRP})) + 4\mu^2\lambda_{max}]Tr(\mathbf{R}\boldsymbol{B}(n)) \\ &+ 2\mu^2\gamma^2 L Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2) + 2\mu^2\hat{\boldsymbol{W}}^H\mathbf{R}\hat{\boldsymbol{W}}Tr(\mathbf{P}Diag(\mathbf{PRP})) \end{aligned} \quad \text{(c.64)}$$

Taking the limit as $n \to \infty$ if the conditions for convergence of the ssemsp, Theorem A.2, (c.36), are satisfied then

$$\lim_{n \to \infty} Tr(\boldsymbol{B}(n+1)) = \lim_{n \to \infty} Tr(\boldsymbol{B}(n)) = q < \infty \quad \text{(c.65)}$$

By applying Lemmas A.1 and A.2 (c.64) yields

$$\lim_{n \to \infty} E[\boldsymbol{V}^H(n)\mathbf{R}\boldsymbol{V}(n)] \le \frac{\mu\gamma^2 L Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2) + \mu\hat{\boldsymbol{W}}^H\mathbf{R}\hat{\boldsymbol{W}}Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu[Tr(\mathbf{P}Diag(\mathbf{PRP})) + 2\lambda_{max}]} \quad \text{(c.66)}$$

and similarly

$$\lim_{n \to \infty} E[\boldsymbol{V}^H(n)\mathbf{R}\boldsymbol{V}(n)] \ge \frac{\mu\gamma^2 L Tr(\mathbf{P}(Diag(\mathbf{PRP}))^2) + \mu\hat{\boldsymbol{W}}^H\mathbf{R}\hat{\boldsymbol{W}}Tr(\mathbf{P}Diag(\mathbf{PRP}))}{2 - \mu[Tr(\mathbf{P}Diag(\mathbf{PRP}))]} \quad \text{(c.67)}$$

As the bounds for the ssemsp exists *Result 3.5.b* can be derived using (c.66), (c.67), (3.56) and (3.57).

Derivation of Result 3.5.c

Taking the trace of (c.22), substituting (c.18) and using Lemma A.2

$$Tr[\boldsymbol{B}(n+1)] = Tr[\boldsymbol{B}(n)] - 4\mu Tr[\mathbf{R}\boldsymbol{B}(n)] + 4\mu^2 Tr(\mathbf{P}\mathbf{R}\boldsymbol{B}(n)\mathbf{R}\mathbf{P})$$
$$+ a\mu^2 Tr(\mathbf{R}\boldsymbol{B}(n)) Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P})) + a\mu^2 \hat{\boldsymbol{W}}^H \mathbf{R} \hat{\boldsymbol{W}} Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P}))$$
$$+ 2\mu^2 Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\boldsymbol{B}(n)\mathbf{R}\mathbf{P})) \quad \text{(c.68)}$$

using Lemma A.1, A.2(c), (c.68) becomes

$$Tr[\boldsymbol{B}(n+1)] \le Tr[\boldsymbol{B}(n)] + a\mu^2 \hat{\boldsymbol{W}}^H \mathbf{R} \hat{\boldsymbol{W}} Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P}))$$
$$+ [-4\mu + a\mu^2 Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P})) + 6\mu^2\lambda_{max}] Tr[\mathbf{R}\boldsymbol{B}(n)] \quad \text{(c.69)}$$

Taking the limit as $n \to \infty$ if the conditions for convergence of ssemsp, TheoremA.2, (c.52) are satisfied then

$$\lim_{n \to \infty} Tr(\boldsymbol{B}(n+1)) = \lim_{n \to \infty} Tr(\boldsymbol{B}(n)) = q < \infty \quad \text{(c.70)}$$

By applying Lemmas A.1 and A.2 (c.69) yields

$$\lim_{n \to \infty} E[V^H(n)\mathbf{R}V(n)] \le \frac{a\mu \hat{\boldsymbol{W}}^H \mathbf{R} \hat{\boldsymbol{W}} Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P}))}{4 - \mu[aTr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P})) + 6\lambda_{max}]} \quad \text{(c.71)}$$

and similarly

$$\lim_{n \to \infty} E[V^H(n)\mathbf{R}V(n)] \ge \frac{a\mu \hat{\boldsymbol{W}}^H \mathbf{R} \hat{\boldsymbol{W}} Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P}))}{4 - a\mu Tr(\mathbf{P}Diag(\mathbf{P}\mathbf{R}\mathbf{P}))} \quad \text{(c.72)}$$

Since bounds on the ssemsp exists *Result 3.5.c* can be derived using (c.71), (c.72), (3.56) and (3.57).

Appendix D

This Appendix contains the derivation of intermediate results for the New Bounds analysis contained in Chapter 3.

Derivation of (3.73), (3.74) and (3.75)

Re-arranging the weight covariance matrix and using

$$\mathbf{P}\mathbf{K}_{WW}(n)\mathbf{P} = \mathbf{P}\mathbf{K}_{WW}(n) = \mathbf{K}_{WW}(n)\mathbf{P} = \mathbf{K}_{WW}(n) \tag{d.1}$$

gives

$$\mathbf{K}_{WW}(n+1) = (\mathbf{I}_{LL} - 2\mu\mathbf{PRP})\mathbf{K}_{WW}(n) - (\mathbf{I}_{LL} - 2\mu\mathbf{PRP})2\mu\mathbf{K}_{WW}(n)\mathbf{PRP} + \mu^2\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P}$$

$$\Rightarrow \mathbf{K}_{WW}(n+1) = (\mathbf{I}_{LL} - 2\mu\mathbf{PRP})(\mathbf{K}_{WW} - 2\mu\mathbf{K}_{WW}\mathbf{PRP}) + \mu^2\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P}$$

$$\Rightarrow \mathbf{K}_{WW}(n+1) = (\mathbf{I}_{LL} - 2\mu\mathbf{PRP})\mathbf{K}_{WW}(\mathbf{I}_{LL} - 2\mu\mathbf{PRP}) + \mu^2\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P} \tag{d.2}$$

For the three receiver structures the last term in (d.2) has been evaluated previously in (3.25), (3.26) and (3.27). It can be represented generically by

$$\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P} = gTr(\mathbf{K}_{WW}(n)\mathbf{R})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + b\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \mathbf{D} \tag{d.3}$$

where g,b,**D** are defined as

1. Dual Receiver Dual Perturbation System

g=2, b=2$\overline{\mathbf{W}}^H(n)\mathbf{R}\overline{\mathbf{W}}(n)$, $\mathbf{D}$=0

2. Dual Receiver Reference Receiver System

g=2, b=2$\overline{\mathbf{W}}^H(n)\mathbf{R}\overline{\mathbf{W}}(n)$, $\mathbf{D} = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P}$

3. Single Receiver System

$g = \left(c + \frac{1}{c}\right)^2$, c is defined in (3.17)

$b = \left(c + \frac{1}{c}\right)^2 \overline{\mathbf{W}}^H(n)\mathbf{R}\overline{\mathbf{W}}(n)$, $\mathbf{D} = 2\mathbf{P}Diag(\mathbf{PRR}_{WW}\mathbf{RP})\mathbf{P}$

Substituting (d.3) into (d.2) and re-arranging

$$\mathbf{K}_{WW}(n+1) = \mathbf{P}\mathbf{K}_{WW}(n)\mathbf{P} - 2\mu\mathbf{P}[\mathbf{R}\mathbf{K}_{WW}(n) + \mathbf{K}_{WW}(n)\mathbf{R}]\mathbf{P} + 4\mu^2\mathbf{PR}\mathbf{K}_{WW}(n)\mathbf{RP} + \mu^2 aTr(\mathbf{K}_{WW}(n)\mathbf{R})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \mu^2 b\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \mu^2\mathbf{D} \tag{d.4}$$

Now assuming that the conditions for convergence of the weight covariance matrix as established in Section 3.4 are satisfied, the trace of the weight covariance matrix in the limit as $n \to \infty$ can be taken and is given by

$$\lim_{n \to \infty} [Tr(\mathbf{K}_{WW}(n+1))] = \lim_{n \to \infty} [Tr(\mathbf{PK}_{WW}(n)\mathbf{P}) - 4\mu Tr(\mathbf{PRK}_{WW}(n))$$
$$+ 4\mu^2 Tr(\mathbf{PRK}_{WW}(n)\mathbf{RP}) + \mu^2 a Tr(\mathbf{K}_{WW}(n)\mathbf{R}) Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P})$$
$$+ \mu^2 b Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \mu^2 Tr(\mathbf{D})] \quad \text{(d.5)}$$

Considering the term $Tr(\mathbf{PRK}_{WW}(n)\mathbf{RP})$, using Lemma A.2(c) and (d.1)

$$Tr(\mathbf{PRK}_{WW}(n)\mathbf{RP}) \le \lambda_{max}(\mathbf{PR}) Tr(\mathbf{PRK}_{WW}(n))$$

$$\Rightarrow Tr(\mathbf{RPRK}_{WW}(n)) \le \lambda_{max}(\mathbf{PR}) Tr(\mathbf{RK}_{WW}(n)) \quad \text{(d.6)}$$

Substituting (d.6) into (d.5) and using

$\lim_{n \to \infty} [\mathbf{K}_{WW}(n+1)] = \lim_{n \to \infty} [\mathbf{K}_{WW}(n)]$ and rearranging, an upper bound for the $\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}]$ can be established and is given by

$$\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}] \le \frac{\lim_{n \to \infty} [\mu b Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \mu Tr(\mathbf{D})]}{4 - a\mu Tr(\mathbf{P}Diag(\mathbf{PRP})) - 4\mu\lambda_{max}(\mathbf{PR})} \quad \text{(d.7)}$$

and similarly

$$\lim_{n \to \infty} Tr[\mathbf{K}_{WW}(n)\mathbf{R}] \ge \frac{\lim_{n \to \infty} [\mu b Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \mu Tr(\mathbf{D})]}{4 - a\mu Tr(\mathbf{P}Diag(\mathbf{PRP}))} \quad \text{(d.8)}$$

(d.7) and (d.8) are (3.73) and (3.74) respectively. When μ is small it is easily observed that the upper and lower bounds approach eachother hence the asymptotic misadjustment for a suitably small gradient step size can be expressed as

$$M = \frac{\lim_{n \to \infty} [\mu b Tr(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \mu Tr(\mathbf{D})]}{4\hat{W}^H \mathbf{R} \hat{W}} \quad \text{(d.9)}$$

(d.9) is (3.75).

Solving the Weight Covariance Matrix:

By rearranging (d.2) and using

$\mathbf{B} = (\mathbf{I}_{LL} - 2\mu\mathbf{PRP})$ the weight covariance matrix can be expressed as

$$\mathbf{K}_{WW}(n+1) = \mathbf{B}\mathbf{K}_{WW}(n)\mathbf{B} + \mu^2\mathbf{P}E[\mathbf{V}_G(W(n))]\mathbf{P} \tag{d.10}$$

substituting (d.3)

$$\mathbf{K}_{WW}(n+1) = \mathbf{B}\mathbf{K}_{WW}(n)\mathbf{B} + aTr(\mathbf{K}_{WW}(n)\mathbf{R})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + b\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \mathbf{D} \tag{d.11}$$

From this expression a solution for the weight covariance matrix at the n^{th} update can be determined it is given by

$$\mathbf{K}_{WW}(n) = \mathbf{B}^n\mathbf{K}_{WW}(0)\mathbf{B}^n + \sum_{i=1}^{n-1}\mathbf{B}^{n-i}\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}\mathbf{B}^{n-i}(aTr(\mathbf{K}_{WW}(i-1)\mathbf{R}) + b) + \sum_{i=1}^{n-1}\mathbf{B}^{n-i}\mathbf{D}\mathbf{B}^{n-i} \qquad \text{for } n \geq 2 \tag{d.12}$$

Appendix E

This Appendix contains the derivation of the gradient covariance expression and the misadjustment analysis contained in Chapter 4. In this appendix to simplify the notation expressions such as *W(n)*, *X(l+i)* and $E[G(\boldsymbol{W})|\boldsymbol{W}]$ may be abbreviated to ***W***, ***X*** and $E[G(\boldsymbol{W})]$ respectively. Where this occurs it is assumed that the reduced notation is obvious.

Firstly, for each receiver structure, we determine the expected values of the gradient with the two quantisation methods. The gradient covariance is then determined.

A. Expected Gradient Estimate, Dual Receiver Dual Perturbation System

Quantisation Method 1.

Considering the dual receiver dual perturbation system, and applying Quantisation *Method 1* to the weights, (2.38) and (2.39), the quantised weights are given by

$$Q(\boldsymbol{W}_p, i) = Q(\boldsymbol{W}, i) + Q(\gamma\delta_p(i)) = \boldsymbol{W}(n) + \eta + \gamma\delta_p(i) + \eta_2(i) \tag{e.1}$$

and

$$Q(\boldsymbol{W}_m, i) = Q(\boldsymbol{W}, i) - Q(\gamma\delta_p(i)) = \boldsymbol{W}(n) + \eta - \gamma\delta_p(i) - \eta_2(i) \tag{e.2}$$

where η corresponds to quantisation error vector for $\boldsymbol{W}(n)$ and η_2 corresponds to the quantised error vector of the perturbation sequence.

Substituting (e.1), (e.2) into the gradient expression, (2.40), gives

$$G_1(\boldsymbol{W}_q(n)) = \frac{1}{2\gamma m}\sum_{i=1}^{m}[(\boldsymbol{W} + \eta + \gamma\delta_p(i) + \eta_2(i))\boldsymbol{X}\boldsymbol{X}^H(\boldsymbol{W} + \eta + \gamma\delta_p(i) + \eta_2(i))$$
$$-(\boldsymbol{W} + \eta - \gamma\delta_p(i) - \eta_2(i))^H\boldsymbol{X}\boldsymbol{X}^H(\boldsymbol{W} + \eta - \gamma\delta_p(i) - \eta_2(i))]\delta_p(i)$$
$$= \frac{1}{2\gamma m}\sum_{i=1}^{m}2(\boldsymbol{W}^H\boldsymbol{X}\boldsymbol{X}^H\gamma\delta_p(i) + \gamma\delta_p^H(i)\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{W})\delta_p(i)$$
$$+\frac{1}{2\gamma m}\sum_{i=1}^{m}2(\boldsymbol{W}^H\boldsymbol{X}\boldsymbol{X}^H\eta_2(i) + \eta_2^H(i)\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{W})\delta_p(i)$$
$$+\frac{1}{2\gamma m}\sum_{i=1}^{m}2(\eta^H\boldsymbol{X}\boldsymbol{X}^H\gamma\delta_p(i) + \gamma\delta_p^H(i)\boldsymbol{X}\boldsymbol{X}^H\eta)\delta_p(i)$$
$$+\frac{1}{2\gamma m}\sum_{i=1}^{m}2(\eta^H\boldsymbol{X}\boldsymbol{X}^H\eta_2(i) + \eta_2^H(i)\boldsymbol{X}\boldsymbol{X}^H\eta)\delta_p(i) \tag{e.3}$$

Taking the conditional expectation, with respect to ***X***, of both sides of (e.3) given ***W***(*n*) and using the *Assumptions 4.1-4.3* and applying Lemma A.8

$$E[\eta^H\eta_2(i)] = 0 \tag{e.4}$$

$$E[G_1(\boldsymbol{W}_q(n))|\boldsymbol{W}(n)] = 2\mathbf{PR}\boldsymbol{W}(n) + 2\mathbf{R}E[\eta|\boldsymbol{W}(n)] \tag{e.5}$$

If the expectation is extended to occur over all $\boldsymbol{W}$ an additional assumption may be made that the quantised error vector for the weight vector has zero mean then

$$E[\boldsymbol{G}_1(\boldsymbol{W}_q(n)) | \boldsymbol{W}(n)] = 2\mathbf{PR}\boldsymbol{W}(n) \tag{e.6}$$

This last assumption is equivalent to assuming there is no quantisation error for the weight and isolates the effects of quantisation on the perturbation sequence.
The assumption is weak in the sense that quantisation errors are fixed at system equilibrium and are not treatable as random.

Quantisation Method 2.

Using quantisation *Method 2*, the array weights are given by

$$Q(\boldsymbol{W}_p, i) = Q(\boldsymbol{W} + \gamma\delta_p(i)) = \boldsymbol{W}(n) + \gamma\delta_p(i) + \eta_1(i) \tag{e.7}$$

and

$$Q(\boldsymbol{W}_m, i) = Q(\boldsymbol{W} - \gamma\delta_p(i)) = \boldsymbol{W}(n) - \gamma\delta_p(i) + \eta_2(i) \tag{e.8}$$

substituting the quantised values for the weights into the expression for the gradient estimate, (2.40), gives

$$\begin{aligned} \boldsymbol{G}_1(\boldsymbol{W}_q(n)) = & \frac{1}{2\gamma m}\sum_{i=1}^{m}[2(\boldsymbol{W}^H\boldsymbol{X}\boldsymbol{X}^H\gamma\delta_p(i) + \gamma\delta_p^H(i)\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{W})]\delta_p(i) \\ & + \frac{1}{2\gamma m}\sum_{i=1}^{m}(\boldsymbol{W}^H\boldsymbol{X}\boldsymbol{X}^H\eta_1(i) + \eta_1^H(i)\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{W} + \eta_1^H(i)\boldsymbol{X}\boldsymbol{X}^H\eta_1(i))\delta_p(i) \\ & + \frac{1}{2\gamma m}\sum_{i=1}^{m}(-(\boldsymbol{W}^H\boldsymbol{X}\boldsymbol{X}^H\eta_2(i) + \eta_2^H(i)\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{W} + \eta_2^H(i)\boldsymbol{X}\boldsymbol{X}^H\eta_2(i)))\delta_p(i) \\ & + \frac{1}{2\gamma m}\sum_{i=1}^{m}(\delta_p^H(i)\boldsymbol{X}\boldsymbol{X}^H\eta_1(i) + \eta_1^H(i)\boldsymbol{X}\boldsymbol{X}^H\delta_p(i))\delta_p(i) \\ & + \frac{1}{2\gamma m}\sum_{i=1}^{m}(\delta_p^H(i)\boldsymbol{X}\boldsymbol{X}^H\eta_2(i) + \eta_2^H(i)\boldsymbol{X}\boldsymbol{X}^H\delta_p(i))\delta_p(i) \end{aligned} \tag{e.9}$$

taking the conditional expectation, with respect to $\boldsymbol{X}$, of both sides of (e.9) given $\boldsymbol{W}(n)$, applying Lemma A.8 and *Assumptions 4.1-4.3* gives

$$E[\boldsymbol{G}_1(\boldsymbol{W}_q(n)) | \boldsymbol{W}(n)] = 2\mathbf{PR}\boldsymbol{W}(n) + \sigma_{\eta 1}^2\mathbf{R}\sum_{i=1}^{m}\delta_p(i) - \sigma_{\eta 2}^2\mathbf{R}\sum_{i=1}^{m}\delta_p(i) \tag{e.10}$$

now as the two receivers are quantised in the same fashion

$$\sigma_{\eta 1}^2 = \sigma_{\eta 2}^2 \tag{e.11}$$

then

$$E[G_1(W_q(n))|W(n)] = 2\mathbf{PR}W(n) \tag{e.12}$$

B. Expected Gradient Estimate, Dual Receiver with Reference Receiver System.

Quantisation Method 1.

Using *Method 1* the quantisation of the weights can be represented as

$$Q(W_p, i) = Q(W, i) + Q(\gamma\delta_p(i)) = W(n) + \eta + \gamma\delta_p(i) + \eta_2(i) \tag{e.13}$$

and

$$Q(W, i) = W(n) + \eta \tag{e.14}$$

where η corresponds to quantisation error vector for $W(n)$ and η_2 corresponds to the quantised error vector of the perturbation sequence. Substituting (e.13) and (e.14) into the expression for the gradient estimate, (2.44), and expanding gives

$$\begin{aligned}
G_2(W_q(n)) = {} & \frac{1}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^HW + W^HXX^H\delta_p(i))\delta_p(i) \\
& + \frac{\gamma}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^H\delta_p(i))\delta_p(i) \\
& + \frac{1}{\gamma m}\sum_{i=1}^{m}((\eta_2^H(i)XX^HW + W^HXX^H\eta_2(i))\delta_p(i) \\
& \qquad + (\eta_2^H(i)XX^H\eta + \eta^HXX^H\eta_2(i))\delta_p(i) \\
& \qquad + \gamma(\delta_p^H(i)XX^H\eta(i) + \eta^H(i)XX^H\delta_p(i))\delta_p(i)) \\
& + \frac{1}{\gamma m}\sum_{i=1}^{m}(\eta_2^H(i)XX^H\eta_2(i) + \gamma\delta_p^H(i)XX^H\eta_2(i) + \gamma\eta_2^H(i)XX^H\delta_p(i))\delta_p(i)
\end{aligned} \tag{e.15}$$

Examining equation (e.15) it can be observed that the first two terms correspond to the original gradient estimate with no quantisation effects, the second term being the gradient bias. The third and fourth terms are additional quantisation bias effects.

Taking the conditional expectation of (e.15) given $W(n)$ and taking the expectation over X and η separately as they are assumed to be independent processes, applying Lemma A.3, A.8 and using *Assumptions 4.1-4.3*

$$E[G_2(W_q(n))|W(n)] = 2\mathbf{PR}W + \frac{\gamma}{m}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))\delta_p(i) + \frac{\sigma_{\eta 2}^2}{\gamma m}\sum_{i=1}^{m}\mathbf{R}\delta_p(i) \quad (e.16)$$

Assuming an odd length sequence is used the contribution of the quantisation components to the gradient estimate is zero, by (a.11). Hence

$$E[G_2(W(n))|W(n)] = 2\mathbf{PR}W(n) \quad (e.17)$$

Quantisation Method 2.

When the weights of the dual receiver system with reference receiver are quantised using *Method 2*, the quantised weights can be represented by

$$Q(W_p, i) = Q(W + \gamma\delta_p(i)) = W(n) + \gamma\delta_p(i) + \eta_1(i) \quad (e.18)$$

and

$$Q(W, i) = Q(W, i) = W(n) + \eta_2 \quad (e.19)$$

substituting these quantised weight expressions into the expression for the gradient estimate, (2.44), and expanding gives

$$\begin{aligned} G_2(W_q(n)) = {} & \frac{1}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^HW + W^HXX^H\delta_p(i))\delta_p(i) \\ & + \frac{\gamma}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^H\delta_p(i))\delta_p(i) \\ & + \frac{1}{\gamma m}\sum_{i=1}^{m}((\eta_1^H(i)XX^HW + W^HXX^H\eta_1(i))\delta_p(i) \\ & \qquad + \gamma(\eta_1^H(i)XX^H\delta_p(i) + \delta_p^H(i)XX^H\eta_1(i))\delta_p(i) \\ & \qquad + (\eta_1^H(i)XX^H\eta_1(i))\delta_p(i)) \\ & - \frac{1}{\gamma m}\sum_{i=1}^{m}(\eta_2^HXX^HW + W^HXX^H\eta_2 + \eta_2^HXX^H\eta_2)\delta_p(i) \end{aligned} \quad (e.20)$$

Examining (e.20) it can be observed that the first two terms are the usual gradient estimate and the third and fourth terms are the quantisation error effects from receivers one and two respectively.

Taking the conditional expectation of (e.20) given $W(n)$ and taking the expectation over X and η separately as they are assumed to be independent processes, applying Lemma A.3, A.8 and using *Assumptions 4.1-4.3*

$$E[G_2(W_q(n))|W(n)] = 2\mathbf{PR}W + \frac{\gamma}{m}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i))\delta_p(i) + \frac{\sigma_{\eta 1}^2}{\gamma m}\sum_{i=1}^{m}\mathbf{R}\delta_p(i)$$

$$-\frac{1}{\gamma m}\sum_{i=1}^{m} E(\eta_2^H \mathbf{R} W + W^H \mathbf{R}\eta_2 + \eta_2^H \mathbf{R}\eta_2)\delta_p(i) \tag{e.21}$$

Assuming an odd length zero mean sequence is used it can be observed that the additional quantisation biases are equal to zero, hence the gradient estimate is unbiased, $E[G_2(W_q(n)) | W(n)] = 2\mathbf{PR}W(n)$

C. Expected Gradient Estimate, Single Receiver System

Quantisation Method 1.

When the weights of the single receiver are quantised according to *Method1* they can be represented by

$$Q(W_p, i) = Q(W, i) + Q(\gamma\delta_p(i)) = W(n) + \eta + \gamma\delta_p(i) + \eta_2(i) \tag{e.22}$$

η and η_2 are the quantisation errors for the weight and the perturbation step size respectively. Substituting (e.22) into the expression for the gradient estimate, (2.49), and expanding

$$\begin{aligned}
G_3(W_q(n)) = {} & \frac{1}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^H W + W^H XX^H \delta_p(i))\delta_p(i) \\
& + \frac{\gamma}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^H \delta_p(i) + W^H XX^H W)\delta_p(i) \\
& + \frac{1}{\gamma m}\sum_{i=1}^{m}((\eta_2^H(i)XX^H W + W^H XX^H \eta_2(i))\delta_p(i) \\
& \qquad + (\eta_2^H(i)XX^H \eta + \eta^H XX^H \eta_2(i))\delta_p(i) \\
& \qquad + (\eta^H(i)XX^H W + W^H XX^H \eta(i))\delta_p(i) \\
& \qquad + \gamma(\delta_p^H(i)XX^H \eta(i) + \eta^H(i)XX^H \delta_p(i))\delta_p(i)) \\
& + \frac{1}{\gamma m}\sum_{i=1}^{m}(\eta_2^H(i)XX^H \eta_2(i) + \gamma\delta_p^H(i)XX^H \eta_2(i) + \gamma\eta_2^H(i)XX^H \delta_p(i))\delta_p(i) \\
& + \frac{1}{\gamma m}\sum_{i=1}^{m}(\eta_1^H(i)XX^H \eta_1(i))\delta_p(i)
\end{aligned} \tag{e.23}$$

Examining (e.23) it can be observed that the first two terms are the normal gradient estimate with the second term being the gradient bias. By similar analysis to the dual receiver cases the error contributions of the quantisation error in the single receiver system also sum to zero. Hence

$$E[G_3(W_q(n))|W(n)] = 2\mathbf{PR}W(n) \tag{e.24}$$

Quantisation Method 2.

When the weights of the single receiver system are quantised using *Method 2*, given by

$$Q(W_p, i) = Q(W + \gamma\delta_p(i)) = W(n) + \gamma\delta_p(i) + \eta_1(i) \tag{e.25}$$

where η_1 represents the total quantisation error. Substituting (e.25) into the expression for the gradient estimate, (2.49), and expanding gives

$$\begin{aligned} G_3(W_q(n)) = {} & \frac{1}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^HW + W^HXX^H\delta_p(i))\delta_p(i) \\ & + \frac{\gamma}{m}\sum_{i=1}^{m}(\delta_p^H(i)XX^H\delta_p(i) + W^HXX^HW)\delta_p(i) \\ & + \frac{1}{\gamma m}\sum_{i=1}^{m}((\eta_1^H(i)XX^HW + W^HXX^H\eta_1(i))\delta_p(i) \\ & \quad + \gamma(\eta_1^H(i)XX^H\delta_p(i) + \delta_p^H(i)XX^H\eta_1(i))\delta_p(i) \\ & \quad + (\eta_1^H(i)XX^H\eta_1(i))\delta_p(i)) \end{aligned} \tag{e.26}$$

Taking the conditional expectation of (e.26) given $W(n)$ and taking the expectation over X and η separately as they are assumed to be independent processes, applying Lemma A.3, A.8 and using *Assumptions 4.1-4.3*

$$E[G_3(W_q(n))] = 2\mathbf{PR}W + \frac{\gamma}{m}\sum_{i=1}^{m}(\delta_p^H(i)\mathbf{R}\delta_p(i) + W^H\mathbf{R}W)\delta_p(i) + \frac{\sigma_{\eta 1}^2}{\gamma m}\sum_{i=1}^{m}\mathbf{R}\delta_p(i) \tag{e.27}$$

Assuming an odd symmetry zero mean perturbation sequence is used the additional quantisation biases are equal to zero, hence the gradient estimate is unbiased $E[G_3(W_q(n))|W(n)] = 2\mathbf{PR}W(n)$.

Derivation of Result 4.4.a. Gradient Covariance Dual Receiver Dual Perturbation

Assuming the quantisation process is performed using *Method 2* the output power sequence of the array can be represented as

$$op_i(W) = d_{1i}(W) + d_{\eta 1i}(W) + d_{\eta 2i}(W) \quad \text{(e.28)}$$

where

$$d_{1i}(W) = 2(W^H XX^H \delta_g(i) + \delta_g^H(i) XX^H W) \quad \text{(e.29)}$$

$$d_{\eta 1i}(W) = W^H XX^H \eta_1(i) + \eta_1^H(i) XX^H W + \delta_g^H(i) XX^H \eta_1(i) + \eta_1^H(i) XX^H \delta_g(i) + \eta_1^H(i) XX^H \eta_1(i) \quad \text{(e.30)}$$

$$d_{\eta 2i}(W) = -(W^H XX^H \eta_2(i) + \eta_2^H(i) XX^H W) (\delta_g^H(i) XX^H \eta_2(i) + \eta_2^H(i) XX^H \delta_g(i)) - \eta_2^H(i) XX^H \eta_2(i) \quad \text{(e.31)}$$

and where $\delta_g(i) = \gamma \delta_p(i)$. (e.32)

By setting $d_i(W) = op_i(W)/2$ in (a.7) one obtains $T = G_1(W_q(n))$. By assumption $\{X(i)\}$ is a sequence of independent random vectors and this implies that (a.9) is satisfied. Applying Lemma A.4 and substituting (e.29)-(e.31) gives

$$\mathbf{V}_{G1}(W_q(n)) = \frac{1}{4\gamma^2 m^2} \sum_{i=1}^{m} \{E[d_i(W) d_i^H(W) | W(n)] - E[d_i(W) | W(n)] E[d_i^H(W) | W(n)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.33)}$$

$$= \frac{1}{4\gamma^2 m^2} \sum_{i=1}^{m} \{E[d_{1i}(W) d_{1i}^H(W) | W(n)] - E[d_{1i}(W) | W(n)] E[d_{1i}^H(W) | W(n)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.34)}$$

$$+ \frac{1}{4\gamma^2 m^2} \sum_{i=1}^{m} \{2E[d_{1i}(W) d_{\eta 1i}^H(W)] + 2E[d_{1i}(W) d_{\eta 2i}^H(W)] + 2E[d_{\eta 1i}(W) d_{\eta 2i}^H(W)] + E[d_{\eta 1i}(W) d_{\eta 1i}^H(W)] + E[d_{\eta 2i}(W) d_{\eta 2i}^H(W)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.35)}$$

$$- \frac{1}{4\gamma^2 m^2} \sum_{i=1}^{m} \{E[d_{1i}(W) | W(n)] E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W)) | W(n)] + E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W)) | W(n)] \times E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W) + d_{1i}^H(W)) | W(n)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.36)}$$

The expressions on the left hand side of the above equation can now be evaluated. Note that the first term, (e.34), is equivalent to the gradient covariance when no quantisation effects are considered.

Evaluating (e.35).

The individual terms of (e.35) are evaluated separately. Using *Assumptions 4.1-4.3*

$$2E[d_{1i}(W)d^H_{\eta 1i}(W)] = 4E[(W^H XX^H \delta_g(i) + \delta^H_g(i)XX^H W)\eta^H_1(i)XX^H\eta_1(i)] \quad \text{(e.37)}$$

Note that it is not necessary to consider terms with an odd number of quantisation error vectors as the expected values of these components will be zero by definition.

Assuming X and η are independent, expectation over X and η can be taken separately.

Taking expectation over X and using Lemma A.3

$$\begin{aligned}2E[d_{1i}(W)d^H_{\eta 1i}(W)] &= 4E[W^H(\mathbf{R}\delta_g(i)\eta^H_1(i)\mathbf{R} + \eta^H_1(i)\mathbf{R}\delta_g(i)\mathbf{R})\eta_1(i) \\ &\quad + \delta^H_g(i)(\mathbf{R}W\eta^H_1(i)\mathbf{R} + \eta^H_1(i)\mathbf{R}W\mathbf{R})\eta(i)] \\ &= 4E[W^H\mathbf{R}\delta_g(i)\eta^H_1(i)\mathbf{R}\eta_1(i) + \eta^H_1(i)\mathbf{R}\delta_g(i)W^H\mathbf{R}\eta_1(i) \\ &\quad + \delta^H_g(i)\mathbf{R}W\eta^H_1(i)\mathbf{R}\eta_1(i) + \eta^H_1(i)\mathbf{R}W\delta^H_g(i)\mathbf{R}\eta(i)] \quad \text{(e.38)}\end{aligned}$$

taking expectation over η

$$\begin{aligned}2E[d_{1i}(W)d^H_{\eta 1i}(W)] &= 4\sigma^2_{\eta 1}[Tr(\mathbf{R})W^H\mathbf{R}\delta_g(i) + W^H\mathbf{R}^2\delta_g(i) \\ &\quad + Tr(\mathbf{R})\delta^H_g(i)\mathbf{R}W + \delta^H_g(i)\mathbf{R}^2 W] \quad \text{(e.39)}\end{aligned}$$

similarly

$$\begin{aligned}2E[d_{1i}(W)d^H_{\eta 2i}(W)] &= -4\sigma^2_{\eta 2}[Tr(\mathbf{R})W^H\mathbf{R}\delta_g(i) + W^H\mathbf{R}^2\delta_g(i) \\ &\quad + Tr(\mathbf{R})\delta^H_g(i)\mathbf{R}W + \delta^H_g(i)\mathbf{R}^2 W] \quad \text{(e.40)}\end{aligned}$$

Consider

$$\begin{aligned}E[d_{\eta 1i}(W)d^H_{\eta 1i}(W)] &= E[(\delta^H_g(i)XX^H\eta_1(i) + \eta^H_1(i)XX^H\delta_g(i)) \times \\ &\quad (\delta^H_g(i)XX^H\eta_1(i) + \eta^H_1(i)XX^H\delta_g(i))] \quad \text{(e.41)}\end{aligned}$$

$$+ E[(W^H XX^H\eta_1(i) + \eta^H_1(i)XX^H W)(W^H XX^H\eta_1(i) + \eta^H_1(i)XX^H W)] \quad \text{(e.42)}$$

$$+ 2E[(W^H XX^H\eta_1(i) + \eta^H_1(i)XX^H W)(\delta^H_g(i)XX^H\eta_1(i) + \eta^H_1(i)XX^H\delta_g(i))] \quad \text{(e.43)}$$

$$+ 2E[(W^H XX^H\eta_1(i) + \eta^H_1(i)XX^H W)(\eta^H_1(i)XX^H\eta_1(i))] \quad \text{(e.44)}$$

$$+ 2E[(\delta^H_g(i)XX^H\eta_1(i) + \eta^H_1(i)XX^H\delta_g(i))(\eta^H_1(i)XX^H\eta_1(i))] \quad \text{(e.45)}$$

$$+ E[\eta^H_1(i)XX^H\eta_1(i)\eta^H_1(i)XX^H\eta_1(i)] \quad \text{(e.46)}$$

(e.41)-(e.46) are now evaluated separately.

Applying Lemma A.3 and Theorem A.4 to (e.41) and taking expectation over X and η separately

$$E[(\delta_g^H(i)XX^H\eta_1(i) + \eta_1^H(i)XX^H\delta_g(i))(\delta_g^H(i)XX^H\eta_1(i) + \eta_1^H(i)XX^H\delta_g(i))]$$

$$\begin{aligned} = \; & E[\delta_g^H(i)(\mathbf{R}\eta(i)\delta_g^H(i)\mathbf{R} + \delta_g^H(i)\mathbf{R}\eta(i)\mathbf{R})\eta_1(i) \\ & + \delta_g^H(i)(\mathbf{R}\eta_1(i)\eta_1^H(i)\mathbf{R} + \eta_1^H(i)\mathbf{R}\eta_1(i)\mathbf{R})\delta_g(i) \\ & + \eta_1^H(i)(\mathbf{R}\delta_g(i)\delta_g^H(i)\mathbf{R} + \delta_g^H(i)\mathbf{R}\delta_g(i)\mathbf{R})\eta_1(i) \qquad \text{taking } E \text{ w.r.t. } X \\ & + \eta_1^H(i)(\mathbf{R}\delta_g(i)\eta_1^H(i)\mathbf{R} + \eta_1^H(i)\mathbf{R}\delta_g(i)\mathbf{R})\delta_g(i)] \end{aligned} \qquad (e.47)$$

re-arranging the terms

$$\begin{aligned} = \; & 2E[\delta_g^H(i)\mathbf{R}\eta_1(i)\delta_g^H(i)\mathbf{R}\eta_1(i) + \delta_g^H(i)\mathbf{R}\eta_1(i)\eta_1^H(i)\mathbf{R}\delta_g(i) \\ & + \eta_1^H(i)\mathbf{R}\eta_1(i)\delta_g^H(i)\mathbf{R}\delta_g(i) + \eta_1^H(i)\mathbf{R}\delta_g(i)\eta_1^H(i)\mathbf{R}\delta_g(i)] \end{aligned} \qquad (e.48)$$

taking the expectation over η and using (5.1)

$$= 2\sigma_{\eta 1}^2(\delta_g^H(i)\mathbf{R}^2\delta_g(i) + Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}\delta_g(i)) \qquad (e.49)$$

By similar methods it can be shown that (e.42)

$$= 2\sigma_{\eta 1}^2(W^H\mathbf{R}^2W + Tr(\mathbf{R})W^H\mathbf{R}W) \qquad (e.50)$$

Evaluating (e.46). Taking expectation over X and η separately and applying Lemma A.2, A.3 and Theorem A.4

$$E[\eta_1^H(i)XX^H\eta_1(i)\eta_1^H(i)XX^H\eta_1(i)]$$

$$= E[\eta_1^H(i)(\mathbf{R}\eta_1(i)\eta_i^H(i)\mathbf{R} + \eta_1^H(i)\mathbf{R}\eta_1(i)\mathbf{R})\eta_1(i)] \text{ , taking } E \text{ w.r.t. } X$$

$$= E[2\eta_1^H(i)\mathbf{R}\eta_1(i)\eta_1^H(i)\mathbf{R}\eta_1(i)]$$

$$= 2(Tr(\sigma_{\eta 1}^2\mathbf{R}\sigma_{\eta 1}^2\mathbf{R}) + Tr(\sigma_{\eta 1}^2\mathbf{R})Tr(\sigma_{\eta 1}^2\mathbf{R})) \text{ , taking } E \text{ w.r.t. } \eta$$

$$= 2(\sigma_{\eta 1}^4 Tr(\mathbf{R}^2) + \sigma_{\eta 1}^4(Tr(\mathbf{R}))^2) \qquad (e.51)$$

Evaluating (e.43). Taking expectation over X and using lemma A.3 and Theorem A.4

$$2E[(W^HXX^H\eta_1(i) + \eta_1^H(i)XX^HW)(\delta_g^H(i)XX^H\eta_1(i) + \eta_1^H(i)XX^H\delta_g(i))]$$

$$\begin{aligned} = \; & 2E[W^H(\mathbf{R}\eta_1(i)\delta_g^H(i)\mathbf{R} + \delta_g^H(i)\mathbf{R}\eta_1(i)\mathbf{R})\eta_1(i) \\ & + W^H(\mathbf{R}\eta_1(i)\eta_1^H(i)\mathbf{R} + \eta_1^H(i)\mathbf{R}\eta_1(i)\mathbf{R})\delta_g(i) \\ & + \eta_1^H(i)(\mathbf{R}W\delta_g^H(i)\mathbf{R} + \delta_g^H(i)\mathbf{R}W\mathbf{R})\eta(i) \end{aligned}$$

$$+ \eta_1^H(i)(\mathbf{R}W\eta_1^H(i)\mathbf{R} + \eta_1^H(i)\mathbf{R}W\mathbf{R})\delta_g(i)] \qquad \text{taking } E \text{ w.r.t. } X \quad \text{(e.52)}$$

multiplying the terms and re-arranging

$$\begin{aligned} &= 2E[W^H\mathbf{R}\eta_1(i)\delta_g^H(i)\mathbf{R}\eta_1(i) + \delta_g^H(i)\mathbf{R}\eta_1(i)W^H\mathbf{R}\eta_1(i) \\ &\quad + W^H\mathbf{R}\eta_1(i)\eta_1^H(i)\mathbf{R}\delta_g(i) + \eta_1^H(i)\mathbf{R}\eta_1(i)W^H\mathbf{R}\delta_g(i) \\ &\quad + \eta_1^H(i)\mathbf{R}W\delta_g^H(i)\mathbf{R}\eta_1(i) + \delta_g^H(i)\mathbf{R}W\eta_1^H(i)\mathbf{R}\eta_1(i) \\ &\quad + 2\eta_1^H(i)\mathbf{R}W\eta_1^H(i)\mathbf{R}\delta_g(i)] \end{aligned} \qquad \text{(e.53)}$$

taking the expectation over η and using Lemmas A.3 and TheoremA.4

$$= 2\sigma_{\eta 1}^2[W^H\mathbf{R}^2\delta_g(i) + Tr(\mathbf{R})W^H\mathbf{R}\delta_g(i) + Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}W + \delta_g^H(i)\mathbf{R}^2W] \quad \text{(e.54)}$$

using Lemma A.3 and Theorem A.4 the terms in (e.44) and (e.45) can be evaluated as:

$$E[W^HXX^H\eta_1(i)\eta_1^H(i)XX^H\eta_1(i)] = 0 \qquad \text{(e.55)}$$

$$E[\delta_g^H(i)XX^H\eta_1(i)\eta_1^H(i)XX^H\eta_1(i)] = 0 \qquad \text{(e.56)}$$

$$E[\eta_1^H(i)XX^H\delta_g(i)\eta_1^H(i)XX^H\eta_1(i)] = 0 \qquad \text{(e.57)}$$

$$E[\eta_1^H(i)XX^HW\eta_1^H(i)XX^H\eta_1(i)] = 0 \qquad \text{(e.58)}$$

substituting (e.49), (e.50), (e.51), (e.54), (e.55)- (e.58) back into the expression for $E[d_{\eta 1i}(W)d_{\eta 1i}^H(W)|W(n)]$ then

$$\begin{aligned} E[d_{\eta 1i}(W)d_{\eta 1i}^H(W)|W(n)] &= 2\sigma_{\eta 1}^2(W^H\mathbf{R}^2W + Tr(\mathbf{R})W^H\mathbf{R}W \\ &\quad + \delta_g^H(i)\mathbf{R}^2\delta_g(i) + Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}\delta_g(i) \\ &\quad + \sigma_{\eta 1}^2(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2) \\ &\quad + Tr(\mathbf{R})(W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W) \\ &\quad + W^H\mathbf{R}^2\delta_g(i) + \delta_g^H(i)\mathbf{R}^2W) \end{aligned} \qquad \text{(e.59)}$$

by similar methods it can be shown that

$$\begin{aligned} E[d_{\eta 2i}(W)d_{\eta 2i}^H(W)|W(n)] &= 2\sigma_{\eta 2}^2(W^H\mathbf{R}^2W + Tr(\mathbf{R})W^H\mathbf{R}W \\ &\quad + \delta_g^H(i)\mathbf{R}^2\delta_g(i) + Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}\delta_g(i) \\ &\quad + \sigma_{\eta 2}^2(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2) \\ &\quad - Tr(\mathbf{R})(W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W) \\ &\quad - W^H\mathbf{R}^2\delta_g(i) - \delta_g^H(i)\mathbf{R}^2W) \end{aligned} \qquad \text{(e.60)}$$

Using the assumed property that quantisation errors between the two receivers are independent of eachother

$$E[\eta_1(i)\eta_2^H(i)\big|W(n)] = E[\eta_2(i)\eta_1^H(i)\big|W(n)] = 0 \quad \text{(e.61)}$$

Evaluating the other terms in (e.35) and using (e.61)

$$2E[d_{\eta 1i}d_{\eta 2i}^H\big|W(n)] = 2E[-\eta_1^H(i)XX^H\eta_1(i)\eta_2^H(i)XX^H\eta_2(i)] \quad \text{(e.62)}$$

taking expectation of (e.62) over X and η separately and using Lemma A.3 and Theorem A.4

$$2E[d_{\eta 1i}d_{\eta 2i}^H\big|W(n)] = -2(\sigma_{\eta 1}^2\sigma_{\eta 2}^2)((Tr(\mathbf{R}))^2 + Tr(\mathbf{R}^2)) \quad \text{(e.63)}$$

Evaluating (e.36)

Using (e.61) and the following

$$E[d_{1i}(W)|W(n)] = 2(W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W) = E[d_{1i}^H(W)|W(n)] \quad \text{(e.64)}$$

$$E[d_{\eta 1i}(W)|W(n)] = E[d_{\eta 1i}^H(W)|W(n)] = \sigma_{\eta 1}^2 Tr(\mathbf{R}) \quad \text{(e.65)}$$

$$E[d_{\eta 2i}(W)|W(n)] = E[d_{\eta 2i}^H(W)|W(n)] = -\sigma_{\eta 2}^2 Tr(\mathbf{R}) \quad \text{(e.66)}$$

then

$$E[d_{1i}(W)|W(n)]E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W))\big|W(n)] =$$
$$2\sigma_{\eta 1}^2 Tr(\mathbf{R})(W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W) - 2\sigma_{\eta 2}^2 Tr(\mathbf{R})(W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W) \quad \text{(e.67)}$$

and

$$E[(d_{\eta 1i}(W) + d_{\eta 2i}(W))|W(n)]E[(d_{1i}^H(W) + d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W))\big|W(n)]$$
$$= (\sigma_{\eta 1}^2 Tr(\mathbf{R}) - \sigma_{\eta 2}^2 Tr(\mathbf{R})) \times$$
$$(\sigma_{\eta 1}^2 Tr(\mathbf{R}) + \sigma_{\eta 2}^2 Tr(\mathbf{R}) + 2(W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W)) \quad \text{(e.68)}$$

using $\sigma_{\eta 1}^2 = \sigma_{\eta 2}^2$, (e.67) and (e.68) are equal to zero.

substituting equations (e.39), (e.40), (e.67), (e.68),(e.59), (e.60), (e.63) back into the expression for the covariance and using $\sigma_{\eta 1}^2 = \sigma_{\eta 2}^2 = \sigma_\eta^2$ the covariance of the gradient estimate is given by

$$\mathbf{V}_{G1}(W_q(n)) = \mathbf{V}_{G1_{orig}}(W(n))$$
$$+ \frac{\sigma_\eta^2}{m^2}\sum_{i=1}^{m}\{\delta_p^H(i)\mathbf{R}^2\delta_p(i) + Tr(\mathbf{R})\delta_p^H(i)\mathbf{R}\delta_p(i)\}\delta_p(i)\delta_p^H(i)$$

$$+\frac{\sigma_\eta^2}{2\gamma^2 m^2}\sum_{i=1}^{m}\{2W^H\mathbf{R}^2W+2Tr(\mathbf{R})W^H\mathbf{R}W+\sigma_\eta^2 Tr(\mathbf{R}^2)+$$

$$\sigma_\eta^2(Tr(\mathbf{R}))^2\}\delta_p(i)\delta_p^H(i) \quad \text{(e.69)}$$

where $\mathbf{V}_{G1_{orig}}(W(n))$ is given by (3.1).

Substituting for the projected perturbation sequence, (2.67), and using Lemma A.6 the covariance of the gradient is given by

$$\mathbf{V}_{G1}(W_q(n)) = \mathbf{V}_{G1_{orig}}(W(n))+\sigma_\eta^2\mathbf{P}(Diag(\mathbf{P}\mathbf{R}^2\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{P}\mathbf{R}\mathbf{P}))\mathbf{P}$$

$$+\frac{\sigma_\eta^2}{\gamma^2 2L}\left(W^H\mathbf{R}^2W+Tr(\mathbf{R})W^H\mathbf{R}W+\frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}^2)+(Tr(\mathbf{R}))^2)\right)\mathbf{P} \quad \text{(e.70)}$$

for a $4L$ length sequence

$$\mathbf{V}_{G1}(W_q(n)) = \mathbf{V}_{G1_{orig}}(W(n))+2\sigma_\eta^2\mathbf{P}(Diag(\mathbf{P}\mathbf{R}^2\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{P}\mathbf{R}\mathbf{P}))\mathbf{P}$$

$$+\frac{\sigma_\eta^2}{\gamma^2 L}\left(W^H\mathbf{R}^2W+Tr(\mathbf{R})W^H\mathbf{R}W+\frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}^2)+(Tr(\mathbf{R}))^2)\right)\mathbf{P} \quad \text{(e.71)}$$

for a $2L$ length sequence

In (e.70) and (e.71), $\mathbf{V}_{G1_{orig}}(W(n))$ is given by (3.8) and (3.9) respectively.

When the non projected sequence is used the additional covariance terms are given by

$$\mathbf{V}_{G1}(W_q(n)) = \mathbf{V}_{G1_{orig}}(W(n))+\sigma_\eta^2(Diag(\mathbf{R}^2)+Tr(\mathbf{R})Diag(\mathbf{R}))$$

$$+\frac{\sigma_\eta^2}{\gamma^2 2L}\left(W^H\mathbf{R}^2W+Tr(\mathbf{R})W^H\mathbf{R}W+\frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}^2)+(Tr(\mathbf{R}))^2)\right)\mathbf{I}_{LL} \quad \text{(e.72)}$$

for a $4L$ length sequence

$$\mathbf{V}_{G1}(W_q(n)) = \mathbf{V}_{G1_{orig}}(W(n))+2\sigma_\eta^2(Diag(\mathbf{R}^2)+Tr(\mathbf{R})Diag(\mathbf{R}))$$

$$+\frac{\sigma_\eta^2}{\gamma^2 L}\left(W^H\mathbf{R}^2W+Tr(\mathbf{R})W^H\mathbf{R}W+\frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}^2)+(Tr(\mathbf{R}))^2)\right)\mathbf{I}_{LL} \quad \text{(e.73)}$$

for a $2L$ length sequence and

In (e.72) and (e.73) $\mathbf{V}_{G1_{orig}}(W(n))$ is given by (f.1) and (f.2) respectively.

This establishes the result.

Derivation of Result 4.4.b, Gradient Covariance Dual Receiver Reference Receiver

Assuming the quantisation process is performed using *Method 2* the output power sequence can be represented as

$$d_i(W) = d_{1i}(W) + d_{\eta 1i}(W) + d_{\eta 2i}(W) \quad \text{(e.74)}$$

where

$$d_{1i}(W) = W^H XX^H \delta_g(i) + \delta_g^H(i) XX^H W + \delta_g^H(i) XX \delta_g(i) \quad \text{(e.75)}$$

$$d_{\eta 1i}(W) = W^H XX^H \eta_1(i) + \eta_1^H(i) XX^H W + \delta_g^H(i) XX^H \eta_1(i) + \eta_1^H(i) XX^H \delta_g(i) + \eta_1^H(i) XX^H \eta_1(i) \quad \text{(e.76)}$$

$$d_{\eta 2i}(W) = -(W^H XX^H \eta_2(i) + \eta_2^H(i) XX^H W + \eta_2^H(i) XX^H \eta_2(i)) \quad \text{(e.77)}$$

By setting $d_i(W)$ = (e.74) in (a.7) one obtains $T = G_2(W_q(n))$. By assumption $\{X(i)\}$ is a sequence of independent random vectors and this implies that (a.9) is satisfied. Applying Lemma A.4 and substituting (e.75)-(e.77) gives

$$V_{G2}(W_q(n)) = \frac{1}{\gamma^2 m^2} \sum_{i=1}^{m} \{E[d_i(W) d_i^H(W) | W(n)] - E[d_i(W) | W(n)] E[d_i^H(W) | W(n)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.78)}$$

$$= \frac{1}{\gamma^2 m^2} \sum_{i=1}^{m} \{E[d_{1i}(W) d_{1i}^H(W) | W(n)] - E[d_{1i}(W) | W(n)] E[d_{1i}^H(W) | W(n)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.79)}$$

$$+ \frac{1}{\gamma^2 m^2} \sum_{i=1}^{m} \{2E[d_{1i}(W) d_{\eta 1i}^H(W)] + 2E[d_{1i}(W) d_{\eta 2i}^H(W)] + 2E[d_{\eta 1i}(W) d_{\eta 2i}^H(W)] + E[d_{\eta 1i}(W) d_{\eta 1i}^H(W)] \quad \text{(e.80)}$$

$$+ E[d_{\eta 2i}(W) d_{\eta 2i}^H(W)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.81)}$$

$$- \frac{1}{\gamma^2 m^2} \sum_{i=1}^{m} \{E[d_{1i}(W) | W(n)] E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W)) | W(n)] + E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W)) | W(n)] \times E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W) + d_{1i}^H(W)) | W(n)]\} \delta_p(i) \delta_p^H(i) \quad \text{(e.82)}$$

The expressions on the left hand side of the above equation can now be evaluated. Note that the first term, (e.79), is equivalent to the covariance of the gradient estimate when no quantisation effects are considered.

Two approaches can be taken here to determine the gradient covariance. The first is to consider the quantisation errors on the reference receiver to be modelled the same as the perturbed receiver and the second is to assume that the quantisation error on the reference receiver is constant over the gradient estimation period.

Approach 1

Evaluating (e.80)

Using the assumed properties of the quantisation error vectors and definitions

$$2E[d_{1i}(W)d_{\eta 1i}^H(W)]$$
$$= 2E[(W^H XX^H \delta_g(i) + \delta_g^H(i)XX^H W)\eta_1^H(i)XX^H\eta_1(i)]$$
$$+ 2E[\delta_g^H(i)XX^H\delta_g(i)\eta_1^H(i)XX^H\eta_1(i)] \qquad \text{(e.83)}$$

The first term in (e.83) has been evaluated previously in (e.37)-(e.39), the second term is derived below. Applying lemmas A.3 and Theorem A.4 and taking the expected value over X and η separately

$$2E[\delta_g^H(i)XX^H\delta_g(i)\eta_1^H(i)XX^H\eta_1(i)]$$
$$= 2E[\delta_g^H(i)(\mathbf{R}\delta_g(i)\eta_1^H(i)\mathbf{R} + Tr(\mathbf{R}\delta_g(i)\eta_1^H(i))\mathbf{R})\eta_1(i)] \text{ taking } E \text{ w.r.t. } X$$
$$= 2E[\delta_g^H(i)\mathbf{R}\delta_g(i)\eta_1^H(i)\mathbf{R}\eta_1(i) + \eta_1^H(i)\mathbf{R}\delta_g(i)\delta_g^H(i)\mathbf{R}\eta_1(i)] \text{ taking } E \text{ w.r.t. } \eta$$
$$= 2\sigma_{\eta 1}^2\delta_g^H(i)\mathbf{R}\delta_g(i)Tr(\mathbf{R}) + 2\sigma_{\eta 1}^2\delta_g^H(i)\mathbf{R}^2\delta_g(i) \qquad \text{(e.84)}$$

substituting this back into (e.83) and using (e.39)

$$2E[d_{1i}(W)d_{\eta 1i}^H(W)] = 2\sigma_{\eta 1}^2[Tr(\mathbf{R})W^H\mathbf{R}\delta_g(i) + W^H\mathbf{R}^2\delta_g(i)$$
$$+ Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}W + \delta_g^H(i)\mathbf{R}^2W]$$
$$+ 2\sigma_{\eta 1}^2(\delta_g^H(i)\mathbf{R}\delta_g(i)Tr(\mathbf{R}) + \delta_g^H(i)\mathbf{R}^2\delta_g(i)) \qquad \text{(e.85)}$$

By similar methods using the assumption that the reference receiver quantisation errors has the same distribution as the perturbed receiver then

$$2E[d_{1i}(W)d_{\eta 2i}^H(W)] = -2\sigma_{\eta 1}^2[Tr(\mathbf{R})W^H\mathbf{R}\delta_g(i) + W^H\mathbf{R}^2\delta_g(i)$$
$$+ Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}W + \delta_g^H(i)\mathbf{R}^2W]$$
$$-2\sigma_{\eta 2}^2(\delta_g^H(i)\mathbf{R}\delta_g(i)Tr(\mathbf{R}) + \delta_g^H(i)\mathbf{R}^2\delta_g(i)) \qquad \text{(e.86)}$$

As was shown in the dual receiver with dual perturbation case, (e.59)

$$E[d_{\eta 1i}(W)d_{\eta 1i}^H(W)\,|\,W(n)] = 2\sigma_{\eta 1}^2(W^H\mathbf{R}^2W + Tr(\mathbf{R})W^H\mathbf{R}W$$
$$+ \delta_g^H(i)\mathbf{R}^2\delta_g(i) + Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}\delta_g(i)$$
$$+ \sigma_{\eta 1}^2(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)$$

$$+ Tr(\mathbf{R})(W^H \mathbf{R} \delta_g(i) + \delta_g^H(i) \mathbf{R} W)$$
$$+ W^H \mathbf{R}^2 \delta_g(i) + \delta_g^H(i) \mathbf{R}^2 W) \quad \text{(e.87)}$$

$$E[d_{\eta 2i}(W) d_{\eta 2i}^H(W) \big| W(n)]$$
$$= E[(W^H X X^H \eta_2(i) + \eta_2^H(i) X X^H W)(W^H X X^H \eta_2(i) + \eta_2^H(i) X X^H W)]$$
$$+ E[\eta_2^H(i) X X^H \eta_2(i) \eta_2^H(i) X X^H \eta_2(i)] \quad \text{(e.88)}$$

Taking the expectation over X and η separately and applying lemmas A.3 and Theorem A.4, then

$$E[d_{\eta 2i}(W) d_{\eta 2i}^H(W) \big| W(n)]$$
$$= 2\sigma_{\eta 2}^2 (W^H \mathbf{R}^2 W + Tr(\mathbf{R}) W^H \mathbf{R} W + \sigma_{\eta 2}^2 (Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)) \quad \text{(e.89)}$$

As shown in the dual receiver with dual perturbation case, (e.63)

$$2E[d_{\eta 1i} d_{\eta 2i}^H \big| W(n)] = -2(\sigma_{\eta 1}^2 \sigma_{\eta 2}^2)((Tr(\mathbf{R}))^2 + Tr(\mathbf{R}^2)) \quad \text{(e.90)}$$

$$E[\eta_{1i}(W) \eta_{2i}^H(W) \big| W(n)] = E[\eta_2(i) \eta_1^H(i) \big| W(n)] = 0 \quad \text{(e.91)}$$

Evaluating (e.82)

Using (e.61), $\sigma_{\eta 1}^2 = \sigma_{\eta 2}^2$ and the following

$$E[d_{1i}(W) | W(n)] = W^H \mathbf{R} \delta_g(i) + \delta_g^H(i) \mathbf{R} W + \delta_g^H(i) \mathbf{R} \delta_g(i) \quad \text{(e.92)}$$

$$E[d_{\eta 1i}(W) | W(n)] = E[d_{\eta 1i}^H(W) | W(n)] = \sigma_{\eta 1}^2 Tr(\mathbf{R}) \quad \text{(e.93)}$$

$$E[d_{\eta 2i}(W) | W(n)] = E[d_{\eta 2i}^H(W) | W(n)] = -\sigma_{\eta 2}^2 Tr(\mathbf{R}) \quad \text{(e.94)}$$

then

$$E[d_{1i}(W) | W(n)] E[(d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W)) \big| W(n)] +$$
$$E[(d_{\eta 1i}(W) + d_{\eta 2i}(W)) \big| W(n)] \times$$
$$E[(d_{1i}^H(W) + d_{\eta 1i}^H(W) + d_{\eta 2i}^H(W)) \big| W(n)] = 0 \quad \text{(e.95)}$$

substituting equations (e.85), (e.86), (e.87), (e.89), (e.90) and (e.95) back into the expression for the covariance of the gradient estimate and using $\sigma_{\eta 1}^2 = \sigma_{\eta 2}^2 = \sigma_\eta^2$ the covariance of the gradient estimate is given by

$$\mathbf{V}_{G2}(W_q(n)) = \mathbf{V}_{G2_{orig}}(W(n))$$
$$+ \frac{2\sigma_\eta^2}{m^2} \sum_{i=1}^{m} \{\delta_p^H(i) \mathbf{R}^2 \delta_p(i) + Tr(\mathbf{R}) \delta_p^H(i) \mathbf{R} \delta_p(i)\} \delta_p(i) \delta_p^H(i)$$
$$+ \frac{2\sigma_\eta^2}{\gamma^2 m^2} \sum_{i=1}^{m} \{2W^H \mathbf{R}^2 W + 2Tr(\mathbf{R}) W^H \mathbf{R} W + \sigma_\eta^2 Tr(\mathbf{R}^2) + \sigma_\eta^2 (Tr(\mathbf{R}))\} \delta_p(i) \delta_p^H(i)$$

$$+\frac{2\sigma_{\eta}^{2}}{\gamma^{2}m^{2}}\sum_{i=1}^{m}\{Tr(\mathbf{R})(W^{H}\mathbf{R}\delta_{p}(i)+\delta_{p}^{H}(i)\mathbf{R}W)\}\delta_{p}(i)\delta_{p}^{H}(i) \quad \text{(e.96)}$$

$$+\frac{2\sigma_{\eta}^{2}}{\gamma^{2}m^{2}}\sum_{i=1}^{m}\{W^{H}\mathbf{R}^{2}\delta_{p}(i)+\delta_{p}^{H}(i)\mathbf{R}^{2}W\}\delta_{p}(i)\delta_{p}^{H}(i)$$

where $\mathbf{V}_{G2_{orig}}(W(n))$ is given by (3.2)

Substituting for the projected perturbation sequence, (2.67), and using Lemma A.6 the covariance of the gradient is given by

$$\mathbf{V}_{G2}(W_{q}(n)) = \mathbf{V}_{G2_{orig}}(W(n))+2\sigma_{\eta}^{2}\mathbf{P}(Diag(\mathbf{PR}^{2}\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$
$$+\frac{\sigma_{\eta}^{2}}{\gamma^{2}L}(2W^{H}\mathbf{R}^{2}W+2Tr(\mathbf{R})W^{H}\mathbf{R}W+\sigma_{\eta}^{2}Tr(\mathbf{R}^{2})+\sigma_{\eta}^{2}(Tr(\mathbf{R})))\mathbf{P} \quad \text{(e.97)}$$

When the non projected sequence is used, the additional covariance terms are given by.

$$\mathbf{V}_{G2}(W_{q}(n)) = \mathbf{V}_{G2_{orig}}(W(n))+2\sigma_{\eta}^{2}(Diag(\mathbf{R}^{2})+Tr(\mathbf{R})Diag(\mathbf{R}))$$
$$+\frac{\sigma_{\eta}^{2}}{\gamma^{2}L}(2W^{H}\mathbf{R}^{2}W+2Tr(\mathbf{R})W^{H}\mathbf{R}W+\sigma_{\eta}^{2}Tr(\mathbf{R}^{2})+\sigma_{\eta}^{2}(Tr(\mathbf{R})))\mathbf{I}_{LL} \quad \text{(e.98)}$$

Approach 2

In this second approach the quantisation errors on the reference receiver is considered to be constant in the estimation period. It is also assumed that this error is small and can be ignored. The following terms will be affected in the previous analysis.

$$d_{\eta 2i}(W) = 0 \quad \text{(e.99)}$$

$$2E[d_{1i}(W)d_{\eta 2i}^{H}(W)\big|W(n)] = 0 \quad \text{(e.100)}$$

$$E[d_{\eta 2i}(W)d_{\eta 2i}^{H}(W)\big|W(n)] = 0 \quad \text{(e.101)}$$

$$2E[d_{\eta 1i}(W)d_{\eta 2i}^{H}(W)\big|W(n)] = 0 \quad \text{(e.102)}$$

$$E[d_{\eta 2i}(W)|W(n)] = E[d_{\eta 2i}^{H}(W)|W(n)] = 0 \quad \text{(e.103)}$$

$$E[d_{1i}(W)|W(n)]E[(d_{\eta 1i}^{H}(W)+d_{\eta 2i}^{H}(W))\big|W(n)]+$$
$$E[(d_{\eta 1i}(W)+d_{\eta 2i}(W))\big|W(n)]\times$$
$$E[(d_{1i}^{H}(W)+d_{\eta 1i}^{H}(W)+d_{\eta 2i}^{H}(W))\big|W(n)]$$
$$= 2\sigma_{\eta 1}^{2}Tr(\mathbf{R})(W^{H}\mathbf{R}\delta_{g}(i)+\delta_{g}^{H}(i)\mathbf{R}W+\delta_{g}^{H}(i)\mathbf{R}\delta_{g}(i)) \quad \text{(e.104)}$$

Substituting (e.85), (e.87), (e.101), (e.101), (e.102) and (e.104) into the expression for the covariance of the gradient estimate then

$$\mathbf{V}_{G2}(W_q(n)) = \mathbf{V}_{G2_{orig}}(W(n))$$
$$+\frac{\sigma_\eta^2}{m^2}\sum_{i=1}^{m}\{4\delta_p^H(i)\mathbf{R}^2\delta_p(i)+2Tr(\mathbf{R})\delta_p^H(i)\mathbf{R}\delta_p(i)\}\delta_p(i)\delta_p^H(i)$$
$$+\frac{2\sigma_\eta^2}{\gamma^2 m^2}\sum_{i=1}^{m}\{W^H\mathbf{R}^2W+Tr(\mathbf{R})W^H\mathbf{R}W+\sigma_\eta^2 Tr(\mathbf{R}^2)+\sigma_\eta^2(Tr(\mathbf{R}))^2\}\delta_p(i)\delta_p^H(i)$$
$$+\frac{2\sigma_\eta^2}{\gamma^2 m^2}\sum_{i=1}^{m}\{Tr(\mathbf{R})(W^H\mathbf{R}\delta_p(i)+\delta_p^H(i)\mathbf{R}W)\}\delta_p(i)\delta_p^H(i)$$
$$+\frac{2\sigma_\eta^2}{\gamma^2 m^2}\sum_{i=1}^{m}\{W^H\mathbf{R}^2\delta_p(i)+\delta_p^H(i)\mathbf{R}^2W\}\delta_p(i)\delta_p^H(i) \quad \text{(e.105)}$$

where $\mathbf{V}_{G2_{orig}}(W(n))$ is given by (3.2). For an odd symmetry sequence the last two terms in (e.105) sum to zero.

Substituting for the projected perturbation sequence, (2.67), and using Lemma A.5, A.6 the covariance of the gradient is given by

$$\mathbf{V}_{G2}(W_q(n)) = \mathbf{V}_{G2_{orig}}(W(n))+\sigma_\eta^2\mathbf{P}(4Diag(\mathbf{P}\mathbf{R}^2\mathbf{P})+2Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$
$$+\frac{\sigma_\eta^2}{\gamma^2 2L}(W^H\mathbf{R}^2W+Tr(\mathbf{R})W^H\mathbf{R}W+\sigma_\eta^2 Tr(\mathbf{R}^2)+\sigma_\eta^2(Tr(\mathbf{R}))^2)\mathbf{P} \quad \text{(e.106)}$$

When the non projected sequence is used the gradient covariance is given by

$$\mathbf{V}_{G2}(W_q(n)) = \mathbf{V}_{G2_{orig}}(W(n))+\sigma_\eta^2(4Diag(\mathbf{R}^2)+2Tr(\mathbf{R})Diag(\mathbf{R}))$$
$$+\frac{\sigma_\eta^2}{\gamma^2 2L}(W^H\mathbf{R}^2W+Tr(\mathbf{R})W^H\mathbf{R}W+\sigma_\eta^2 Tr(\mathbf{R}^2)+\sigma_\eta^2(Tr(\mathbf{R}))^2)\mathbf{I}_{LL} \quad \text{(e.107)}$$

Derivation of Result 4.4.c, Gradient Covariance Single Receiver System

Assuming the quantisation process is performed using *Method* 2 the output power of receiver 1 defined by (2.50) can be represented as

$$d_i(W) = d_{1i}(W)+d_{\eta 1i}(W) \quad \text{(e.108)}$$

where

$$d_{1i}(W) = W^HXX^H\delta_g(i)+\delta_g^H(i)XX^HW+W^HXX^HW+\delta_g^H(i)XX^H\delta_g(i) \quad \text{(e.109)}$$
$$d_{\eta 1i}(W) = W^HXX^H\eta_1(i)+\eta_1^H(i)XX^HW+\delta_g^H(i)XX^H\eta_1(i)+\eta_1^H(i)XX^H\delta_g(i)$$

$$+\eta_1^H(i)XX^H\eta_1(i) \qquad \text{(e.110)}$$

By setting $d_i(W)$ = (e.108) in (a.7) one obtains $T = G_3(W_q(n))$. By assumption $\{X(i)\}$ is a sequence of independent random vectors and this implies that (a.9) is satisfied. Applying Lemma A.4 and substituting (e.109) and (e.110) gives

$$\mathbf{V}_{G3}(W_q(n)) = \frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}\{E[d_i(W)d_i^H(W)\big|W(n)]$$

$$-E[d_i(W)\big|W(n)]E[d_i^H(W)\big|W(n)]\}\delta_p(i)\delta_p^H(i) \qquad \text{(e.111)}$$

$$= \frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}\{E[d_{1i}(W)d_{1i}^H(W)\big|W(n)]$$

$$-E[d_{1i}(W)\big|W(n)]E[d_{1i}^H(W)\big|W(n)]\}\delta_p(i)\delta_p^H(i) \qquad \text{(e.112)}$$

$$+\frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}\{2E[d_{1i}(W)d_{\eta 1i}^H(W)] + E[d_{\eta 1i}(W)d_{\eta 1i}^H(W)]\}\delta_p(i)\delta_p^H(i) \qquad \text{(e.113)}$$

$$-\frac{1}{\gamma^2 m^2}\sum_{i=1}^{m}\{E[d_{1i}(W)\big|W(n)]E[d_{\eta 1i}^H(W)\big|W(n)]$$

$$+E[d_{\eta 1i}^H(W)\big|W(n)]E[d_{\eta 1i}^H(W) + d_{1i}^H(W)\big|W(n)]\}\delta_p(i)\delta_p^H(i) \qquad \text{(e.114)}$$

The expressions on the left hand side of the above equation can now be evaluated. Note that the first term, (e.112), is equivalent to the gradient covariance when no quantisation effects are considered.
Using the assumed properties of the quantisation error vectors and definitions

Evaluating (e.113)

$$2E[d_{1i}(W)d_{\eta 1i}^H(W)] = 2E[(W^H XX^H\delta_g(i) + \delta_g^H(i)XX^H W)\eta_1^H(i)XX^H\eta_1(i)]$$

$$+2E[\delta_g^H(i)XX^H\delta_g(i)\eta_1^H(i)XX^H\eta_1(i)]$$

$$+2E[W^H XX^H W\eta_1^H(i)XX^H\eta_1(i)] \qquad \text{(e.115)}$$

The first and second term in (e.115) has been evaluated previously in (e.37)-(e.39) and (e.84) respectively. The last term can be derived by applying Lemma A.3, Theorem A.4 and taking the expectant value over X and η separately

$$2E[W^H XX^H W\eta_1^H(i)XX^H\eta_1(i)]$$

$$= 2\sigma_{\eta 1}^2 W^H\mathbf{R}WTr(\mathbf{R}) + 2\sigma_{\eta 1}^2 W^H\mathbf{R}^2 W \qquad \text{(e.116)}$$

using (e.39), (e.84) and (e.116) in (e.115) gives

$$
\begin{aligned}
2E[d_{1i}(W)d_{\eta 1i}^H(W)] &= 2\sigma_{\eta 1}^2[Tr(\mathbf{R})W^H\mathbf{R}\delta_g(i) + W^H\mathbf{R}^2\delta_g(i) \\
&\quad + Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}W + \delta_g^H(i)\mathbf{R}^2 W] \\
&\quad + 2\sigma_{\eta 1}^2(\delta_g^H(i)\mathbf{R}\delta_g(i)Tr(\mathbf{R}) + \delta_g^H(i)\mathbf{R}^2\delta_g(i)) \\
&\quad + 2\sigma_{\eta 1}^2(W^H\mathbf{R}WTr(\mathbf{R}) + W^H\mathbf{R}^2 W) \qquad \text{(e.117)}
\end{aligned}
$$

As was shown for the dual receiver dual perturbation case, (e.59)

$$
\begin{aligned}
E[d_{\eta 1i}(W)d_{\eta 1i}^H(W)\big|W(n)] &= 2\sigma_{\eta 1}^2(W^H\mathbf{R}^2 W + Tr(\mathbf{R})W^H\mathbf{R}W \\
&\quad + \delta_g^H(i)\mathbf{R}^2\delta_g(i) + Tr(\mathbf{R})\delta_g^H(i)\mathbf{R}\delta_g(i) \\
&\quad + \sigma_{\eta 1}^2(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2) \\
&\quad + Tr(\mathbf{R})(W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W) \\
&\quad + W^H\mathbf{R}^2\delta_g(i) + \delta_g^H(i)\mathbf{R}^2 W) \qquad \text{(e.118)}
\end{aligned}
$$

Evaluating (e.114)

Using the following

$$
E[d_{1i}(W)|W(n)] = (W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W + \delta_g^H(i)\mathbf{R}\delta_g(i) + W^H\mathbf{R}W) \qquad \text{(e.119)}
$$

$$
E[d_{\eta 1i}(W)|W(n)] = E[d_{\eta 1i}^H(W)|W(n)] = \sigma_{\eta 1}^2 Tr(\mathbf{R}) \qquad \text{(e.120)}
$$

then

$$
\begin{aligned}
&E[d_{1i}(W)|W(n)]E[(d_{\eta 1i}^H(W))\big|W(n)] + \\
&E[d_{\eta 1i}(W)\big|W(n)]E[d_{1i}^H(W) + d_{\eta 1i}^H(W)\big|W(n)] \\
&\quad = 2(W^H\mathbf{R}W + W^H\mathbf{R}\delta_g(i) + \delta_g^H(i)\mathbf{R}W + \delta_g^H(i)\mathbf{R}\delta_g(i))\sigma_{\eta 1}^2 Tr(\mathbf{R}) \\
&\qquad\qquad + \sigma_{\eta 1}^4(Tr(\mathbf{R}))^2 \qquad \text{(e.121)}
\end{aligned}
$$

substituting (e.117), (e.118), (e.121) into the expression for the gradient covariance then

$$
\begin{aligned}
\mathbf{V}_{G3}(W_q(n)) &= \mathbf{V}_{G3_{orig}}(W(n)) \\
&\quad + \frac{4\sigma_\eta^2}{\gamma^2 m^2}\sum_{i=1}^{m}\{W^H\mathbf{R}^2 W + \gamma^2\delta_p^H(i)\mathbf{R}^2\delta_p(i) + \\
&\qquad \gamma(W^H\mathbf{R}^2\delta_p(i) + \delta_p^H(i)\mathbf{R}^2 W)\}\delta_p(i)\delta_p^H(i)
\end{aligned}
$$

$$+\frac{2\sigma_\eta^2}{\gamma^2 m^2}\sum_{i=1}^{m}\left\{Tr(\mathbf{R})(\gamma^2\delta_p^H(i)\mathbf{R}\delta_p(i)+\gamma(W^H\mathbf{R}\delta_p(i)+\delta_p^H(i)\mathbf{R}W)+W^H\mathbf{R}W)\right.$$

$$\left.\sigma_\eta^2 Tr(\mathbf{R}^2)+\frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}))^2\right\}\delta_p(i)\delta_p^H(i) \qquad \text{(e.122)}$$

where $\mathbf{V}_{G3_{orig}}(W(n))$ is given by (3.5).

Substituting for the projected perturbation sequence, (2.67), and using Lemmas A.5 and A.6 the covariance is given by

$$\begin{aligned}\mathbf{V}_{G3}(W_q(n)) &= \mathbf{V}_{G3_{orig}}(W(n))\\ &+4\sigma_\eta^2\left(\frac{W^H\mathbf{R}^2W}{\gamma^2 4L}\mathbf{P}+\mathbf{P}Diag(\mathbf{P}\mathbf{R}^2\mathbf{P})\mathbf{P}\right)\\ &+\frac{\sigma_\eta^2}{L\gamma^2}\left(\sigma_\eta^2 Tr(\mathbf{R}^2)+\frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}))^2\right)\mathbf{P}\\ &+2\sigma_\eta^2\left(Tr(\mathbf{R})\left(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}+\frac{W^H\mathbf{R}W}{2L\gamma^2}\mathbf{P}\right)\right)\end{aligned} \qquad \text{(e.123)}$$

When the non projected sequence is used the gradient covariance is given by

$$\begin{aligned}\mathbf{V}_{G3}(W_q(n)) &= \mathbf{V}_{G3_{orig}}(W(n))\\ &+4\sigma_\eta^2\left(\frac{W^H\mathbf{R}^2W}{\gamma^2 4L}\mathbf{I}_{LL}+Diag(\mathbf{R}^2)\right)\\ &+\frac{\sigma_\eta^2}{L\gamma^2}\left(\sigma_\eta^2 Tr(\mathbf{R}^2)+\frac{\sigma_\eta^2}{2}(Tr(\mathbf{R}))^2\right)\mathbf{I}_{LL}\\ &+2\sigma_\eta^2\left(Tr(\mathbf{R})\left(Diag(\mathbf{R})+\frac{W^H\mathbf{R}W}{2L\gamma^2}\mathbf{I}_{LL}\right)\right)\end{aligned} \qquad \text{(e.124)}$$

where $\mathbf{V}_{G3_{orig}}(W(n))$ is given by (f.9).

Gradient Covariance With Respect to the Weight Error Vector

To determine the misadjustment we require the gradient covariance to be expressed in terms of the weight error vector. Here we evaluate these expressions.

Dual Receiver Dual Perturbation System- Gradient Covariance

Substituting the expression for the weight error vector, (3.54), into (4.8), one obtains the following

$$\mathbf{V}_{G1}(W_q(n)) = 2(V(n)+\hat{W})^H\mathbf{R}(V(n)+\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$
$$+ \frac{\sigma_\eta^2}{\gamma^2 2L}((V(n)+\hat{W})^H\mathbf{R}^2(V(n)+\hat{W}) + Tr(\mathbf{R})(V(n)+\hat{W})^H\mathbf{R}(V(n)+\hat{W})^H)\mathbf{P}$$
$$+ \frac{\sigma_\eta^4}{\gamma^2 4L}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\mathbf{P} \quad \text{(e.125)}$$

Using (c.3) and (c.7) in (e.125) gives

$$\mathbf{V}_{G1}(W_q(n)) = 2(V^H(n)\mathbf{R}V(n) + \hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$
$$+ \frac{\sigma_\eta^2}{\gamma^2 2L}(V^H(n)\mathbf{R}^2V(n) + \hat{W}^H\mathbf{R}^2\hat{W} + V^H(n)\mathbf{R}^2\hat{W} + \hat{W}^H\mathbf{R}^2V(n))\mathbf{P}$$
$$+ \frac{\sigma_\eta^2}{\gamma^2 2L}(Tr(\mathbf{R})(V^H(n)\mathbf{R}V(n) + \hat{W}^H\mathbf{R}\hat{W}))\mathbf{P}$$
$$+ \frac{\sigma_\eta^4}{\gamma^2 4L}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\mathbf{P} \quad \text{(e.126)}$$

taking expectation with respect to $V(n)$ and using the following

$$B(n) = E[V(n)V^H(n)] \quad \text{(e.127)}$$

$$E[V^H(n)\mathbf{R}^2\hat{W}] = E[\hat{W}^H\mathbf{R}^2V(n)] = 0 \quad \text{(e.128)}$$

gives

$$\mathbf{V}_{G1}(W_q(n)) = 2(Tr(\mathbf{R}B(n)) + \hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$
$$+ \frac{\sigma_\eta^2}{\gamma^2 2L}(Tr(\mathbf{R}^2B(n)) + \hat{W}^H\mathbf{R}^2\hat{W})\mathbf{P}$$
$$+ \frac{\sigma_\eta^2}{\gamma^2 2L}(Tr(\mathbf{R})(Tr(\mathbf{R}B(n)) + \hat{W}^H\mathbf{R}\hat{W}))\mathbf{P}$$
$$+ \frac{\sigma_\eta^4}{\gamma^2 4L}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\mathbf{P} \quad \text{(e.129)}$$

and re-arranging the terms gives

$$\mathbf{V}_{G1}(W_q(n)) = Tr(\mathbf{R}B(n))\left(2(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \frac{\sigma_\eta^2}{\gamma^2 2L}Tr(\mathbf{R})\mathbf{P}\right)$$

$$+ \frac{\sigma_\eta^2}{\gamma^2 2L}Tr(\mathbf{R}^2 B(n))\mathbf{P} + 2\hat{W}^H\mathbf{R}\hat{W}\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}$$

$$+ \sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$

$$+ \frac{\sigma_\eta^2}{\gamma^2 2L}\left(Tr(\mathbf{R})\hat{W}^H\mathbf{R}\hat{W} + \hat{W}^H\mathbf{R}^2\hat{W} + \frac{\sigma_\eta^2}{\gamma^2 2}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\right)\mathbf{P} \quad \text{(e.130)}$$

Dual Receiver Reference Receiver System - Gradient Covariance

Substituting the expression for the weight error vector, (3.54), into (4.10), one obtains the following

$$\mathbf{V}_{G2}(W_q(n)) = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P}$$

$$+ 2(V(n) + \hat{W})^H\mathbf{R}(V(n) + \hat{W})\mathbf{P}(Diag(\mathbf{PRP}))\mathbf{P}$$

$$+ 2\sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$

$$+ \frac{2\sigma_\eta^2}{\gamma^2 L}(V^H(n)\mathbf{R}^2V(n) + \hat{W}^H\mathbf{R}^2\hat{W} + V^H(n)\mathbf{R}^2\hat{W} + \hat{W}^H\mathbf{R}^2V(n))\mathbf{P}$$

$$+ \frac{2\sigma_\eta^2}{\gamma^2 L}(Tr(\mathbf{R})(V^H(n)\mathbf{R}V(n) + \hat{W}^H\mathbf{R}\hat{W}))\mathbf{P}$$

$$+ \frac{\sigma_\eta^4}{\gamma^2 L}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\mathbf{P} \quad \text{(e.131)}$$

Using (c.3), (c.7), (e.127) in (e.131) and taking expectation with respect to $V(n)$ gives

$$\mathbf{V}_{G2}(W_q(n)) = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P}$$

$$+ 2(Tr(\mathbf{R}B(n)) + \hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}$$

$$+ 2\sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$

$$+ \frac{2\sigma_\eta^2}{\gamma^2 L}(Tr(\mathbf{R}^2B(n)) + \hat{W}^H\mathbf{R}^2\hat{W})\mathbf{P}$$

$$+ \frac{2\sigma_\eta^2}{\gamma^2 L}(Tr(\mathbf{R})(Tr(\mathbf{R}(B(n))) + \hat{W}^H\mathbf{R}\hat{W}))\mathbf{P}$$

$$+\frac{\sigma_\eta^4}{\gamma^2 L}(Tr(\mathbf{R}^2)+(Tr(\mathbf{R}))^2)\mathbf{P} \tag{e.132}$$

rearranging the terms gives

$$\begin{aligned}\mathbf{V}_{G2}(W_q(n)) &= Tr(\mathbf{R}B(n))\left(2\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}+\frac{2\sigma_\eta^2}{\gamma^2 L}Tr(\mathbf{R})\mathbf{P}\right)\\ &+2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P}+2(\hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}\\ &+2\sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}\\ &+\frac{2\sigma_\eta^2}{\gamma^2 L}Tr(\mathbf{R}^2B(n))\mathbf{P}\\ &+\frac{2\sigma_\eta^2}{\gamma^2 L}\left(\hat{W}^H\mathbf{R}^2\hat{W}+Tr(\mathbf{R})\hat{W}^H\mathbf{R}\hat{W}+\frac{\sigma_\eta^2}{\gamma^2 2L}(Tr(\mathbf{R}^2)+(Tr(\mathbf{R}))^2)\right)\mathbf{P}\end{aligned} \tag{e.133}$$

Single Receiver System - Gradient Covariance

Substituting the expression for the weight error vector, (3.54), into (4.12) one obtains the following

$$\begin{aligned}\hat{\mathbf{V}}_{G3}(W_q(n)) &= 2\mathbf{P}[Diag(\mathbf{PR}(V(n)+\hat{W})(V(n)+\hat{W})^H\mathbf{RP})]\mathbf{P}\\ &+\mathbf{P}[a(V(n)+\hat{W})^H\mathbf{R}(V(n)+\hat{W})Diag(\mathbf{PRP})]\mathbf{P}\\ &+2\sigma_\eta^2\mathbf{P}(2Diag(\mathbf{PR}^2\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}\\ &+\frac{\sigma_\eta^2}{\gamma^2 L}((V(n)+\hat{W})^H\mathbf{R}^2(V(n)+\hat{W})+Tr(\mathbf{R})(V(n)+\hat{W})^H\mathbf{R}(V(n)+\hat{W})^H)\mathbf{P}\\ &+\frac{\sigma_\eta^4}{\gamma^2 L}\left(Tr(\mathbf{R}^2)+\frac{1}{2}(Tr(\mathbf{R}))^2\right)\mathbf{P}\end{aligned} \tag{e.134}$$

Using (c.3), (c.7) in (e.134) gives

$$\begin{aligned}\hat{\mathbf{V}}_{G3}(W_q(n)) &= a(V^H(n)\mathbf{R}V(n)+\hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}\\ &+2\mathbf{P}Diag(\mathbf{PR}V(n)V^H(n)\mathbf{RP})\mathbf{P}\\ &+2\sigma_\eta^2\mathbf{P}(2Diag(\mathbf{PR}^2\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}\\ &+\frac{\sigma_\eta^2}{\gamma^2 L}(V^H(n)\mathbf{R}^2V(n)+\hat{W}^H\mathbf{R}^2\hat{W}+V^H(n)\mathbf{R}^2\hat{W}+\hat{W}^H\mathbf{R}^2V(n))\mathbf{P}\end{aligned}$$

$$+\frac{\sigma_\eta^2}{\gamma^2 L}(Tr(\mathbf{R})(V^H(n)\mathbf{R}V(n)+\hat{W}^H\mathbf{R}\hat{W}))\mathbf{P}$$

$$+\frac{\sigma_\eta^4}{\gamma^2 L}\left(Tr(\mathbf{R}^2)+\frac{1}{2}(Tr(\mathbf{R}))^2\right)\mathbf{P} \qquad \text{(e.135)}$$

and taking expectation with respect to $V(n)$ and using (e.127) in (e.135)gives

$$\hat{\mathbf{V}}_{G3}(W_q(n)) = a(Tr(\mathbf{R}B(n))+\hat{W}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}$$

$$+2\mathbf{P}Diag(\mathbf{PR}B(n)\mathbf{RP})\mathbf{P}$$

$$+2\sigma_\eta^2\mathbf{P}(2Diag(\mathbf{PR}^2\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$

$$+\frac{\sigma_\eta^2}{\gamma^2 L}(Tr(\mathbf{R}^2B(n))+\hat{W}^H\mathbf{R}^2\hat{W})\mathbf{P}$$

$$+\frac{\sigma_\eta^2}{\gamma^2 L}(Tr(\mathbf{R})(Tr(\mathbf{R}(B(n)))+\hat{W}^H\mathbf{R}\hat{W}))\mathbf{P}$$

$$+\frac{\sigma_\eta^4}{\gamma^2 L}\left(Tr(\mathbf{R}^2)+\frac{1}{2}(Tr(\mathbf{R}))^2\right)\mathbf{P} \qquad \text{(e.136)}$$

re-arranging the terms

$$\hat{\mathbf{V}}_{G3}(W_q(n)) = Tr(\mathbf{R}B(n))\left(a\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}+\frac{\sigma_\eta^2}{\gamma^2 L}Tr(\mathbf{R})\mathbf{P}\right)$$

$$+2\mathbf{P}Diag(\mathbf{PR}B(n)\mathbf{RP})\mathbf{P}+\frac{\sigma_\eta^2}{\gamma^2 L}Tr(\mathbf{R}^2B(n))\mathbf{P}$$

$$+a\hat{W}^H\mathbf{R}\hat{W}\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}+2\sigma_\eta^2\mathbf{P}(2Diag(\mathbf{PR}^2\mathbf{P})+Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P}$$

$$+\frac{\sigma_\eta^2}{\gamma^2 L}\left(\hat{W}^H\mathbf{R}^2\hat{W}+Tr(\mathbf{R})\hat{W}^H\mathbf{R}\hat{W}+\sigma_\eta^2\left(Tr(\mathbf{R}^2)+\frac{1}{2}(Tr(\mathbf{R}))^2\right)\right)\mathbf{P} \qquad \text{(e.137)}$$

Gradient Covariance - Generic Expression

Examining equations (e.130), (e.133) and (e.137) it can be observed that the covariance expressions are similar and can be expressed in a generic form given by the following

$$\mathbf{V}_G(W_q(n)) = Tr(\mathbf{R}B(n))A_1 + Tr(\mathbf{R}^2B(n))A_2 + A_3 + A_4Tr(\mathbf{P}Diag(\mathbf{PR}B(n)\mathbf{RP})) \qquad (e.138)$$

The variables A_1, A_2, A_3 and A_4 are defined below.

Dual Receiver Dual Perturbation System

$$A_1 = \left(2(\mathbf{P}Diag(\mathbf{PRP})\mathbf{P}) + \frac{\sigma_\eta^2}{\gamma^2 2L}Tr(\mathbf{R})\mathbf{P}\right)$$

$$A_2 = \frac{\sigma_\eta^2}{\gamma^2 2L}\mathbf{P}$$

$$A_3 = 2\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P} + \frac{\sigma_\eta^2}{\gamma^2 2L}\left(Tr(\mathbf{R})\hat{W}^H\mathbf{R}\hat{W} + \hat{W}^H\mathbf{R}^2\hat{W} + \frac{\sigma_\eta^2}{\gamma^2 2}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\right)\mathbf{P} \qquad (e.139)$$

Dual Receiver Reference Receiver System

$$A_1 = \left(2\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \frac{2\sigma_\eta^2}{\gamma^2 L}Tr(\mathbf{R})\mathbf{P}\right)$$

$$A_2 = \frac{2\sigma_\eta^2}{\gamma^2 L}\mathbf{P}$$

$$A_3 = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{PRP}))^2\mathbf{P} + 2(\hat{\mathbf{W}}^H\mathbf{R}\hat{W})\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + 2\sigma_\eta^2\mathbf{P}(Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P} + \frac{2\sigma_\eta^2}{\gamma^2 L}\left(\hat{W}^H\mathbf{R}^2\hat{W} + Tr(\mathbf{R})\hat{W}^H\mathbf{R}\hat{W} + \frac{\sigma_\eta^2}{\gamma^2 2L}(Tr(\mathbf{R}^2) + (Tr(\mathbf{R}))^2)\right)\mathbf{P} \qquad (e.140)$$

Single Receiver System

$$A_1 = a\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + \frac{\sigma_\eta^2}{\gamma^2 L}Tr(\mathbf{R})\mathbf{P}$$

$$A_2 = \frac{\sigma_\eta^2}{\gamma^2 L}\mathbf{P}$$

$$A_3 = a\hat{\mathbf{W}}^H\mathbf{R}\hat{\mathbf{W}}\mathbf{P}Diag(\mathbf{PRP})\mathbf{P} + 2\sigma_\eta^2\mathbf{P}(2Diag(\mathbf{PR}^2\mathbf{P}) + Tr(\mathbf{R})Diag(\mathbf{PRP}))\mathbf{P} + \frac{\sigma_\eta^2}{\gamma^2 L}\left(\hat{W}^H\mathbf{R}^2\hat{W} + Tr(\mathbf{R})\hat{W}^H\mathbf{R}\hat{W} + \sigma_\eta^2\left(Tr(\mathbf{R}^2) + \frac{1}{2}(Tr(\mathbf{R}))^2\right)\right)\mathbf{P}$$

$$A_4 = 2 \qquad (e.141)$$

Misadjustment - Bounds Analysis

In this abbreviated analysis we assume that the conditions for convergence of the norm of the weight error vector as established in Section 3.5 and Appendix B are satisfied, the trace of the norm of the weight error vector in the limit as $n \to \infty$ can then be taken. By substituting (e.138) into (3.58), and by using Lemma A.2 an upper and lower bound for $\lim_{n \to \infty} Tr(\mathbf{R}\boldsymbol{B}(n))$ can be obtained. They are given by:

$$\lim_{n \to \infty} Tr(\mathbf{R}\boldsymbol{B}(n)) \le \frac{\mu A_3}{4 - \mu(A_1 + \lambda_{max} A_2 + \lambda_{max} A_4 + 4\lambda_{max})} \tag{e.142}$$

$$\lim_{n \to \infty} Tr(\mathbf{R}\boldsymbol{B}(n)) \ge \frac{\mu A_3}{4 - \mu A_1} \tag{e.143}$$

The equations (e.142) and (e.143) correspond to the upper and lower bounds for the ssemsp. The bounds on the misadjustment can be obtained from these.

The asymptotic misadjustment can be obtained by considering μ to be small in (e.142) and (e.143). The asymptotic misadjustment is then given by the following

$$M = \frac{\mu A_3}{4 \hat{W}^H \mathbf{R} \hat{W}} \tag{e.144}$$

The misadjustment expression (4.15), (4.16) and (4.17) are obtained by making the appropriate substitutions from (e.139), (e.140) and (e.141) into (e.144).

Appendix F

For reference purposes this appendix contains results derived in [2], [10], [39], [40]. For all the results shown here, the non projected Time Multiplex sequence is used throughout the receiver structures.

Gradient Covariance of Dual Receiver Dual Perturbation System

$$\mathbf{V}_{G1}(W(n)) = 2W^H(n)\mathbf{R}W(n)Diag(\mathbf{R}) \text{ for a } 4L \text{ length sequence.} \quad \text{(f.1)}$$

$$\mathbf{V}_{G1}(W(n)) = 4W^H(n)\mathbf{R}W(n)Diag(\mathbf{R}) \text{ for a } 2L \text{ length sequence.} \quad \text{(f.2)}$$

Misadjustment of the Dual Receiver Dual Perturbation System - Direct Approach

For a given *W(n)* if the covariance of the gradient used in (2.27) is given by (f.1) and

$$0 < \mu < \frac{1}{\lambda_{max} + \frac{Tr(\mathbf{R})}{2L}} \text{ and } \frac{\mu Tr(\mathbf{R})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i} < 1$$

$$\text{then } M = \frac{\frac{\mu Tr(\mathbf{R})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}}{1-\frac{\mu Tr(\mathbf{R})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}} \quad \text{(f.3)}$$

where λ represents the eigenvalues of **PRP**.

Misadjustment of the Dual Receiver Dual Perturbation System - Bounds Approach

For a given *W*(*n*) if the covariance of the gradient used in (2.27) is given by (f.1) and

$$\text{if } E[\|V(0)\|^2] < \infty \text{ and } 0 < \mu < \frac{1}{0.5\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + 2\lambda_{max}}$$

then the steady state excess mean square power is bounded and the corresponding misadjustment is bounded by $b_L \le M \le b_h$

where

$$b_L = \frac{\frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R})\right]}{1-\frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R})\right]} \text{ and } b_h = \frac{\frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R})\right]}{1-\frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + 2\lambda_{max}\right]} \quad \text{(f.4)}$$

where λ represents the eigenvalues of $\mathbf{R}$

Gradient Covariance of Dual Receiver Reference Receiver System

$$\mathbf{V}_{G2}(\boldsymbol{W}(n)) = 2\gamma^2 L\mathbf{P}(Diag(\mathbf{R}))^2\mathbf{P} + 2W^H(n)\mathbf{R}W(n)Diag(\mathbf{R}) \quad \text{(f.5)}$$

for a $4L$ length sequence

Misadjustment of the Dual Receiver Reference Receiver System - Direct Approach

For a given $\boldsymbol{W}(n)$ if the covariance of the gradient used in (2.27) is given by (f.5) and

$$0 < \mu < \frac{1}{\lambda_{max} + \frac{Tr(\mathbf{R})}{2L}} \text{ and } \frac{\mu Tr(\mathbf{R})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i} < 1$$

$$\text{then } M = \frac{\left[1 + \frac{\gamma^2 Tr(\mathbf{R})}{\hat{W}^H\mathbf{R}\hat{W}}\right]\frac{\mu Tr(\mathbf{R})}{L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}}{1 - \frac{\mu Tr(\mathbf{R})}{2L}\sum_{i=1}^{L-1}\frac{1}{1-\mu\lambda_i}} \quad \text{(f.6)}$$

where λ represent the eigenvalues of $\mathbf{PRP}$.

Misadjustment of the Dual Receiver Reference Receiver System - Bounds Approach

For a given $\boldsymbol{W}(n)$ if the covariance of the gradient used in (2.27) is given by (f.5) and

$$\text{if } E[\|V(0)\|^2] < \infty \text{ and } 0 < \mu < \frac{1}{0.5\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + \lambda_{max}}$$

then the steady state excess mean square power is bounded and the corresponding misadjustment is bounded by $b_L \le M \le b_h$

$$\text{where } b_L = \frac{\frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R})\left(1 + \gamma^2\frac{Tr(\mathbf{R})}{\hat{W}^H\mathbf{R}\hat{W}}\right)\right]}{1 - \frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R})\right]} \quad \text{(f.7)}$$

$$\text{and } b_h = \frac{\frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R})\left(1 + \gamma^2\frac{Tr(\mathbf{R})}{\hat{W}^H\mathbf{R}\hat{W}}\right)\right]}{1 - \frac{\mu}{2}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + 2\lambda_{max}\right]} \quad \text{(f.8)}$$

where λ represents the eigenvalues of $\mathbf{R}$.

Gradient Covariance of Single Receiver System

$$\mathbf{V}_{G3}(W(n)) = \mathbf{P}\Big(\gamma^2 2L(Diag(\mathbf{R}))^2 + \frac{(W^H(n)\mathbf{R}W(n))^2}{\gamma^2 2L} + 2Diag(\mathbf{R}W(n)W^H(n)\mathbf{R}) + 2W^H(n)\mathbf{R}W(n)Diag(\mathbf{R})\Big)\mathbf{P} \quad \text{(f.9)}$$

Misadjustment of the Single Receiver System - Bounds Approach

For a given $W(n)$ if the covariance of the gradient used in (2.27) is given by (f.9) with

$$\hat{\gamma}(W(n)) = c\left[\frac{W^H(n)\mathbf{R}W(n)}{2Tr(\mathbf{R})}\right]^{\frac{1}{2}}$$

If $E[\|V(0)\|^2] < \infty$ and $0 < \mu < \dfrac{1}{\frac{a}{4}\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + 1.5\lambda_{max}}$

then the steady state excess mean square power is bounded and the corresponding misadjustment is bounded by $b_L \le M \le b_h$

$$\text{where } b_L = \frac{\frac{\mu}{2}\left[\frac{a}{4}\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + 0.5(L-1)\hat{W}^H\mathbf{R}\hat{W}\right]}{1 - \frac{\mu a}{4}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R})\right]} \quad \text{(f.10)}$$

$$\text{and } b_h = \frac{\frac{\mu}{2}\left[\frac{a}{4}\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + 0.5(L-1)\hat{W}^H\mathbf{R}\hat{W}\right]}{1 - \frac{\mu a}{4}\left[\left(\frac{L-1}{L}\right)Tr(\mathbf{R}) + 1.5\lambda_{max}\right]} \quad \text{(f.11)}$$

where λ represents the eigenvalues of $\mathbf{R}$ and $a = \left(c + \frac{1}{c}\right)^2$