Real Time Characterisation of the Mobile Multipath Channel

Paul Teal

B.E. (Honours), University of Sydney

November 2001

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University

Department of Telecommunications Engineering
Research School of Information Sciences and Engineering
The Australian National University

This Ph.D. thesis was examined by Professor Jørgen Bach Andersen, Center for PersonCommunikation, Aalborg University, Denmark and Jack Winters, AT&T Labs Research, Redbank, New Jersey, in December 2001 and March 2002.

The subfigures in the pdf version of this document sometimes do not quite match in size. But the file is a *lot* smaller than the postscript file. ☺

Declaration

The contents of this thesis are the results of original research and have not been submitted for a higher degree to any other university or institution.

Much of the work in this thesis has been published or has been submitted for publication as journal papers or conference proceedings. These papers are:

1. P. D. Teal, R. Raich, and R. G. Vaughan. Prediction of fading in the mobile multipath environment. *IEE Proceedings — Communications*, 2000. Submitted.

2. P. D. Teal, R. C. Williamson, and R. A. Kennedy. Error performance of a channel of known impulse response. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 2733-2736, Istanbul, June 2000.

3. R. Vaughan, P. Teal, and R. Raich. Prediction of fading signals in a multipath environment. In *Proc. IEEE Vehicular Technology Conference*, volume 1, pages 751–758, Fall, Boston, Sep 2000.

4. O. Nørklit, P. D. Teal, and R. G. Vaughan. Measurement and evaluation of multi-antenna handsets in indoor mobile communication. *IEEE Trans. Antennas & Propagation*, 49(3):429–437, Mar 2001.

5. P. D. Teal and R. G. Vaughan. Simulation and performance bounds for real-time prediction of the mobile multipath channel. In *Proc. IEEE Workshop on Statistical Signal Processing*, pages 548–551, Singapore, Aug 2001.

6. P. D. Teal, T. A. Abhayapala, and R. A. Kennedy. Spatial correlation for general distributions of scatterers. *IEEE Signal Processing Letters*, 2001. (to appear).

7. P. D. Teal, T. A. Abhayapala, and R. A. Kennedy. Spatial correlation in non-isotropic scattering scenarios. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, Orlando, Florida, 2002. (to appear).

8. P. D. Teal. Rough surface scattering and prediction of the mobile channel. *IEEE Communications Letters*, 2001. To be submitted.

The following informal presentations were also published.

1. P. Teal. Error performance of a channel of known impulse response. In *2nd Australian Communications Theory Workshop*, Adelaide, Feb 2001. http://www.itr.unisa.edu.au/~alex/AusCTW2001/talks/teal1.html.

2. P. Teal. Performance bounds for real time prediction of the mobile multipath channel. In *2nd Australian Communications Theory Workshop*, Adelaide, Feb 2001. http://www.itr.unisa.edu.au/~alex/AusCTW2001/talks/teal2.html.

The research represented in this thesis has been performed jointly with Professor Robert C. Williamson, Professor Rodney A. Kennedy, and Dr Rodney G. Vaughan. The substantial majority of this work is my own.

Paul Teal
The Australian National University
November 2001

Acknowledgements

Several people have contributed to this thesis, and I would like to take this opportunity to thank them.

I am grateful to Rodney Vaughan whose research programme formed the basis for this project, and who has supported it all the way through.

Thank you also to my supervisors at the Australian National University, Bob Williamson and Rodney Kennedy who contributed quite a lot of time to proposing various ideas, encouraging me in some of them, and saving me from unnecessary cul-de-sacs on others.

I am grateful to the management of Industrial Research Limited (Lower Hutt, New Zealand) who have supported my study in Australia, and in particular to Sunil Vather who encouraged me to embark on this project, and who smoothed the way for me on more than one occasion.

My colleagues at Industrial Research Limited, particularly Raviv Raich have been involved in the research of which this thesis is only a part, and have consequently contributed to discussions relating to it. In the early stages I'm not sure how I would have got started without Raviv. Thank you also to many who have answered questions via email when I've been in Australia.

Thank you Jan for being willing to take our family to Canberra, three times, and especially for believing in me enough to think that it was all worth while. I know that not every man is blessed with so supportive a wife. And thank you too Marta and Jemima, for your patience with all the upheaval.

Without any pretensions with regard to the relative importance of my work and Newton's I am grateful to God for the opportunity to explore his world, and to submit this contribution that "we may now more nearly behold the beauties of Nature, and entertain ourselves with the delightful contemplation; and which is the best and most valuable fruit of philosophy, be thence incited the more profoundly to reverence and adore the great Maker and Lord of all." [101]

Abstract

In this thesis a new approach for characterisation of digital mobile radio channels is investigated. The new approach is based on recognition of the fact that while the fading which is characteristic of the mobile radio channel is very rapid, the processes underlying this fading may vary much more slowly. The comparative stability of these underlying processes has not been exploited in system designs to date.

Channel models are proposed which take account of the stability of the channel. Estimators for the parameters of the models are proposed, and their performance is analysed theoretically and by simulation and measurement.

Bounds are derived for the extent to which the mobile channel can be predicted, and the critical factors which define these bounds are identified.

Two main applications arise for these channel models. The first is the possibility of prediction of the overall system performance. This may be used to avoid channel fading (for instance by change of frequency), or compensate for it (by change of the signal rate or by power control). The second application is in channel equalisation. An equaliser based on a model which has parameters varying only very slowly can offer improved performance especially in the case of channels which appear to be varying so rapidly that the convergence rate of an equaliser based on the conventional model is not adequate.

The first of these applications is explored, and a relationship is derived between the channel impulse response and the performance of a broadband system.

Notation and Symbols

$.^*$	Complex conjugate
$\|\| \cdot \|\|$	Frobenius norm of a matrix or 2-norm of a vector
$.^+$	Moore-Penrose generalised inverse
$\odot$	Element by element (Hadamard) product
$\circledast$	Convolution
$\equiv$	Equivalence
$\triangleq$	Definition
$\boldsymbol{x} \cdot \boldsymbol{y}$	Dot product between two spatial vectors $\boldsymbol{x}$ and $\boldsymbol{y}$
$\mathbf{A}$	Array steering matrix
$\mathbf{a}$	Array steering vector
$b(\cdot)$	Scattering function
c	Speed of light
$\mathcal{CN}(\cdot, \cdot)$	Complex normally distributed with given mean and covariance
D, d	Distances between source points and measurement points
$\mathbf{D}$	Diagonal matrix of eigenvalues
$E\{\cdot\}$	Expectation operator
$e^{\cdot}$	Exponential operator
$\mathbf{e}_i$	A vector of zeros with 1 in the i-th position
f	Discrete channel impulse response co-efficients
$g(\cdot)$	Antenna gain pattern
$.^H$	Complex Hermitian conjugate
$H_n^{(1)}(\cdot)$	Order n Hankel function of the first kind
h	System or channel impulse response
$I_n(\cdot)$	Order n Modified Bessel function of the first kind
$\mathbf{I}$	Identity matrix
$J_n(\cdot)$	Order n Bessel function of the first kind
j	$\sqrt{-1}$
$j_n(\cdot)$	Order n Spherical Bessel function
$\mathbf{J}$	Fisher Information matrix
$\mathbf{K}$	Row/Column reversing matrix
k	Wave number of signal carrier $= 2\pi/\lambda$

l	Snapshot index
L	The number of terms of a finite impulse response channel model or the number of "snapshots" taken by an array
M	The number of sensors in an array
m	Array sensor index
N	The number of sources or scatterers in the environment of an array
n	Source or scatterer index
$\mathcal{N}(\cdot,\cdot)$	Normally distributed with given mean and covariance
P	dimension of a covariance or correlation matrix
$P_n^m(\cdot)$	Associated Legendre functions
$p(t)$	Transmit pulse shaping function
$s(t)$ $s[m]$	Transmitted signal or transmitted symbol
T	System symbol period
$.^T$	Matrix or vector transpose
t	Time
$U(\cdot)$	Unit step function (Heaviside's unit function)
$u(\cdot)$	Baseband transmitted signal
$\mathbf{v}$	an eigenvector
$\mathbf{V}$	Matrix of eigenvectors
W	Bandwidth of a baseband transmitted signal
x, y, z	3 dimensional spatial co-ordinates
$\delta(t)$	Dirac delta
$\delta_i, \delta_{i,j}, \delta[i]$	Kronecker delta
ϵ	Error in decoded symbol
ζ	Gain (attenuation) of a path between transmitter and receiver
η	Additive Noise process
θ	Arrival angle of a signal (measured from broadside)
λ	Wavelength of signal carrier
$\sum$	Summation operator
τ	Path delay
ω_c	Carrier frequency
ϖ_n	Spatial frequency corresponding to path n
ω	Baseband channel frequency
cdf	cumulative distribution function
pdf	probability density function
SNR	Signal to Noise Ratio

Contents

Declaration iii

Acknowledgements v

Abstract vii

Notation and Symbols ix

1 Introduction 1

1.1 The Mobile Radio Propagation Channel 1
1.2 Channel Prediction 2
1.2.1 Prediction Concept 2
1.2.2 Approach to Predicting Channel Behaviour 3
1.3 Applications of the Prediction Information 3
1.3.1 Avoidance of "bad" channels 3
1.3.2 Adaptive Modulation 4
1.3.3 Parametric Equalisation 5
1.3.4 Improved Combining 6
1.4 Required Prediction Lengths 6
1.5 Structure of this Thesis 7

2 Multipath Propagation Models 9

2.1 General Channel Models 9
2.1.1 Single Path 9
2.1.2 Moving Receiver 10
2.1.3 Time Varying Impulse Response Notation 11
2.1.4 Multiple Paths 12
2.2 Stochastic Channel Characterisations 14
2.3 Spatial Correlation 16
2.3.1 Spatial Correlation for General Distributions of Scatterers 18
2.3.2 Three Dimensional Scattering Environment 19
2.3.3 Two Dimensional Scattering Environment 21

2.3.4 Mutual Coupling 26
2.3.5 Correlation for a Small Number of Scatterers 27
2.4 Diversity 29
2.5 Real Time Channel Characterisation 33
2.6 Deterministic Channel Models 34
2.6.1 Narrowband Far Field Sources 34
2.6.2 Array Processing 35
2.6.3 Linear Prediction 37
2.6.4 Narrowband Near Field Sources 38
2.6.5 Wide-band Sources 38
2.7 Model Limitations 39
2.7.1 Source Extent 39
2.7.2 Number of Paths 40
2.7.3 Near field sources 40
2.7.4 Changes in the Parameters 41
2.8 Summary and Contributions 41

3 Prediction Algorithms 43
3.1 Deterministic Point Sources Channel Model 43
3.1.1 Likelihood Function 43
3.1.2 First Derivative 46
3.1.3 Far field Sources 47
3.1.4 Near Field Sources 48
3.1.5 Parameter Estimation 49
3.2 Subspace Spectral Estimation 49
3.2.1 MUSIC 52
3.2.2 Near Field MUSIC 53
3.2.3 Aliasing 55
3.2.4 Minimum Norm Spectral Estimation 56
3.2.5 Principal Components Linear Prediction (PCLP) 57
3.2.6 ESPRIT Algorithm 58
3.3 Gaussian Model 59
3.4 Polynomial Prediction 61
3.5 Fixed Sector Prediction 61
3.6 Determining the Model Order $\hat{N}$ 64
3.7 Support Vector Machines 65
3.8 Summary and Contributions 66

4 Simulations, Measurements and Bounds **67**
4.1 Performance Measures . . . 67
4.2 Simulations . . . 69
4.2.1 Simulation Setup . . . 69
4.2.2 Simulation Results . . . 69
4.2.3 Prediction Length versus Measurement Segment Length . . . 70
4.2.4 Distribution of Prediction Lengths . . . 72
4.2.5 Prediction Length versus SNR . . . 72
4.2.6 Prediction Length versus Number of Samples . . . 73
4.2.7 Prediction Length versus Model Order . . . 73
4.2.8 Prediction Length versus Performance Threshold . . . 74
4.3 Measurements . . . 75
4.4 Cramer Rao Bounds . . . 79
4.4.1 Invertibility of the Fisher Information Matrix . . . 82
4.5 Rough Surface Scattering . . . 87
4.5.1 Scattering formulation . . . 88
4.5.2 Moving Receiver . . . 90
4.5.3 Numerical Study . . . 91
4.5.4 Conclusion . . . 92
4.6 Mutual Information Considerations . . . 93
4.7 Multipath Dimensionality . . . 94
4.7.1 General Two Dimensional Multipath Field . . . 94
4.7.2 Dimensionality of Multipath . . . 96
4.7.3 Bounding the Relative Error . . . 97
4.7.4 Dimensionality of 3 Dimensional Field . . . 99
4.7.5 Plane Wave Synthesis . . . 101
4.7.6 Extrapolation Bounds . . . 102
4.7.7 Conclusions . . . 104
4.8 Summary and Contributions . . . 105

5 Channel Performance **107**
5.1 MLSE Error Probability . . . 108
5.2 Reduced Complexity Error Rate Calculation . . . 112
5.2.1 Minimum Distance Error Events . . . 112
5.2.2 Contributing Error Events . . . 113
5.3 Algorithm for finding the set $\mathcal{U}$ . . . 114
5.3.1 Stage 1 . . . 114
5.3.2 Stage 2 . . . 115
5.3.3 Stage 3 . . . 116
5.4 Symmetry of Contributing Error Events . . . 119

5.5 Visualisation . . . 122
5.5.1 Three Dimensional Slices With One Co-ordinate Fixed . . . 124
5.5.2 Rotating Hyper-plane Three Dimensional Slice . . . 125
5.6 Effect of Imperfect Channel Knowledge . . . 126
5.6.1 Stochastic Channel Estimate and Noise . . . 126
5.6.2 Stochastic transmitted symbols $\mathbf{s}$ . . . 129
5.7 Summary and Contributions . . . 132
5A Proof of Theorem 1 . . . 133
5A.1 The term $t_1 = \mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{f}$ . . . 133
5A.2 The term $t_2 = {\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f}$ . . . 133
5A.3 The term $t_3 = 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f})$ . . . 134
5A.4 The term $t_4 = 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})$ . . . 135
5A.5 The term $t_5 = 2\,\mathrm{Re}({\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})$ . . . 135
5A.6 The term $t_6 = -2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\boldsymbol{\eta})$ . . . 135
5A.7 The term $t_7 = -2\,\mathrm{Re}({\Delta_\mathbf{f}}^H\mathbf{E}^H\boldsymbol{\eta})$ . . . 136
5A.8 The cross-correlation of t_3 and t_4 . . . 137
5A.9 The cross correlation of t_2 and t_5 . . . 137
5A.10 The covariance of t_2 and t_3 . . . 138
5A.11 The Combined Expression . . . 138
5B Proof of Theorem 2 . . . 139
5B.1 The term $t_4 = 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})$ . . . 139
5B.2 The term $t_5 = 2\,\mathrm{Re}({\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})$ . . . 140
5B.3 The combined expression for random $\mathbf{s}$ . . . 140
5C Note on 'A New Method for Determining "Unknown" Worst-Case Channels for Maximum Likelihood Sequence Estimation' . . . 142

6 Conclusions and Further Research 145
6.1 Conclusions . . . 145
6.2 Future Research . . . 146

A Contributing Error Events 149
A.1 BPSK . . . 150
A.2 4PAM . . . 151
A.3 8PAM . . . 153
A.4 QPSK . . . 155
A.5 8-PSK . . . 156
A.6 16-QAM . . . 158
A.7 16-PSK . . . 160
A.8 Asymmetric Sequences . . . 164
A.8.1 8-PSK . . . 164

A.8.2 16-PSK 165

References **167**

Index **181**

Chapter 1

Introduction

1.1 The Mobile Radio Propagation Channel

BY an interesting co-incidence the wavelength of the electromagnetic wave used by a mobile telephone is approximately the same as that of the sound wave which it is used to transmit. Both electromagnetic waves and pressure waves can reach the intended receiver by many paths; by line-of-sight propagation and by reflections and scattering from objects such as buildings and furniture, and also by diffraction. The difficulty which can be experienced in communicating in a reverberant room is similar to that experienced in digital mobile communications, although because of different reflectivities and propagation speeds the "reverberant room" experienced by a mobile communications system may be an entire city. The human ear is extremely adept at interpreting corrupted speech signals. The receivers of digital communications signals are generally less robust.

The presence of several paths by which a signal can propagate between a transmitter and a receiver is known in mobile communications as *multipath propagation*. The received signal is the resulting summation or interference of all the fields incident on the sensor. Even though the average received power may be large, at particular places, times and frequencies the complex summation of the fields may result in a very small signal. This phenomenon is called *multipath fading*.

Multipath propagation is a principal cause of unreliability of mobile radio systems. Not only does it cause short term fluctuations in the received signal power, but it introduces time-dispersion and phase-distortion into the received signal. To obtain the same performance as an unvarying channel, the fading channel must have a considerably larger signal to noise ratio (SNR).

In addition the mobile nature of the receiver and/or transmitter causes the channel characteristics to continually change. This can make adaptive techniques for mitigating the effects of multipath and limited channel bandwidth (such as equalisation) difficult to implement.

The usual approach used to date in designing systems to operate over fading channels is simply to allow a large *fade-margin* when calculating the power budget. Transmitters spend much of the time transmitting several tens of decibels more power than is actually required. Although simple, this technique has several disadvantages. First, it produces far more interference than is necessary, with consequent inefficiencies in channel re-use. Second, it has a dramatic effect on the running time of battery operated mobile transmitters. Third, it results in higher levels of electromagnetic radiation exposure to users. Although most studies do not support the hypothesis that cell phone use is a risk factor for brain tumours, the risk factor for long term heavy users with a delayed onset cannot be excluded [8, 75, 99, 115]. Hence lower operating power levels for transmitters in close proximity to users may be desirable.

Another approach, which has been used with considerable success, is to introduce *diversity* into the channel, in either time, frequency, space or polarisation. In its most simple form, diversity involves sending the intended signal more than once, either at a different frequency, at different time, or from a different location with more than one transmitting antenna (space diversity is more often achieved with multiple receiving antennas). If the receiver can be supplied with several replicas of the same information signal transmitted over *independently* fading channels, the probability that all the signal components will fade simultaneously is reduced considerably. The frequencies, times and locations corresponding to the channels must of course be sufficiently different that the independence requirement is satisfied (in the case of time diversity, this is achieved using interleaving). The signal can be recovered by choosing whichever channel is most reliable at any given instant (selection diversity) or by *combining* the signals from each channel. Achieving diversity by sending the same signal multiple times, is in effect repetition coding. Considerably superior coding techniques may also be used, with consequent improvements in the effect of the diversity.

1.2 Channel Prediction

This thesis investigates an entirely different approach to the problem of fading channels.

1.2.1 Prediction Concept

The *concept* is that if the behaviour of the channel is known, the performance of the system can be optimised for the particular channel being experienced at any instant. Of course for the *system* to optimise its performance, information of the channel state must be common to both the transmitter and the receiver. Since normally only the receiver has this information, there is a requirement for a feedback channel. Unless the channel is changing very slowly, the channel state information available to the transmitter will be outdated, unless the channel behaviour can be *predicted* a short time into the future.

Thus there are many situations in mobile communications in which it would be advantageous for a communications system to have information on when and how a signal will fade in advance of the fade actually occurring. If the timing of a fade is known far enough in advance, there will be sufficient time for corrective action to be negotiated between the transmitter and receiver.

It is perhaps unfortunate that the term *prediction* is most often used in the mobile communications literature to mean *off-line* prediction of the radio coverage which may be expected in certain topographical areas if base stations of a given type and antenna configuration are operating in some given locations. Throughout this thesis, the term prediction is applied strictly to real time prediction.

This concept was first introduced by Vaughan and Bach Andersen in [162], although there have been a small number of papers published by other researchers since, and even a newspaper article [67].

1.2.2 Approach to Predicting Channel Behaviour

The approach taken in this thesis to the problem of how channel behaviour is predicted is discussed in Chapter 2. It may be described briefly as follows.

The key to the approach is finding an appropriate model for the channel. Samples of the channel information are acquired as the mobile moves through some spatial trajectory. These samples are used to estimate the parameters of the model. If the model has been chosen appropriately these parameters will vary much more slowly than the channel itself. The channel model can thus be used to extrapolate beyond the region of the measurements.

1.3 Applications of the Prediction Information

The applications of prediction information can be classified loosely into four areas:

1.3.1 Avoidance of "bad" channels

Even for a channel severely degraded by multipath fading, the percentage of time that a channel is "faded" is comparatively small. (The statistical distribution of the power level has been the subject of much study and is examined briefly in Section 2.2). The most basic corrective action which may be taken then is to simply "avoid" times when the channel does not support reliable communication. This avoidance may take the form of suspension of transmission for a time, or change of time slot for a time division multiple access (TDMA) system, or change of frequency for a frequency division multiple access (FDMA) system (although change of frequency may require prediction of multiple channels which may not be feasible in some systems).

This last case has some resemblance to frequency hopping such as is used in the Groupe Speciale Mobile (GSM) system [122] although with the significant difference that a change would only be made to a channel which is predicted to offer superior performance, rather than to a random choice from available channels.

1.3.2 Adaptive Modulation

Another level of sophistication may be applied to the prediction concept; rather than simply avoiding transmission into channels known to be faded, parameters of the transmission can be adapted on a more continuous basis. These include the power level, modulation type, and coding. The effect of any one or a combination of these is to increase the overall efficiency in the use of power and bandwidth. With the increase in demand for mobile bandwidth, and the growing concern about the effects of non-ionising radiation among mobile telephony users, these are significant gains. A useful review of this approach is [40] (although the present author and the authors of [162] claim originality for what follows).

For example, in truncated channel inversion power control (TCI) [60] transmission is avoided when the instantaneous received channel power falls below a certain threshold, and the transmitted power is proportional to the inverse of the fading channel power when it is above this threshold. Such a power control scheme would of course require knowledge of the transmitted power level at the receiver so that a correction can be made for the effects of the power control on calculation of the channel parameters.

Various adaptive modulation schemes have been proposed and analysed [26, 60, 61, 117, 166]. These increase the data *rate* during periods of high SNR, and decrease the rate during fades. The variation is typically achieved by varying the constellation size of a linear modulation between large bandwidth efficient constellations, and small power efficient ones. The implicit need for prediction for such systems to be viable in a rapidly changing environment does not appear to have been adequately addressed in the literature [54].

Transmit diversity is another application of predicted channel information. In the case of mobile down-link channels, there is more space available at the transmit end for widely spaced antennas, (although placing multiple antennas on the mobile unit has advantages also [102], see Section 2.4). Transmit diversity has the potential to be as effective as receive diversity, provided the channel information is available at the transmitter. For the same reasons as discussed above, this requires the behaviour of each of the channels to be predicted.

There is increasing interest in the literature in multiple input multiple output (MIMO) systems, where there are multiple antennas at both the receiver and transmitter. Several of the proposed methods [118, 119] require channel state information at the transmitter. Once again, in a time varying channel, this information will generally

be obsolete unless some form of prediction is used.

1.3.3 Parametric Equalisation

Equalisation of communication channels requires an adaptive algorithm to adjust a channel model to best reflect the channel impulse response. In the case of linear and decision feedback equalisation (DFE) this is used to update filters to best compensate for the intersymbol interference introduced by the channel. This algorithm is usually the minimum mean square error (MSE) or recursive least squares (RLS) algorithm.

If the system's variation is slow compared to the algorithm's convergence time, the equaliser will be able to adapt to the changing channel.

The popularity of adaptive methods is an indication of the importance of time varying channel equalisation. Despite their wide use and simplicity, however, they have a number of limitations. They are derived under the *stationarity* assumption, and they do not explicitly take into account the time varying nature of the channel. Thus they cannot follow rapid changes in the channel or deep fades.

If the channel is varying faster than the algorithm can converge, convergence is never achieved, and correct equaliser operation is not obtained.

In order to overcome these problems, more elaborate modelling of the time variation of the coefficients is required. The key here is the same as that used in the prediction problem above — find a model for the channel, the parameters of which (or *some* of the parameters of which) vary more slowly than the channel impulse response itself does. This is the approach used in [27, 28, 68, 155, 156]. Equalisers which use such a model for the fading channel may outperform the traditional adaptive schemes. They are not sensitive to rapid variations or deep fades, since these are anticipated by the model.

The approach just described is not strictly *prediction*; however the same channel models and algorithms are applicable. Predicted channel information could in principle also be used to reduce the latency of an equalisation system, in which case the term prediction is again applicable.

Receiver forms which make use of predicted channel information are described in [19]. Equivalent forms exist which do not require channel prediction, and have the advantage of removing a potential source of error if the prediction is not accurate. They have been shown in [19] to obtain this advantage at the expense of considerably more complexity. Once again a reliable prediction method could provide a significant breakthrough.

Yet another potential application of predicted channel information is in construction of an optimal receiver. It is well known that a form of optimal digital receiver includes a matched filter, followed by symbol rate sampling. In the absence of information on the form a signal has taken as a result of transmission through a channel, the required filter to match the received pulse is not known. As a result, the suboptimal approach

of matching to the *transmitted* pulse is often used. Another approach, which results in increased complexity, is fractional spaced sampling. An alternative would be to use predicted channel information along with knowledge of the transmitted pulse shape, to design the optimal matched filter. The application has the advantage over some of those listed above that the prediction would only need to be accurate for a very short time scale.

1.3.4 Improved Combining

An effective technique to overcome the effects of fading is to use spatial diversity — using multiple receive antennas and combining the multiple independently-faded received signals [113, 130, 138]. Such schemes require that the fading process is estimated (and perhaps tracked) in order to determine the appropriate combiner weights [98]. Under slow fading conditions, fading estimation is often performed by a windowing and averaging technique. However, because the fading rate is proportional to the carrier frequency, for high frequencies the high fading rate requires that a shorter window length be used. This results in degraded performance. Furthermore, the estimated fading process obtained by the conventional windowing method also suffers from a time lag. Unless a decision delay of half the window size is introduced, the lag will also lead to performance degradation. Using even short term prediction of a channel could substantially alleviate these difficulties. This approach was first suggested in [74].

1.4 Required Prediction Lengths

In this section the parameters which are used in just one system, GSM, are examined, in order to determine prediction lengths which would be required in order to be effective. Other systems, and certainly any future system which incorporated prediction, will of course have quite different parameters, but this example at least gives an idea of the range of prediction which *could* be required [120, 122].

In the "normalised" column of Table 1.1, times have been normalised to the symbol period T, frequencies to the symbol rate $1/T$, and distances to the wavelength λ. The normalised units allow fair comparison to be made across different systems — the usefulness and achievability of a certain prediction range will depend on the frequency of operation, and the number of symbols transmitted inside the time taken to travel that range.

A mobile transceiver moving at the maximum velocity v travels through vfT/c wavelengths per symbol period (refer to Table 1.1 for the symbol nomenclature), or $vfDS/c \approx 0.95$–1.03 wavelengths per frame. If the protocol allowed the mobile to request a change of frequency for instance, in any up-link frame, and the base station acknowledged the request in the next down-link frame, the mobile would need to have

GSM	Parameter	Specification	Normalised
T	Symbol Period	$1/(270\frac{5}{6})$ ms	1 T
D	Time Slot Period	576.9 μs	156.25 T
f	Operating Frequency	Down-link 935-960 MHz	3.45-3.54 kT^{-1}
		Up-link 890-915 MHz	3.29-3.38 kT^{-1}
Δf	Channel Spacing	200 kHz	0.3846 T^{-1}
S	Time slots per Frame	8	8
τ	Maximum Delay Spread	16 μs	4.33 T
R	Maximum Cell Size	35 km (one way)	63 T
v	Maximum Travel Speed	250 km/h (69 ms^{-1})	761-821 $\mu\lambda T^{-1}$
δf	Maximum Doppler Shift	206-222 Hz	761-821 μT^{-1}

Table 1.1: Parameters of GSM which relate to required prediction ranges.

predicted out to two times the frame duration, or about two wavelengths. In many situations, the mobile velocity is much less than 250 km/h, and so the required prediction range correspondingly less.

If fades (minima of narrowband received power) occur about 0.6–0.7λ apart, this equates to prediction of about three fades, certainly beyond what could be achieved with knowledge of the current level and its gradient and also beyond the correlation distance in most situations (see section 2.3).

On the other hand, if only prediction (and backwards prediction) is only within one time-slot (for example to assist with equalisation), extrapolation forwards and backwards (58 symbols) in time from the training sequence (of length 26 symbols) is required. The prediction in this case is only be required to $58vfT/c \approx 0.044$–0.048 wavelengths. There would however be only the duration of the training sequence, or 0.020 wavelengths on which to base the extrapolation.

The conclusion of this section is then, that in a typical system, for prediction to allow time for negotiation between the transmitter and receiver, it must be accurate for lengths of the order of a wavelength. This length is small enough to believe prediction may be feasible and hence profitable, but large enough to believe it may not be trivial. Prediction for equalisation appliations does not need as great a range, and hence is easier to achieve.

1.5 Structure of this Thesis

The focus of this thesis is on investigating the nature of the mobile channel in order to identify channel models which have parameters which vary more slowly over time than the impulse or frequency responses. Several algorithms are then developed for robust and accurate channel prediction, and their performance is analysed in various simulated and experimental scenarios. The implementation of any of the applications discussed in Section 1.3 is not developed in detail, although reference is made to the applications

at points at which accuracy of estimation has an impact on their operation, which are different at times for different applications.

Various multipath channel models are presented in Chapter 2, commencing with a review of stochastic models. This leads into presentation of the models which are used for prediction, most of them being deterministic models. These models are fundamental to what follows in the remainder of the thesis — they have the properties that they sufficiently reflect reality that their parameters vary much more slowly than does the channel impulse response, but are parsimonious and simple enough that the parameters can be effectively and reliably estimated. The trade-off between optimality and robustness is an important factor in their design.

New closed form expressions for the correlation of the signal at spatially separated points are derived. The issue of the number of effective scatterers present in some models is discussed, as this becomes important later in the thesis.

In Chapter 3 the algorithms are derived in detail. The algorithms are primarily adaptations of various subspace methods of spectrum estimation and array signal processing. Gradient methods are also used. Some of the work which has been published in this area [6, 41, 45] is discussed. The chapter focusses on narrowband algorithms, but also shows how the concepts extend to the broadband case.

In Chapter 4 the results of numerical simulations are presented and discussed. The measurement techniques and equipment are described, and then the results of the prediction algorithms being applied to the measured data. Conclusions about the propagation models, both indoor and outdoor, are deduced. Some more theoretical considerations are then introduced which provide bounds on the accuracy with which channels may be predicted, and some of the factors which define these bounds are highlighted. The effect of rough surface scattering is discussed briefly, and a bound is derived based on consideration of the multipath dimensionality in a defined region.

Chapter 5 moves to a consideration of the prediction accuracy required for just one application. The performance of a channel as a function of the channel impulse response is derived for the case where a maximum likelihood sequence estimator (MLSE) is used for equalisation of the channel. An approximation based on the union of ellipsoids is described. An algorithm for finding the *a priori* most likely error events for particular channels is presented, along with the results of its application.

The effects of imperfect knowledge of a channel as applied to MLSE equalisation are investigated in detail. A result is derived which allows the probability of error to be approximated in the case where the channel is not known perfectly.

In Chapter 6 the results of the research are summarised and suggestions are made to further explore the potential of the prediction concept. The potential for real time channel prediction to become a significant technology in the future is also evaluated.

Chapter 2

Multipath Propagation Models

IN order to be able to characterise a channel in real time, a model must be proposed which allows useful parameters to be updated in real time from the available channel information. In addition, a desirable property of the model is that it has parameters which change only slowly, thus allowing the channel behaviour to be predicted by extrapolation of the parameters.

This chapter establishes the models used to characterise the multipath radio propagation channel. It discusses the key difference between *stochastic* models, which are used to characterise an environment which consists of an ensemble of channels, and *deterministic* models, which characterise the behaviour of a particular instance of a channel in real time. New closed form results for the spatial correlation of a channel are presented.

The effect of multipath propagation on channel performance is detailed, and measures which are used to counteract and even take advantage of multipath propagation are discussed.

2.1 General Channel Models

2.1.1 Single Path

Suppose a signal $s(t)$ is transmitted, and takes time $\tau(t)$ to travel to the receiver. The received signal will then be given by $\zeta_n(t)s(t-\tau(t))$, where $\zeta_n(t)$ is the gain of the propagation path. This gain or attenuation factor includes such factors as space loss, antenna gain and orientation, and polarisation mismatch.

Suppose now that a complex baseband signal $s(t)$ is modulated on a carrier of angular frequency ω_c to obtain $\mathrm{Re}(s(t)e^{j\omega_c t})$. The complex baseband signal corresponding to the received signal $\mathrm{Re}(r(t)e^{j\omega_c t})$ for a single path is given by

$$r(t) = \zeta_n(t)s(t-\tau(t))e^{-j\omega_c\tau(t)}. \tag{2.1}$$

For the next few sections, the equations are made more clear by ignoring noise. Noise will of course be included, with appropriate covariance assumptions, before practical use is made of the models.

2.1.2 Moving Receiver

If the receiver is moving with respect to some reference location, the delay will change, so $\tau(t) = \tau_n - t(v/c)\sin\theta$, where

- c is the speed of propagation of the waves (light).
- θ is the angle between the direction of propagation of the electromagnetic waves and the direction of travel of the receiver minus $\pi/2$. (This awkward description is to make the terminology consistent with conventions in array processing theory, which will be introduced later, and which are based on elevation angles rather than azimuth angles).
- v is the speed of the receiver. In general both v and θ may be functions of time if the receiver changes speed or direction or if the transmitter or local scatterer is close to the receiver, so that the wavefronts must be considered spherical rather than planar.

Thus the received signal will be

$$r(t) = \zeta_n(t) s\left((1 + \frac{v}{c}\sin\theta)t - \tau_n\right) e^{-j\omega_c\tau_n} e^{jt\omega_c\frac{v}{c}\sin\theta}. \tag{2.2}$$

Signals of interest in this thesis are of sufficiently narrow bandwidth (i.e., are slowly varying) that $s\left((1 + (v/c)\sin\theta)t\right) \approx s(t)$ (note that a signal can be wide-band in the sense defined in Section 2.2 and this approximation still be valid). The modulating frequency is also generally high enough that the phase of $e^{j\omega_c\tau_n}$ can be considered to be independent of any practical estimate of τ_n. This phase is typically incorporated as a phase term in the attenuation factor ζ_n, which now becomes complex. The factor $\varpi_n = (\omega_c v\sin\theta)/c = (2\pi v\sin\theta)\lambda = kv\sin\theta$ (where λ is the wavelength of the carrier signal, and $k = 2\pi/\lambda$ is the wave number) is what is known as the *Doppler frequency*, measured in radians per second.

Thus the received signal can be expressed as

$$r(t) = \zeta_n(t) e^{j\varpi_n(t)t} s\left(t - \tau_n(t)\right). \tag{2.3}$$

The term $\varpi_n t$ is sometimes expressed as xu_n where x is the displacement of the mobile receiver, rather than its velocity, and u_n is the *spatial frequency* measured in radians per metre. This gives the position-dependent behaviour of the received signal as a function of delay time and position.

2.1.3 Time Varying Impulse Response Notation

There is some variation in the literature for notation for a time (or shift) varying impulse response. It is quite common [142, 154] to define $h(t,\tau)$ or $h(t;\tau)$ as the response of a system at time t to an impulse at time τ, so the output of a system, given input $s(t)$ is given by the convolution.

$$r(t) = h(t,\tau) \circledast s(t) = \int_{-\infty}^{\infty} h(t,\tau)s(\tau)d\tau, \tag{2.4}$$

or for discrete time

$$r[m] = h[m,k] \circledast s[k] = \sum_{k=-\infty}^{\infty} h[m,k]s[k]. \tag{2.5}$$

These simplify in the time-invariant case to

$$r(t) = \int_{-\infty}^{\infty} h(t-\tau)s(\tau)d\tau, \tag{2.6}$$

or in the shift-invariant case to

$$\text{or} \quad r[m] = \sum_{k=-\infty}^{\infty} h[m-k]s[k]. \tag{2.7}$$

An alternative notation [10] is for $h(\tau,t)$ to represent the response of the system at time t to an impulse at time $t-\tau$. If the transmitted signal $s(t)$ in (2.3) is $\delta(t-t_1)$, the received signal will be

$$r(t) = \zeta_n e^{j\varpi_n t}\delta(t-\tau_n-t_1). \tag{2.8}$$

The symbol τ is defined as $\tau = t - t_1$, and the time-varying impulse response defined as

$$h(\tau,t) = \zeta_n e^{j\varpi_n t}\delta(\tau-\tau_n). \tag{2.9}$$

Then (2.4) and (2.5) become

$$r(t) = h(\tau,t) \circledast s(\tau) = \int_{-\infty}^{\infty} h(t-\tau,t)s(\tau)d\tau \tag{2.10}$$

and

$$r[m] = h[k,m] \circledast s[k] = \sum_{k=-\infty}^{\infty} h[m-k,m]s[k]. \tag{2.11}$$

This notation has two advantages. First, it has greater resemblance to (2.3). Second,

it has greater consistency with the time-invariant case of (2.6) and the shift-invariant case of (2.7). It is this notation which is adopted in this thesis.

2.1.4 Multiple Paths

It is assumed that the receiver sees the sum of the effects of all the paths, or equivalently, the additive interference of waves incident on the receive antenna. Thus the impulse response may be expressed as a summation of each of the individual impulse responses:

$$h(\tau, t) = \sum_{n=1}^{N} \zeta_n e^{j\varpi_n t} \delta(\tau - \tau_n) \tag{2.12}$$

or expressed as a function of delay time and position as

$$h(\tau, x) = \sum_{n=1}^{N} \zeta_n e^{ju_n x} \delta(\tau - \tau_n). \tag{2.13}$$

If there is a continuum of sources present, the cumulative effect of all the paths may be expressed as an integral over both Doppler frequency and delay:

$$r(t) = \iint \zeta(\varpi, \tau) e^{j\varpi t} s(t - \tau) d\varpi d\tau \tag{2.14}$$

$$= \int h(\tau, t) s(t - \tau) d\tau, \tag{2.15}$$

where the time varying impulse response $h(t, \tau)$ is defined as

$$h(\tau, t) = \int \zeta(\varpi, \tau) e^{j\varpi t} d\varpi, \tag{2.16}$$

and $\zeta(\varpi, \tau)$ is now a complex-valued function which is continuous in both ϖ and τ.

The Fourier transform of (2.13) with respect to the base band frequency ω gives the position dependent transfer function:

$$H(\omega, x) = \int h(\tau, x) e^{-j\omega\tau} d\tau = \sum_{n=1}^{N} \zeta_n e^{j(-\omega\tau_n + x u_n)}. \tag{2.17}$$

The Fourier transform of (2.13) and (2.17) with respect to the position x gives a function of delay time and spatial frequency called the scattering function

$$b(\tau, u) = \int h(\tau, x) e^{-jxu} dx = 2\pi \sum_{n=1}^{N} \zeta_n \delta(\tau - \tau_n) \delta(u - u_n), \tag{2.18}$$

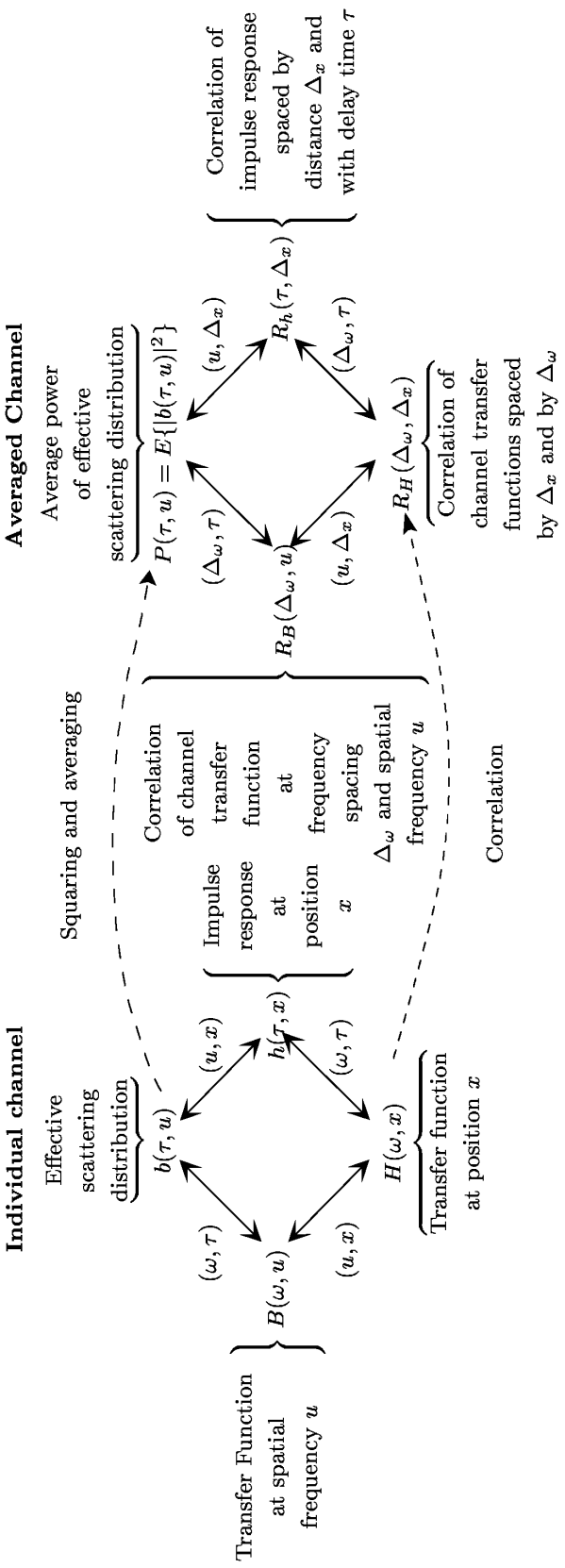


Figure 2.1: Fourier transform relations for the mobile channel functions and for their statistical representations under wide sense stationarity in frequency and position. (From [163].)

and the transfer function

$$B(\omega, u) = \int H(\omega, x) e^{-jxu} dx = 2\pi \sum_{n=1}^{N} \zeta_n e^{-j\omega\tau_n} \delta(u - u_n). \tag{2.19}$$

These relationships apply where the impulse response and transfer functions are continuous rather than discrete functions, and are shown on the left of Fig. 2.1. Functions on opposite corners of the squares are related by two dimensional Fourier transforms.

2.2 Stochastic Channel Characterisations

If the number of paths in the above equations is even moderately large, the impulse response and transfer function behave in a random fashion, hence statistical approaches have been used for their characterisation. The autocorrelation function of the impulse response, for instance, is

$$R_h(\tau_1, \tau_2; x_1, x_2) = E\{h(\tau_1, x_1) h^*(\tau_2, x_2)\}. \tag{2.20}$$

It is often assumed that the channel is *wide sense stationary in the frequency domain*, so the correlation functions do not depend on the choice of frequency ω, but only on the frequency difference Δ_ω. This assumption is equivalent to assuming *uncorrelated scattering* in the delay time (τ) domain (the "US" assumption [14]).

Similarly, *wide sense stationarity in the spatial domain* x corresponds to uncorrelated scattering in the Doppler domain u. This is known in the context of fading channels as the WSS assumption. These two assumptions together (WSSUS) [14] allow the convenience of Fourier transform relations, although the validity of the WSSUS model should always be reviewed in any new situation.

It is assumed in nearly all cases that the statistical quantities being modelled have zero mean, and so the term *covariance* could be used instead of correlation.

If the assumption that scattering is uncorrelated for different delays is applied, the correlation function can be expressed as

$$R_h(\tau, \Delta_x) = R_h(\tau; x, x - \Delta_x) \delta(\tau_1 - \tau_2). \tag{2.21}$$

The Fourier relationships between the correlation functions under these conditions are shown on the right hand side of Fig. 2.1. They are analogous to those of the instantaneous functions shown in the left part of the figure.

If the autocorrelation $R_H(\Delta_\omega, \Delta_x)$ is evaluated at $\Delta_x = 0$, the *spaced-frequency correlation function*[1] $R_H(\Delta_\omega)$ of the channel is obtained. The range of ω over which $R_H(\Delta_\omega)$ is essentially non-zero $(\Delta_\omega)_C$ is known as the *coherence bandwidth* of the

[1] This rather unfortunate terminology is from [116, p708].

channel. This definition is sufficiently loose that a variety of interpretations have been applied to it [90]. The only use made of the term here is to define a *wide-band signal* as one for which the signal bandwidth W is greater than the channel coherence bandwidth, $(W/(\Delta_\omega)_C > 1)$ and a *narrowband signal* as one for which the signal bandwidth is less than the channel coherence bandwidth $(W/(\Delta_\omega)_C < 1)$.

There is of course an interesting philosophical question concerning the relationship between stochastic and deterministic models. To what extent is it appropriate for instance, to talk about the probability of an event $\mathfrak{A}$ occurring, when the opportunity for $\mathfrak{A}$ to occur or not has already passed? The answer is that it is quite appropriate if knowledge of whether the event has occurred is not available. The event to which the probability is assigned is the event $\mathfrak{B}$ that when knowledge *is* obtained, it is found that $\mathfrak{A}$ *did* occur. Such a consideration may bear on the discussion of subjective and objective interpretations of probability, and in particular the *propensity interpretation* of probability found in [59, 96, 114, 143], and the objections to it in [127].

The characterisation of communication channels can be thought of in a similar manner. In principle, given enough knowledge of the distribution of scatterers of (2.12) it is possible to determine in a deterministic manner how a channel will behave. In the absence of such knowledge, a stochastic treatment is quite appropriate, and may be the best that can be hoped for. The *stationarity* of these statistics is another question [22, 23], discussed in Section 2.7.

A large part of this thesis is concerned with determining the extent to which this knowledge *can* be obtained, and hence a deterministic real time model used.

The essence of stochastic models is that they characterise the *ensemble* of channels that a mobile radio receiver may encounter in a region of a particular topography [14, 34, 57, 72, 108]. In stochastic characterisations, the number of paths is assumed to be large.

Statistical descriptions of the narrowband fading channel envelope have been of particular interest throughout the history of mobile radio communications. Some of these are outlined below.

- When a mobile radio channel has a large number of paths, the central limit theorem can be invoked to deduce that the real and imaginary components of the random function $\zeta(\omega, \tau)$ each have a normal distribution. The amplitude of a zero mean complex normal process is Rayleigh distributed [121], hence the term *Rayleigh fading channel* [71, 76].

- Where a single dominant *line of sight* component is present in addition to the indirect components, the complex normal no longer has zero mean [108] and so the narrowband channel envelope has a *Ricean* distribution.

- The central limit theorem applied to the multiplication of random variables has

led to characterisation using the lognormal distribution [71, 109] (the random variables being the coefficients of reflection in the multiple stages of each path).

- If lognormal variation is introduced into the strength of a Rayleigh distributed signal, the *Suzuki* distribution results [70, 71, 109, 144].

- Other distributions which have been applied to the amplitude of the mobile radio channel, without theoretical justification, but with some empirical justification are the Nakagami-m [21, 92, 100], the Weibull [167], and the Generalised Gamma distribution [34].

The usefulness of these characterisation lies in calculating the performance that may be expected from a system operating in a particular environment. If the parameters applicable to a particular amplitude distribution in that environment can be estimated from measurements made in that (or a similar) environment, these parameters can then be used to calculate an expected bit error ratio (BER).

If the probability of obtaining a particular SNR γ for a particular ensemble of channels given by one of the above distributions with parameters $\boldsymbol{\theta}$ is $p(\gamma;\boldsymbol{\theta})$, and the BER for a particular modulation scheme is given by $P(\gamma)$, then the expected BER for that ensemble is given by

$$P(\boldsymbol{\theta}) = E\{P(\gamma)\} = \int_0^\infty P(\gamma)p(\gamma,\boldsymbol{\theta})d\gamma. \tag{2.22}$$

An example given in [116, p707] is for a Rayleigh distributed amplitude and Binary Phase Shift Keying (BPSK).

Such methods have demonstrated reliable calculation of the error performance of systems, and so are useful for system designers in choosing modulation types and coding schemes, and specifying performance limits. They are also useful to planners for locating base stations.

2.3 Spatial Correlation

In establishing a relationship between the channel in two neighbouring locations, the measure typically used is the linear relationship of *correlation*. Spatial correlation is of interest for other reasons also — there is growing interest in the literature in the use of multiple sensors — particularly multiple antennas for transmission and/or reception of wireless signals. As well as diversity reception [76], this includes such areas as MIMO systems, spatio-temporal equalisation, adaptive arrays and space-time coding [50]. Most of the work assumes that each receiving sensor receives uncorrelated signals, and conversely that the signal received from each transmitting source is uncorrelated. A widely used "rule" is that half a wavelength separation is required in order to obtain

de-correlation. This arises from the first null of the sinc(·) function (2.28), which is the spatial correlation function for a three dimensional diffuse field [32].

This section quantifies this *spatial correlation* relationship for a narrowband system. Recall that, by definition, for a narrowband system, the signal bandwidth is less than the coherence bandwidth of the channel. All of the delays τ_n of (2.13) can be considered to be the same, and without loss of generality, set to zero. The correlation function $R_h(\Delta_x)$ in the case of the single path n is then

$$R_h(\Delta_x) = E\left\{\zeta_n e^{ju_n x}\zeta_n^* e^{-ju_n(x-\Delta_x)}\right\} = E\{\zeta_n\zeta_n^*\}e^{ju_n\Delta_x}. \tag{2.23}$$

The correlation coefficient, being the normalised covariance, is defined as

$$\rho(\Delta_x) = \frac{R_h(\Delta_x)}{\sqrt{E\{h(x_1)h^*(x_1)\}E\{h(x_2)h^*(x_2)\}}} = e^{ju_n\Delta_x}. \tag{2.24}$$

At this point the azimuth angle ϕ is introduced, and the Doppler frequency u_n expressed as $k\sin\theta\cos\phi$. The correlation coefficient may then be expressed as

$$\rho(\Delta_x) = e^{j\Delta_x k\sin\theta\ \cos\phi}. \tag{2.25}$$

If waves are incident from an angular sector Θ, over which the (possibly continuous) amplitude distribution of the scatterers is $\zeta(\theta,\phi)$ then the correlation coefficient becomes

$$\rho(\Delta_x) = \iint_{\Theta} \zeta(\theta,\phi)e^{j\Delta_x k\sin\theta\ \cos\phi}\sin\theta\, d\theta\, d\phi \tag{2.26}$$

In the special case that $\zeta(\theta,\phi) = \delta(\theta-\pi/2)/(2\pi)$ (a uniform ring of scatterers in the horizontal plane), using (3.915-2) and (8.411-1) of [64]

$$\rho(\Delta_x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\Delta_x k\cos\phi}d\phi = J_0(k\Delta_x) \tag{2.27}$$

In the special case that $\zeta(\theta,\phi) = 1/(4\pi)$ (a uniform sphere of scatterers), using (6.681-8) and (8.464-1&2) of [64]

$$\begin{aligned}\rho(\Delta_x) &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\int_{0}^{\pi} e^{j\Delta_x k\sin\theta\cos\phi}\sin\theta\ d\theta\, d\phi \\ &= \frac{1}{2}\int_0^{\pi} J_0(k\Delta_x\sin\theta)\sin\theta\ d\theta \\ &= \frac{\sin(k\Delta_x)}{k\Delta_x} \triangleq \mathrm{sinc}(k\Delta_x), \quad \text{for } k\Delta_x \neq 0,\end{aligned} \tag{2.28}$$

and $\rho(0) = \mathrm{sinc}(0) \triangleq 1$.

2.3.1 Spatial Correlation for General Distributions of Scatterers

Several approaches have been used to find the spatial correlation function in the case of signals confined to a limited azimuth or elevation [79]. In this section a modal analysis approach is presented which can obtain closed form expressions for the spatial correlation function for narrowband signals for a wide variety of scattering distribution functions.

The expression for the spatial correlation is first recast in a form which is suitable for applying modal analysis techniques. Consider two sensors located at points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, with $\|\boldsymbol{x}_2 - \boldsymbol{x}_2\| = \Delta_x$. Let $s_1(t)$ and $s_2(t)$ denote the complex envelope of the received signal at the two sensors, respectively. Then the normalised spatial correlation function between the complex envelopes of the two received signals, is defined by

$$\rho(\boldsymbol{x}_2 - \boldsymbol{x}_1) = \frac{E\{s_1(t)s_2^*(t)\}}{E\{s_1(t)s_1^*(t)\}}. \tag{2.29}$$

Consider a general scattering environment with a large number of scatters distributed far from the two sensors. If the transmitted signal is a narrowband signal $e^{i\omega t}$, then the received signal at the lth receiver is

$$s_l(t) = e^{j\omega t} \int A(\hat{\boldsymbol{y}}) e^{-jk\boldsymbol{x}_l \cdot \hat{\boldsymbol{y}}} d\hat{\boldsymbol{y}}, \;\; l = 1, 2 \tag{2.30}$$

where $\hat{\boldsymbol{y}}$ is a unit vector pointing in the direction of wave propagation, and $A(\hat{\boldsymbol{y}})$ is the *complex gain* of scatterers as a function of direction which captures both the amplitude and phase distribution. Also note that the integration in (2.30) is over a unit sphere in the case of a 3-dimensional multipath environment or a unit circle in the 2-dimensional case. Substitute (2.30) in (2.29) and assume that scattering from one direction is independent from another direction, to get

$$\rho(\boldsymbol{x}_2 - \boldsymbol{x}_1) = \int \mathcal{P}(\hat{\boldsymbol{y}}) e^{jk(\boldsymbol{x}_2 - \boldsymbol{x}_1)\cdot \hat{\boldsymbol{y}}} d\hat{\boldsymbol{y}}, \tag{2.31}$$

where

$$\mathcal{P}(\hat{\boldsymbol{y}}) = \frac{E\{|A(\hat{\boldsymbol{y}})|^2\}}{\int E\{|A(\hat{\boldsymbol{y}})|^2\} d\hat{\boldsymbol{y}}}, \tag{2.32}$$

is the normalised average power of a signal received from direction $\hat{\boldsymbol{y}}$, or the *distribution function* of scatterers over all angles.

2.3.2 Three Dimensional Scattering Environment

In order to gain a better understanding of spatial correlation, a spherical harmonic expansion of plane waves in now used, which is given by [31]

$$e^{jk\boldsymbol{x}\cdot\hat{\boldsymbol{y}}} = 4\pi \sum_{n=0}^{\infty} j^n j_n(k\|\boldsymbol{x}\|) \sum_{m=-n}^{n} Y_{nm}(\hat{\boldsymbol{x}}) Y_{nm}^*(\hat{\boldsymbol{y}}), \tag{2.33}$$

where $\hat{\boldsymbol{x}} = \boldsymbol{x}/\|\boldsymbol{x}\|$, $j_n(r) \triangleq \sqrt{\pi/2r} J_{n+1/2}(r)$ are spherical Bessel functions, and

$$\begin{aligned} Y_{nm}(\hat{\boldsymbol{x}}) &\equiv Y_{nm}(\theta_{\boldsymbol{x}}, \phi_{\boldsymbol{x}}) \\ &\triangleq \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\theta_{\boldsymbol{x}}) e^{jm\phi_{\boldsymbol{x}}}, \end{aligned} \tag{2.34}$$

where $\theta_{\boldsymbol{x}}$ and $\phi_{\boldsymbol{x}}$ are the elevation and azimuth respectively of the unit vector $\hat{\boldsymbol{x}}$, and where $P_n^m(\cdot)$ are the associated Legendre functions of the first kind. Equations (2.31) and (2.33) may be combined to obtain

$$\rho(\boldsymbol{x}_2 - \boldsymbol{x}_1) = 4\pi \sum_{n=0}^{\infty} j^n j_n(k\Delta_x) \sum_{m=-n}^{n} \beta_{nm} Y_{nm}\left(\frac{\boldsymbol{x}_2 - \boldsymbol{x}_1}{\Delta_x}\right) \tag{2.35}$$

where

$$\beta_{nm} = \int \mathcal{P}(\hat{\boldsymbol{y}}) Y_{nm}^*(\hat{\boldsymbol{y}}) d\hat{\boldsymbol{y}}. \tag{2.36}$$

The fact that the higher order spherical Bessel functions (and in Section 2.3.3 the higher order Bessel functions) have small values for arguments near zero, means that to evaluate the correlation for points near each other in space, only a few terms in the sum need to be evaluated in order to obtain a very good approximation.

Three Dimensional Omni-directional Diffuse Field

If waves are incident on the two points from all directions in 3-dimensional space, then (2.35) reduces to a single term, and so the correlation coefficient is given by $\rho = j_0(k\Delta_x) = \mathrm{sinc}(k\Delta_x)$, the same classic result [32] as (2.28). The first zero crossing is at $\lambda/2$.

Uniform Limited Azimuth/Elevation Field

Without loss of generality, the co-ordinate system may be chosen so that $\theta_{\boldsymbol{x}_2-\boldsymbol{x}_1} = \pi/2$ and $\phi_{\boldsymbol{x}_2-\boldsymbol{x}_1} = 0$. If the scatters are uniformly distributed over the sector $\Omega \in \{(\theta, \phi); \theta \in$

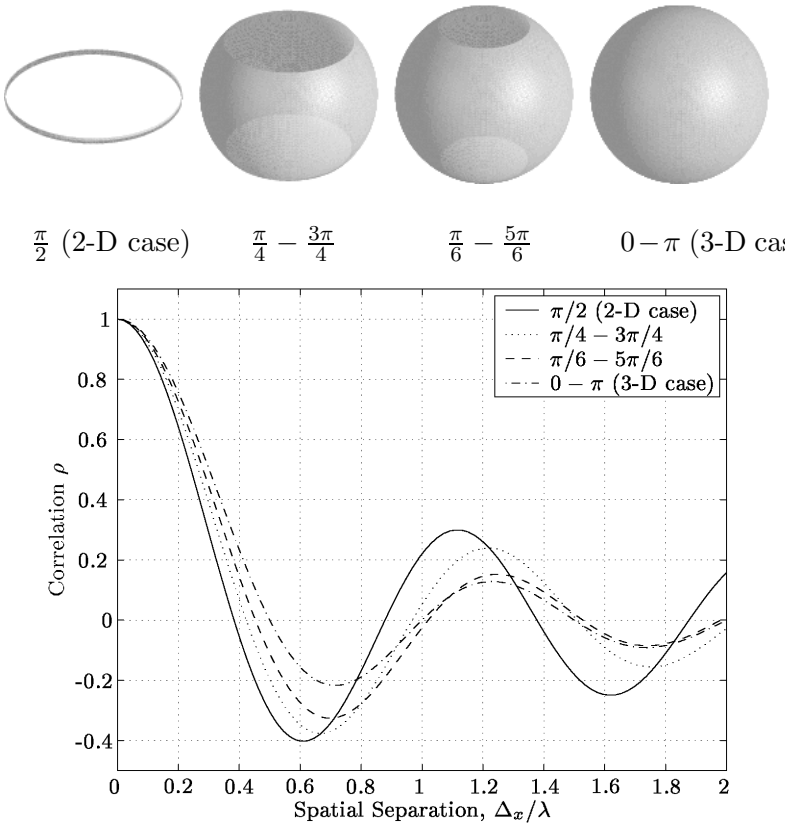


Figure 2.2: Spatial Correlation versus Separation for different elevation ranges. This figure shows that the spread of interference in elevation plays only a secondary role in influencing spatial correlation and hence diversity within the horizontal plane.

$[\theta_1, \theta_2], \phi \in [\phi_1, \phi_2]\}$, then the correlation can be expressed as

$$\rho(\boldsymbol{x}) = \frac{1}{(\cos\theta_1 - \cos\theta_2)} \sum_{n=0}^{\infty} (2n+1) j^n j_n(kr) \left(P_n(0) \int_{\theta_1}^{\theta_2} P_n(\cos\theta) \sin\theta \, d\theta \right.$$
$$\left. + 2 \sum_{m=1}^{n} \frac{(n-m)!}{(n+m)!} \frac{\sin(m\phi_2) - \sin(m\phi_1)}{m(\phi_2 - \phi_1)} P_n^m(0) \int_{\theta_1}^{\theta_2} P_n^m(\cos\theta) \sin\theta \, d\theta \right) \quad (2.37)$$

Note that (2.37) is actually expressible in closed form (the integrals can be evaluated for example by using recursions in m and n). Hence, this result can be used to build up the result for a general scattering situation. Any arbitrary scattering can be regarded as the limiting summation of a weighted set of uniformly distributed incremental solid angle contributions.

The case where energy is arriving from all azimuth directions (uniformly) but the elevation spread is in some range of angles either side of zero elevation is shown in Fig. 2.2. The spatial correlation is shown for four sets of elevation spread: π, $\pi/2$, $\pi/3$ and 0, each centred on $\theta = \pi/2$. Generally it can be concluded that multipath spread

in elevation does not have a great effect on spatial correlation in the horizontal plane. Multipath which may have some small elevation support (in many practical situations not much would be expected) may be modelled as only coming from the horizontal plane.

Spherical Harmonic Model

The distribution function $\mathcal{P}$ in (2.32) may be expressed as a weighted sum of orthonormal spherical basis functions

$$\mathcal{P}(\hat{\boldsymbol{y}}) = \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \gamma_{n'm'} Y_{n'm'}(\hat{\boldsymbol{y}}), \tag{2.38}$$

where the coefficients $\gamma_{n'm'}$ are chosen so that the distribution function is normalised. In many cases the number of basis functions required to approximate the distribution function may be quite small. By substitution into (2.36), β_{nm} can be simply expressed in terms of the coefficients $\gamma_{n'm'}$ as

$$\begin{aligned} \beta_{nm} &= \int \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} \gamma_{n'm'} Y_{n'm'}(\hat{\boldsymbol{y}}) Y^*_{nm}(\hat{\boldsymbol{y}}) d\hat{\boldsymbol{y}} \\ &= \gamma_{nm}. \end{aligned} \tag{2.39}$$

2.3.3 Two Dimensional Scattering Environment

If the fields arrive only from the azimuthal plane, it is more convenient to consider the 2 dimensional modal expansion [31]

$$e^{jk\boldsymbol{x}\cdot\hat{\boldsymbol{y}}} = \sum_{m=-\infty}^{\infty} j^m J_m(k\|\boldsymbol{x}\|) e^{jm(\phi_x - \phi_y)} \tag{2.40}$$

where ϕ_x and ϕ_y are the angles of $\boldsymbol{x}$ and $\hat{\boldsymbol{y}}$. Equations (2.31) and (2.40) may be combined to obtain

$$\rho(\boldsymbol{x}_2 - \boldsymbol{x}_1) = \sum_{m=-\infty}^{\infty} \alpha_m J_m(k\Delta_x) e^{jm\phi_{12}}, \tag{2.41}$$

where

$$\alpha_m = j^m \int_0^{2\pi} \mathcal{P}(\phi) e^{-jm\phi} d\phi, \tag{2.42}$$

$\mathcal{P}(\phi)$ is the distribution function equivalent to $\mathcal{P}(\hat{\boldsymbol{y}})$ in (2.32), and ϕ_{12} is the angle of the vector connecting $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. Without loss of generality, this can be considered to be zero.

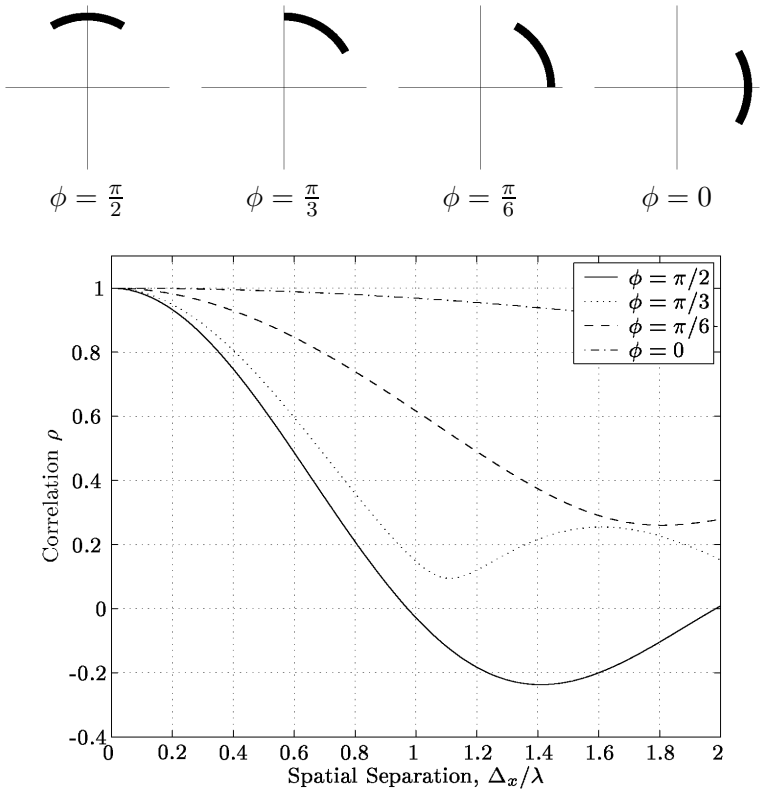


Figure 2.3: Spatial Correlation versus Separation (Wavelengths) for an azimuthal source range of $\pi/3$ centred at various values of the azimuth ϕ. For $\phi = 0$, the source direction is co-linear with the two points, while for $\phi = \pi/2$ the source direction is broadside to the two points. In the upper subfigures, the shaded region indicates the source range, and the two locations are separated on the horizontal axis. For $\phi = \pi/2$ the real part of the correlation coefficient is shown, as the imaginary component is zero. For the other cases the magnitude is shown.

Two Dimensional Omni-directional Diffuse Field

For the special case of scattering over *all* angles in the plane containing two points, (2.41) reduces to a single term, and so the correlation coefficient is given by $\rho = J_0(k\Delta_x)$, another classical result [76]. As can be seen by an examination of Fig. 2.2, $J_0(\cdot)$ and sinc$(\cdot)$ are qualitatively similar.

Uniform Limited Azimuth Field

If energy arrives uniformly from a restricted range of azimuth angles $(\phi_0 - \Delta, \phi_0 + \Delta)$,

$$\alpha_m = e^{jm(\pi/2-\phi_0)}\,\mathrm{sinc}(m\Delta). \tag{2.43}$$

This result is equivalent to that derived in [128].

If attention is restricted to multipath only in the horizontal plane and a restricted

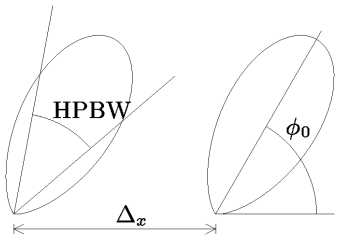


Figure 2.4: The $\cos^{2p}((\phi - \phi_0)/2)$ amplitude pattern

range of azimuth angles is considered then one can better understand the effect of a limited range of multipath and the effect of orientation of the sensors on multipath. The variation with azimuth angle is shown in Fig. 2.3. Where the correlation contains an imaginary component, the magnitude has been shown, since it is the magnitude which provides a measure of the linear relationship between the two signals. The range of azimuth angles is $\pi/3$ centred on the four values of ϕ: $\pi/2$, $\pi/3$, $\pi/6$ and 0. It is immediately apparent by comparison with Fig. 2.2 that restricting the azimuthal range considerably increases the correlation between the signals at the two locations. This is particularly the case when the two sensors are end-fire relative to the central signal direction.

von Mises Distributed Field

Non-isotropic scattering in the azimuthal plane may be modelled by the von Mises distribution [e.g., 1], for which the density is given by

$$\mathcal{P}(\phi) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\phi - \phi_0)}, \quad |\phi - \phi_0| \leq \pi, \tag{2.44}$$

where ϕ_0 represents the mean direction, $\kappa > 0$ represents the degree of non-isotropy, and $I_m(\kappa)$ is the modified Bessel function of the first kind. In this case, using (3.937) of [64]

$$\begin{aligned} \alpha_m &= \frac{j^m}{2\pi I_0(\kappa)} \int_0^{2\pi} e^{\kappa \cos(\phi - \phi_0)} e^{jm\phi} d\phi \\ &= e^{jm(\pi/2 - \phi_0)} \frac{I_{-m}(\kappa)}{I_0(\kappa)}. \end{aligned} \tag{2.45}$$

$\cos^{2p}\phi$ Distributed Field

The restriction in the range of angles in which scatterers are located may be due to a non-isotropic antenna gain pattern. For instance, a real-world amplitude pattern may

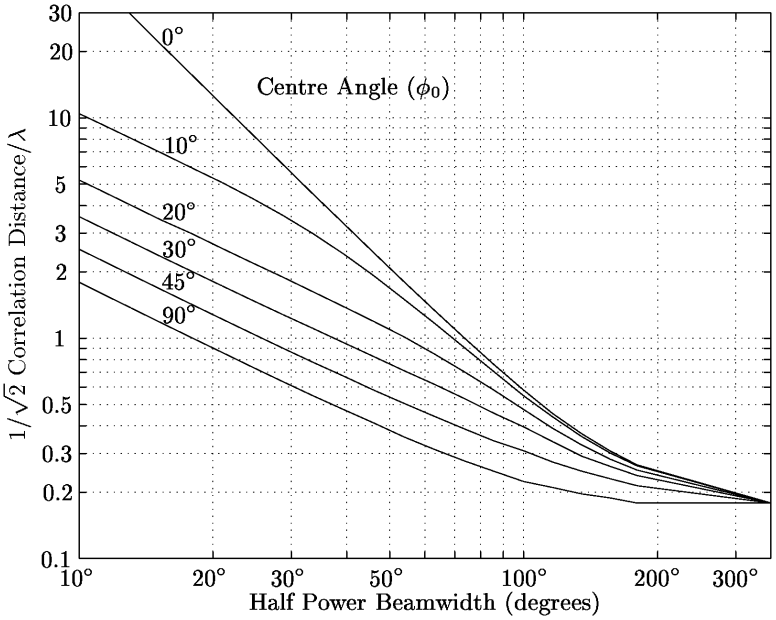


Figure 2.5: Correlation distance for a directional scenario. The directionality is due to the product of the incident power distribution and the pattern of the antenna.

be approximated by the single lobed pattern [160, 161]

$$\mathcal{P}(\phi) = Q\cos^{2p}\left(\frac{\phi-\phi_0}{2}\right), \quad |\phi-\phi_0| \le \pi, \tag{2.46}$$

where Q is a normalisation constant. Using (335.19) of [66],

$$\begin{aligned}\alpha_m &= \frac{j^m 2^{2p-1}\Gamma^2(p+1)}{\pi\Gamma(2p+1)}\int_{\phi_0-\pi}^{\phi_0+\pi}\cos^{2p}\left(\frac{\phi-\phi_0}{2}\right)e^{-jm\phi}d\phi \\ &= e^{jm(\pi/2-\phi_0)}\frac{\Gamma^2(p+1)}{\Gamma(p-m+1)\Gamma(p+m+1)}.\end{aligned} \tag{2.47}$$

The half power beam width of this azimuthal pattern is $4\cos^{-1}(2^{(-1)/(2n)})$ (see Fig. 2.4). Defining the *correlation distance* as the distance at which the correlation coefficient decreases to $\sqrt{1/2}$, the dependence of this correlation distance on the beam width and azimuth angle ϕ_0 is easily calculated using this method, and is represented in Fig. 2.5 (similar to Figure 4 of [161]). The correlation can be seen to be (very approximately) inversely proportional to the beam width of the antennas and/or width of distribution of the scatterers, and to increase for beams near end-fire to the two locations being correlated.

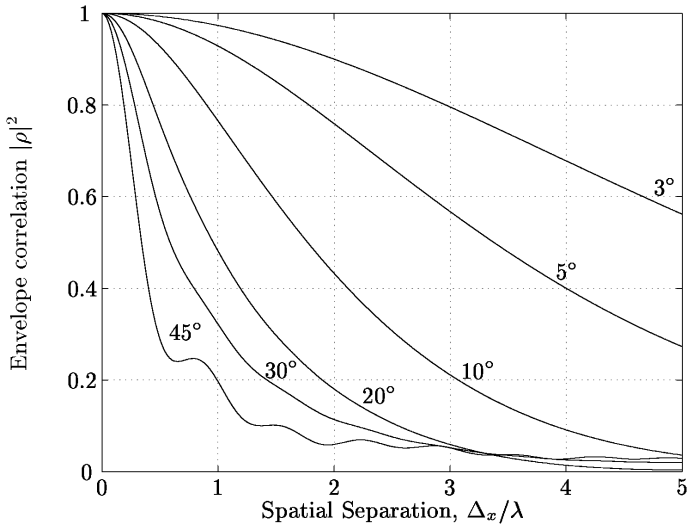


Figure 2.6: Correlation for angle of incidence 60° from broadside against separation of sensors with angular spread of the scattering distribution as parameter, based on a Laplacian power distribution function.

Laplacian Distributed Field

In [112] the symmetric Laplacian distribution is proposed as a realistic model of the power distribution function in some circumstances. Here

$$\mathcal{P}(\phi) = \frac{Q}{\sqrt{2}\sigma} e^{-\sqrt{2}|\phi-\phi_0|/\sigma}, \quad |\phi - \phi_0| \leq \frac{\pi}{2}, \tag{2.48}$$

where Q is a normalisation constant. It is straightforward to show that

$$\alpha_m = e^{jm(\pi/2-\phi_0)} \frac{(1-(-1)^{\lceil m/2 \rceil}\xi F_m)}{(1+\sigma^2 m^2/2)(1-\xi)} \tag{2.49}$$

where $\xi = e^{-\pi/(\sqrt{2}\sigma)}$, and $F_m = 1$ for m even, and $F_m = m\sigma/\sqrt{2}$ for m odd.

To demonstrate the power of the technique, Figure 5b from [53] is reproduced in Fig. 2.6, except that the distribution used here is the Laplacian, considered in [53] to be realistic but mathematically intractable. Approximately 100 terms of the summation (2.41) were required to obtain results to about 10 significant figures at the largest spatial separation on the figure (5λ), or 70 terms for 5 significant figures. For separations up to 2λ only 40–50 terms are required. The angular spread used to distinguish between the data sets in the figure is defined as the square root of the variance, which for this distribution is given by

$$S_\sigma^2 = \frac{1}{1-\xi}\left(\sigma^2 - \frac{\xi}{4}(\pi^2 + 4\sigma^2 + \sqrt{8}\pi\sigma)\right). \tag{2.50}$$

Normally Distributed Field

Several researchers [79, 160] have used the truncated normal distribution for modelling the distribution of scatterers, thus

$$\mathcal{P}(\phi) = \frac{Q}{\sqrt{2\pi}\sigma}\, e^{-\frac{(\phi-\phi_0)^2}{2\sigma^2}}, \quad |\phi - \phi_0| \leq \frac{\pi}{2}, \tag{2.51}$$

where σ is the standard deviation, and Q is a normalisation constant. It can be shown using (313.6) of [65], and the symmetries of the error function for complex arguments discussed in [55], that

$$\alpha_m = e^{jm(\pi/2-\phi_0)-m^2\sigma^2/2}\frac{\mathrm{Re}\left(\mathrm{erf}\left(\frac{\pi/2+jm\sigma^2}{\sqrt{2}\sigma}\right)\right)}{\mathrm{erf}\left(\frac{\pi/2}{\sqrt{2}\sigma}\right)}. \tag{2.52}$$

It can be seen that contrary to the assertion of [53], a closed form solution for the correlation function is available for distribution functions other than the normal (such as the $\cos^{2p}\phi$ distribution), and that the normal distribution is neither a simple nor natural choice. If however, the beam is narrow, a good approximation for the normal case can be obtained by performing integration over the domain $(-\infty, \infty)$, since the tails will cause very little error. In this case using (337.3) of [66], it is straightforward to show that

$$\alpha_m \approx e^{jm(\pi/2-\phi_0)-m^2\sigma^2/2}. \tag{2.53}$$

Cylindric Harmonic Model

The distribution function $\mathcal{P}$ may be expressed as the sum of orthogonal basis functions

$$\mathcal{P}(\phi) = \sum_{k=-\infty}^{\infty} \gamma_k e^{jk\phi}, \tag{2.54}$$

where the coefficients γ_k are chosen to that the distribution function is normalised. In many cases the number of basis functions required to approximate the distribution function may be quite small. The coefficients α_m in (2.41) can be simply expressed in terms of the coefficients γ_m as

$$\alpha_m = j^m 2\pi\gamma_m. \tag{2.55}$$

2.3.4 Mutual Coupling

If the correlation of interest is between two nearby antennas, rather than between measurements from one antenna moved to two different locations, there can be mutual

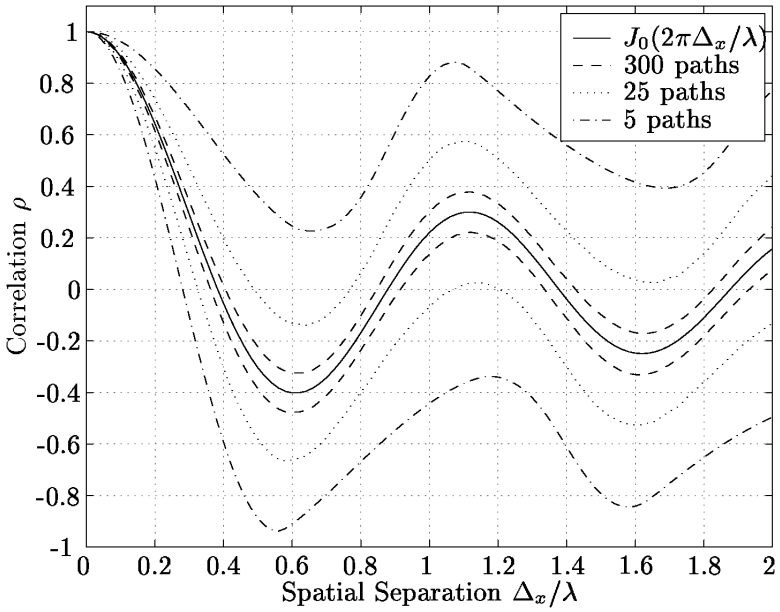


Figure 2.7: Upper and lower 95% Confidence Intervals for correlation resulting from a finite number of scatterers in a azimuthal only environment (only the real part is shown).

coupling effects between the sensors. For a given spacing, mutual coupling of terminated antennas can actually *decrease* the correlation between the sensors from that calculated using the expressions above. One interpretation of this is that the presence of other antennas creates a slow-wave structure which in effect decreases the wavelength of the signal in their vicinity, and thus increases the number of wavelengths separation between the elements. This phenomenon was first reported in [164], and has recently been reported in [145]. If it proves true in general that mutual coupling effects decrease correlation, then the closed form expressions of the previous sections may be considered as upper bounds on the correlation between antennas in an electromagnetic field.

For most situations, the antennas of mobile receivers are omni-directional, and in urban environments the scattering scenario is approximately isotropic, hence the correlation distance may not exceed a few tenths of a wavelength.

2.3.5 Correlation for a Small Number of Scatterers

It is of interest also to note that the correlation can resemble that for a uniform diffuse field, even when the number of paths for a signal is quite small. The case studied here is for paths incident on the receiving antenna from the azimuthal plane only.

Since the correlation in the case of one path is given by (2.25), a uniform distribution of ϕ implies that the probability density function of ρ_{R1}, the real part of ρ for one path is given (using [104, pp126,189]) by finding the probability density for a function of a

random variable. Thus

$$p_{\rho_{R1}}(\rho_{R1}) = \sum_m \sum_{s=\pm 1} \frac{1}{2\pi^2 \frac{\Delta_x}{\lambda}\sqrt{1-\rho_{R1}^2}\sqrt{1-(\frac{\Delta_x}{\lambda}(m+\frac{s}{2\pi}\cos^{-1}(\rho_{R1})))^2}}, \tag{2.56}$$

where the summation over m is for all values of m for which $|\frac{\Delta_x}{\lambda}(m+\frac{s}{2\pi}\cos^{-1}(\rho_{R1}))| < 1$. The probability density function for N paths is then given by the N-fold convolution of the one-path distribution.

Fig. 2.7 shows the 95% confidence intervals of the real part of the correlation function for the case of a finite number of scatterers which are uniformly distributed in the azimuthal plane. The curve for $N = 5$ is produced using N-fold convolution — the others by Monte Carlo simulation.

It can be seen that even for quite a small number of paths, the correlation, especially for small separations, can closely resemble that of a diffuse field. Numerous researchers have also found that the sum of only a small number of sinusoids (e.g., six) is required in order to produce an approximation to the Rayleigh statistics that are so familiar in mobile radio literature [15, 76, 110, 111, 126, 133, 135]. In fact it is shown in Section 4.7 that within a certain region, the "illusion" of there being only a small number of paths is effectively complete. It is also widely assumed in the literature that there are only a small number of significant scatterers in a typical mobile communications environment [27, 44, 45, 58, 71]. The experimental justification for this is rather sparse or may be interpreted as contradicting this assumption [37–39, 47]. The fact that $\mathcal{A} \implies \mathcal{B}$ of course does *not* mean that $\mathcal{B} \implies \mathcal{A}$. The strongest conclusion that can be drawn is that the familiar distributions and correlations may arise in situations where there are either many or few paths.

A numerical study of the relationship between correlation and prediction is now presented. The parameters used for the study are listed here, although the definitions and methods used are defined later in the thesis. Several sets of data points were synthesised, each covering 10 wavelengths and based on between 5 and 30 point sources with an SNR of 40 dB. The length of successful prediction for each was calculated using the methods outlined later in the thesis. Also the root-mean-square (RMS) value of the estimated correlation function (the mean being taken over 10 wavelengths) was calculated. These two quantities are shown in a scatter plot in Fig. 2.8. It might have been expected that there would be some relationship — that if the mean correlation is high, the channel would be highly predictable. It is clear from Fig. 2.8 that there is no strong relationship between these two quantities; a scenario may have a small RMS correlation and yet be "predictable", or a large RMS correlation and be "unpredictable". Some other factors which have a greater influence on predictability are discussed later in Sections 4.2.3 and 4.5.

To conclude this section it is noted that if only a small number of paths is present,

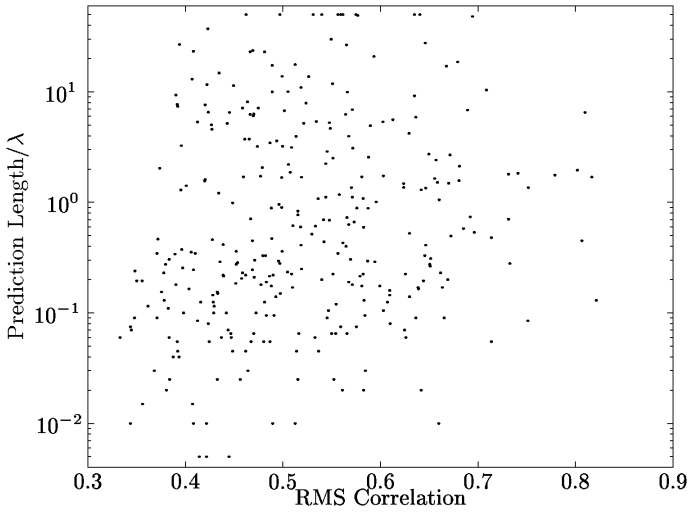


Figure 2.8: Relationship between Correlation and Prediction Length. Although one might expect a strong relationship to exist, none is apparent.

the spatial correlation may fall rapidly for increasing separation of points. This does not necessarily appear to provide a bound to deterministic prediction methods. Such a method may perform quite well if there are few paths, because there may be only a few parameters to estimate in order to obtain reliable long-range prediction. Though the *expectations* of the spatial correlation and the prediction error may be related, there is not a strong relationship for *individual* scenarios.

2.4 Diversity

In this section the important equations relating to diversity are derived, following the format of [130]. This demonstrates the considerable improvement in performance that can be obtained using diversity reception of fading signals. Channel prediction may be considered as a complementary, or in some instances, competing technique, and so the comparison of these techniques is important.

Just one method of utilising diversity — maximal ratio combining — is treated here. Other methods are equal gain combining and selection diversity . In equal gain combining the signals are phase aligned, but not scaled according to their respective signal to noise ratios. In selection combining, the single signal with the largest signal to noise ratio is chosen (or the signal choice is changed if the SNR falls below a certain threshold).

The idea behind diversity is to use several *independent* copies of a fading signal. The probability that the *combined* signal is below the threshold of reliable detection can in this way be reduced well below that for a single source. In maximal-ratio combining, optimality is defined as achievement of the maximum possible SNR at the combiner

output at each instant.

The transmitted signal is denoted by $s(t)$, and the M copies of the received signal which are denoted by $r_m(t)$ have power g_m. The noise terms are $\eta_m(t)$, with noise power N_m. It is assumed that the noise terms are all independent. Optimality is then achieved when the weights for each copy of the received signal are given by

$$\alpha_m = \chi \frac{g_m^*}{N_m} \tag{2.57}$$

where χ is any arbitrary complex constant. In this case the combining results in an instantaneous output SNR which is the sum of the instantaneous SNRs on the individual branches: $\gamma = \sum_{m=1}^{M} \gamma_m$. The probability density function (pdf) of the combined SNR can be derived assuming that the individual branch noise signals are complex normal, and so the individual branch SNRs are Rayleigh distributed. Since γ and all the γ_m are all positive, a Laplace transform may be used to evaluate the characteristic function:

$$\begin{aligned} F(s) &= \int_0^\infty e^{-s\gamma} p(\gamma) d\gamma = E\{e^{s\gamma}\} = \prod_{m=1}^{M} E\{e^{-s\gamma_m}\} \\ &= \prod_{m=1}^{M} \int_0^\infty e^{-s\gamma_m} \frac{1}{\Gamma_m} e^{-\frac{\gamma_m}{\Gamma_k}} d\gamma_m \\ &= \prod_{m=1}^{M} \frac{1}{1+s\Gamma_m}, \end{aligned} \tag{2.58}$$

where Γ_m is the average SNR for branch m, and it is assumed that each of the γ_m are uncorrelated so that the mean of the product is the product of the means. Inversion of this transform gives

$$p(\gamma) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{s\gamma}{\prod_{m=1}^{M}(1+s\Gamma_m)} ds \quad c \geq 0. \tag{2.59}$$

The integral can be solved using the method of residues . If the Γ_k are all equal, $\Gamma_k = \Gamma$, so that

$$\begin{aligned} p(\gamma) &= \frac{1}{\Gamma^M} \frac{1}{(M-1)!} \left[\frac{\partial^{M-1}}{\partial s^{M-1}} e^{s\gamma} \right]_{s=-\frac{1}{\Gamma}} \\ &= \frac{1}{(M-1)!} \frac{\gamma^{M-1}}{\Gamma^M} e^{-\frac{\gamma}{\Gamma}}, \end{aligned} \tag{2.60}$$

which is the chi-square distribution of order $2M$. If the Γ_m are all unequal

$$p(\gamma) = \sum_{m=1}^{M} \frac{1}{\kappa_m} e^{-\frac{\gamma_m}{\Gamma_m}}, \qquad \kappa_m = \Gamma_m \prod_{\substack{m'=1 \\ m' \neq m}}^{M} 1 - \frac{\Gamma_{m'}}{\Gamma_m}. \tag{2.61}$$

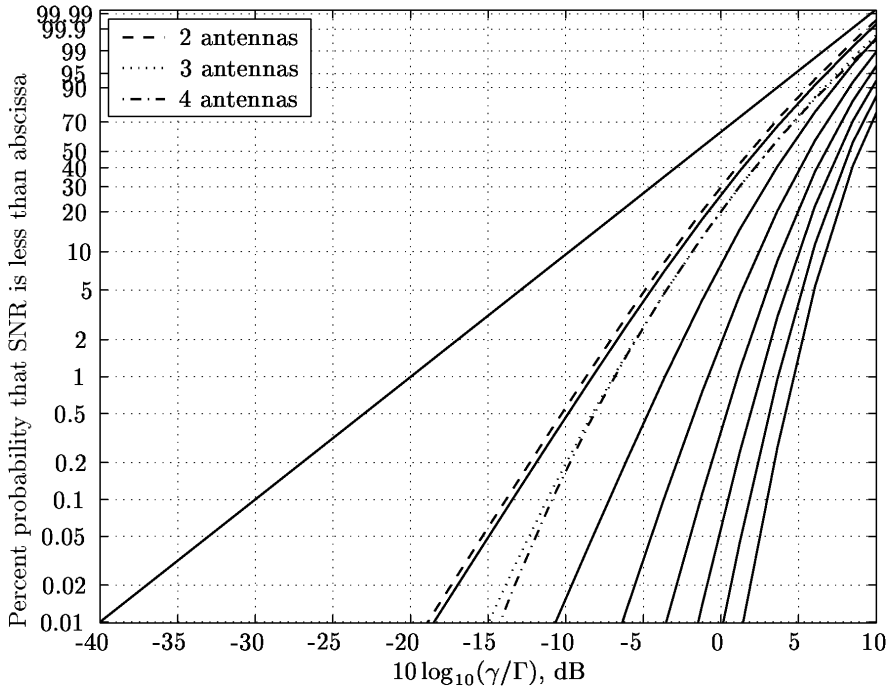


Figure 2.9: Cumulative probability of Maximal Ratio combined signals with 1 to 8 uncorrelated equal SNR branches shown with solid lines from left to right respectively. The broken lines are the measured data, which have 2, 3, or 4 correlated, unequal SNR branches.

The cumulative distribution function (cdf) is given by

$$p(\gamma < x) = \sum_{m=1}^{M} \frac{\Gamma_m}{\kappa_m} \left(1 - e^{-\frac{x}{\Gamma_m}}\right). \tag{2.62}$$

The case of correlated signals can be treated by using eigendecomposition of the signal correlation matrix, which can be estimated from the signal plus noise correlation matrix. Eigendecomposition is a similarity transform and so conserves the energy of the original matrix.

The cdf for the equal SNR case is presented in the solid lines of Fig. 2.9 on "Rayleigh paper" so that the single antenna case is a single straight line.

Antenna diversity is widely used at base stations in mobile communications. At the mobile, the use of diversity has received some attention, but its commercial use is not yet widespread. There are two main differences between antenna diversity design for the base and for the mobile. The first is that usually there is much less space available at the mobile, particularly if the unit is to be hand-held. The second relates to the scattering environments. Generally the base station will receive signals from a limited range of arrival angles, whereas the mobile is often moving in an environment with many nearby scatterers, and so signals will arrive in a more omni-directional manner.

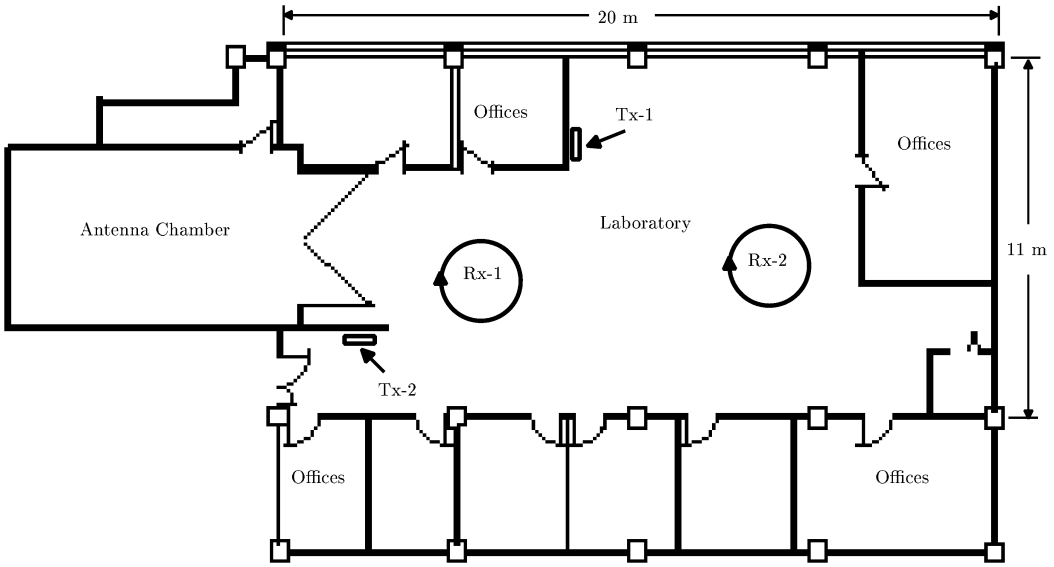


Figure 2.10: Floor plan of the indoor measurement environment for evaluating the effectiveness of diversity in a compact mobile receiver.

As shown in Section 2.3, this results in shorter correlation lengths, and thus a lesser requirement for large antenna spacing.

An experiment was performed[1] to evaluate the diversity performance of a handset with multiple antennas in an indoor environment. The transmitting frequency used was 1.9 GHz corresponding to current PCS systems. The transmitting antenna was a commercially available 1850-1990 MHz bow-tie dipole mounted in front of a ground plane, vertically polarised and wall mounted at a height of 2.5 m. The receiving handset had dimensions 100 × 50 × 20 mm rotated at the end of a 1.2 m wooden arm at a height of 2.2 m. There were two different transmitter and receiver location pairs used, shown in Fig. 2.10. The antenna configurations included a single centrally mounted quarter wavelength monopole, and two, three or four co-linear monopoles with outer monopoles separated by 32 mm. The signals r_{ml} received at each of the M antennas (indexed by m) at each of L positions (indexed by l) was used to estimate the elements of the signal plus noise correlation matrix using

$$\widehat{R}_{m_1 m_2} = \frac{1}{L} \sum_{l=1}^{L} r^*_{m_1 l} r_{m_2 l}. \tag{2.63}$$

There are additional issues associated with antenna mutual impedances, and the fact that only pairs of measurements could be made simultaneously. These are discussed at greater length in [102]. The noise power was estimated and subtracted to obtain an estimate of the signal correlation matrix $\mathbf{R}_{\text{snr}} = \mathbf{R}/\sigma_\eta^2 - \mathbf{I}$. The eigenvalues of this

[1]This work was led by Ole Nørklit.

matrix were then used to estimate the appropriate cdf function for the resulting signal using maximal ratio combining. These functions are the broken lines in Fig. 2.9. It can be seen that the two antenna system gives almost ideal performance. This is because the two antennas have low correlation. The four branch system gives an effective order of diversity of about 2.5.

It is apparent that the combining schemes at the mobile can give considerable immunity to fading. As has been discussed in chapter 1, channel prediction may be able to gain some of the effect of combining by allowing channels to be chosen which are not experiencing fading conditions. It is worth noting however, that even if perfect channel prediction information were available for all of a large number of channels from which one can be chosen, the best that could be hoped for in the way of performance enhancement would be the same as for a *selection* diversity scheme.

On the other hand, a prediction scheme requires no additional sensors, and so may be feasible in some situations where diversity is not. For a MIMO system, where multiple receive sensors are present in any case, prediction may still be of use in overcoming latency in the estimation of channel information at the transmitter.

2.5 Real Time Channel Characterisation

Real time characterisations of a channel have a different purpose to the stochastic characterisations discussed above. These allow a system to *adapt* to the particular *instantaneous* channel conditions being experienced. They form the basis of equalisers and Rake receivers. The parameters of a deterministic model of the channel are updated in real time and are used to optimise the performance of the receiver by removing as far as possible the effects of this channel.

Of interest in this thesis are models of the channel which have parameters which vary more slowly than does the channel itself. The term *syndetic*[1] is here defined as describing such models. Since the parameters of a syndetic model do not change rapidly, extrapolation of these parameters allows *predictions* of channel behaviour for some time into the future to be obtained in real time.

To be useful for equalisation and adaptation purposes, a model must provide not merely information on the ensemble of channels in a particular area, but instantaneous information on the instance of the channel being experienced by the signal in its passage from the transmitter to the receiver. For this purpose *deterministic* channel models are most useful. Although both stochastic and deterministic models are used in this thesis, the focus is on deterministic models.

[1] syndetic: (a) serving to unite or connect; connective, copulative., (b) pertaining to or designating a catalogue, index, etc., which uses cross-references to indicate links between entries. Oxford English Dictionary.

2.6 Deterministic Channel Models

If it is assumed that there are a finite number of discrete paths between the transmitter and receiver, then (2.12) may be used as the basis for a deterministic model.

2.6.1 Narrowband Far Field Sources

The description "far field" is used, because the assumption that the Doppler frequencies ϖ_n are only slowly varying functions of time implies that the angle θ_n is also only slowly varying, and hence that the nearest scatterer of each path is a considerable distance from the receiver. The term "far field" is taken from array-processing theory [77, p113] and implies that the waves incident on the receiving antenna are *plane waves*.

If the system has a narrow bandwidth, the delay of each path cannot be individually resolved. Equation 2.12 then simplifies to

$$h(\tau, t) = \sum_{n=1}^{N} \zeta_n e^{j\varpi_n t} \delta(\tau) \tag{2.64}$$

and so the received signal is

$$r(t) = \sum_{n=1}^{N} \zeta_n e^{j\varpi_n t} s(t). \tag{2.65}$$

In a real receiver of course, the received signal is filtered and demodulated (either synchronously or asynchronously) and the transmitted data is recovered. Provided the error rate is not too large, it is then simple to calculate $s(t)$ and hence $\sum_{n=1}^{N} \zeta_n e^{j\varpi_n t}$. Although demodulation is not a trivial task, it is not the focus of this thesis. For much of this thesis then, the assumption is made that the signal $s(t) = 1 \; \forall \; t$. The received signal $r(t)$ can thus be treated as instantaneous measurements of the complex channel gain.

The channel gain function can be sampled in time with period Δ_t to obtain the discrete function

$$r_m = \sum_{n=1}^{N} \zeta_n e^{j\varpi_n m} + \eta_m, \tag{2.66}$$

where the meaning of ϖ_n has now changed[1] from $\varpi_n = 2\pi v \sin\theta/\lambda$ measured in radians per second, to $\varpi_n = 2\pi\Delta_x \sin\theta/\lambda$ measured in radians per sample interval, where $\Delta_x = v\Delta_t/\lambda$ is the distance between samples measured in wavelengths.

A noise term η_m has been introduced. It will mostly be assumed that η_m is a complex normal process with $E\{\eta_m\} = 0$, $E\{\eta_{m_1}\eta^*_{m_2}\} = \sigma^2_\eta \delta_{m_1 m_2}$, and $E\{\eta_{m_1}\eta_{m_2}\} = 0$.

[1] This potential source of confusion is actually not as great as it may at first appear. ϖ is always multiplied by either a time or an index, and so the use of the symbol is very obvious from its context.

That is, it is assumed that the noise is white, or has been *whitened* after sampling (e.g., [154, pp60,621] or [116, p552]).

Equation 2.66 can be represented in matrix form as

$$\mathbf{r} = \mathbf{A}(\boldsymbol{\varpi})\boldsymbol{\zeta} + \boldsymbol{\eta}, \tag{2.67}$$

where

$$\begin{aligned}
\mathbf{r} &= [r_1, r_2, \ldots, r_M]^T, \\
\mathbf{A}(\boldsymbol{\varpi}) &= \left[\begin{array}{cccc} \mathbf{a}(\varpi_1) & \mathbf{a}(\varpi_2) & \ldots & \mathbf{a}(\varpi_N) \end{array} \right], \\
\mathbf{a}(\varpi) &= \left[e^{j\varpi 0}, e^{j\varpi 1}, \ldots, e^{j\varpi(M-1)} \right]^T, \\
\boldsymbol{\varpi} &= [\varpi_1, \varpi_2, \ldots, \varpi_N]^T, \\
\boldsymbol{\zeta} &= [\zeta_1, \zeta_2, \ldots, \zeta_N]^T, \quad \text{and} \\
\boldsymbol{\eta} &= [\eta_1, \eta_2, \ldots, \eta_M]^T.
\end{aligned} \tag{2.68}$$

2.6.2 Array Processing

At this point it is appropriate to introduce array processing theory. Equation 2.67 is a familiar form in array processing.

The signal at the mth element of an M element array is represented as the sum of signals arriving from N sources by the following sum [77]:

$$r_m(t) = \sum_{n=1}^{N} \zeta_n G_m(\boldsymbol{\theta}_n) s_n \left(t - \tau_m(\boldsymbol{\theta}_n)\right) + \eta_m(t), \qquad m = 1, 2, \ldots, M, \tag{2.69}$$

where

$\zeta_n = \varsigma_n e^{j\psi_n}$ is the complex attenuation of the n-th source,
$\boldsymbol{\theta}_n$ are the spatial parameters of the n-th source,
$G_m(\boldsymbol{\theta})$ is the gain of the mth sensor at $\boldsymbol{\theta}$,
$\tau_m(\boldsymbol{\theta})$ is the delay of a signal coming from $\boldsymbol{\theta}$ towards the m-th sensor,
$s_n(t)$ is the signal due to the n-th source, and
$\eta_m(t)$ is the noise at the m-th sensor.

The data is sampled over a measurement interval T, so that $-T/2 \leq t \leq T/2$. The Fourier transform of the data over the measurement interval T, defined as

$$R(\omega) \triangleq \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} r(t) e^{-j\omega t} dt, \tag{2.70}$$

can be applied to (2.69), resulting in

$$R_m(\omega) = \sum_{n=1}^{N} G_m(\boldsymbol{\theta}_n) e^{-j\omega\tau_m(\boldsymbol{\theta}_n)} \zeta_n S_n(\omega) + H_m(\omega), \tag{2.71}$$

where $R_m(\omega)$, $S_n(\omega)$, and $H_m(\omega)$, are the Fourier transform of $r_m(t)$, $s_n(t)$, and $\eta_m(t)$, respectively, normalised as in (2.70).

The measurements from each of the m elements can be arranged as a vector to yield

$$\begin{aligned} \mathbf{R}(\omega) &= \sum_{n=1}^{N} \mathbf{a}(\omega, \boldsymbol{\theta}_n) \zeta_n S_n(t) + \mathbf{H}(\omega), \\ \mathbf{a}(\omega, \boldsymbol{\theta}_n) &\triangleq \left[G_1(\boldsymbol{\theta}_n) e^{-j\omega\tau_1(\boldsymbol{\phi})}, G_2(\boldsymbol{\theta}_n) e^{-j\omega\tau_2(\boldsymbol{\theta}_n)}, \ldots, G_M(\boldsymbol{\theta}) e^{-j\omega\tau_M(\boldsymbol{\theta})} \right]^T. \end{aligned} \tag{2.72}$$

The vector $\mathbf{a}(\omega, \boldsymbol{\phi}_n)$ is called the *steering vector* of the array. The steering vectors can then be arranged as a matrix so that (2.72) can be presented as:

$$\begin{aligned} \mathbf{R}(\omega) &= \mathbf{A}(\omega, \boldsymbol{\Theta}) \mathbf{Z}(\omega) + \mathbf{H}(\omega) \\ M \times 1 \quad & M \times N \quad N \times 1 \quad M \times 1 \end{aligned} \tag{2.73}$$

where

$$\begin{aligned} \mathbf{A}(\omega, \boldsymbol{\Theta}) &\triangleq [\mathbf{a}(\omega, \boldsymbol{\theta}_1), \mathbf{a}(\omega, \boldsymbol{\theta}_2), \ldots, \mathbf{a}(\omega, \boldsymbol{\theta}_N)], \\ \mathbf{Z}(\omega) &\triangleq [\zeta_1 S_1(\omega), \zeta_2 S_2(\omega), \ldots, \zeta_N S_N(\omega)]^T, \quad \text{and} \\ \boldsymbol{\Theta} &\triangleq [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_N]. \end{aligned}$$

This equation may be used as the basis for several different algorithms, such as

1. Extracting the data, $s_n(t)$, for $n \in \{1, 2, \ldots, N\}$.
2. Extracting the data, $s_n(t)$, in the presence of a other interfering signals $s_{n'}(t)$ for $n' \neq n$.
3. Finding the number of sources, N.
4. Finding the spatial parameters, $\boldsymbol{\Theta}$.

In this thesis it is the last two problems which are of interest. It is assumed (initially at least) that all of the sources contain the desired signal, and are simply scatterers in different locations of the same *desired* signal; that is, there is no interference.

For narrowband signals, the expressions in (2.73) can be considered to be independent of frequency. If the sources are in the far field, the spatial parameter vector $\boldsymbol{\Theta}$ is simply the vector of angles $\boldsymbol{\theta}$, and if the array elements are co-linear and regularly spaced, these can be expressed in terms of Doppler frequencies $\boldsymbol{\varpi} = (2\pi/\lambda) \sin \boldsymbol{\theta}$. Hence

the equation may be considered as the equivalent of (2.67). Thus the measurements of the channel made by a mobile receiver can be considered to be equivalent to the measurements received by the elements of a stationary array. In other words, the moving receiver forms a *synthetic array*.

2.6.3 Linear Prediction

It is apparent from (2.66) for example, that the discrete point sources model is equivalent to the sum of complex sinusoids in noise. Techniques for estimating the parameters of such a model are well known and are discussed in the next chapter. It is also equivalent to what is known as a *predictable process* (where prediction is a linear function of previous data values) or an autoregressive (AR) process. This equivalence follows from Wold's Decomposition theorem [105, 106]. A brief rationale for one of the elements of Wold's decomposition theorem follows.

For a wide-sense stationary process $r[m]$, the prediction error may be defined as

$$W = \min_{a_k} E\left\{\left|r[n] - \sum_{k=1}^{\infty} a_k r[n-k]\right|^2\right\}, \tag{2.74}$$

and a predictable process is defined as one for which $W = 0$. If the process consists only of sinusoids, then its spectrum may be expressed as the sum of lines

$$R(e^{j\omega}) = \sum_i \alpha_i \delta(\omega - \omega_i). \tag{2.75}$$

The polynomial F may be formed as

$$F(z) = \prod_i (1 - z^{-1} e^{j\omega_i}) \equiv 1 - \sum_k a_k z^{-k}, \tag{2.76}$$

and the terms R and F may be multiplied to obtain

$$R(e^{j\omega})F(e^{j\omega}) = \sum_i \alpha_i F(e^{j\omega_i})\delta(\omega - \omega_i) = 0. \tag{2.77}$$

It can be seen that the resulting mean square error W is zero. Hence there is a close relationship between the far field point sources model, and standard time-series AR modelling techniques. The only real difference is that if the process consists of sinusoids of constant magnitude (i.e., not growing or decaying) then the poles of the prediction filter are constrained to lie on the unit circle.

If a scattering distribution is truly point sources (sources having infinitesimal physical extent), so that the Doppler spectrum of the fading process can be expressed as the sum of lines, then the error in linear prediction of the process is zero! That does not necessarily imply that the prediction coefficients can be perfectly estimated of course, since

that will depend on the quantity and quality (SNR) of the data available for estimation of the coefficients. That the sources may not be points is examined in Section 4.5.

2.6.4 Narrowband Near Field Sources

If a mobile receiver is moving near to a point signal source, the wave fronts incident on the receiving antenna are not planar, but spherical, and the angle θ_n may vary quite rapidly. In this case, a model for which the parameters do not vary so rapidly is a near field model, which includes a parameter relating to the distance to the source.

The spatial parameter vector $\mathbf{\Theta}$ in this case may be $[d_1, \theta_1, d_2, \theta_2, \dots, d_N, \theta_N]^T$, where d_n is the distance of the mth source point from some arbitrary reference location, or $[x_1, y_1, x_2, y_2, \dots, x_N, y_N]^T$, where x_n is the distance along and y_n the distance from some axis of the array.

Just as the far field point source model is equivalent to prediction based on *linear* combinations of past measurements, so the near field point source model is equivalent to prediction based on *non-linear* combinations of past measurement. If the non-linear function governing the evolution of the channel is smooth, then a truncated Taylor series expansion may provide an accurate model. Confining the expansion to linear and quadratic terms of past values results in the quadratic Volterra series [134]. Application of this method to real time channel prediction is investigated in [41]. Another general technique of parametric nonlinear modelling investigated in [41] is Multivariate Adaptive Regression Splines (MARS) [52, 91]. Neither of these techniques appear to offer reliable prediction performance. One of the difficulties is again the large number of parameters required.

2.6.5 Wide-band Sources

There are several possible approaches to wide-band source modelling. These are based on either (2.12) or (2.73). In the first approach the entire scenario of scatterers is modelled as one set. The parameters which must be estimated are the number of paths N, and N triples $(\zeta_n, \varpi_n, \tau_n)$.

In the second approach, the wide-band channel is treated as some integral number W of narrowband channels. The parameters of this model are WN pairs (ζ_n, ϖ_n). This approach does not make use of any the spatial "insights" available from the *combination* of the delays, but is as simple to implement as any narrowband algorithm.

An alternative to modelling the channel behaviour at different frequencies ω_n, is to model the channel behaviour at different delays τ_n. This is the approach used in [41]. It is apparent from (2.17) that the parameters $h(\tau, x)$ and $H(\omega, x)$ have a Fourier transform relationship, and so are in most respects equivalent.

Obviously both of these approaches can be extended to the near field case by changing the spatial parameters involved. In general the distance to a source point d_n will

not have a close relationship to its delay τ_n because d_n is the distance of the receiving antenna from the last reflection/diffraction point, whereas τ_n is the total path delay time of the signal since leaving the transmit antenna, which often will involve multiple reflections (although any particular distance d_n must be less than or equal to $c\tau_n$).

Characterisation of wide-band channels is a generalisation (although not a trivial one) of narrowband channels. The focus of this thesis is on an initial study of the feasibility and likely impact of prediction and thus focuses on the narrowband models, which are more simple to analyse. In Chapter 5 however, the probability of error for wide-band communications systems is investigated in detail.

2.7 Model Limitations

Even before use is made of the overall channel model proposed, there are several obvious difficulties, which are briefly discussed here.

These are

1. The sources are often not point sources;
2. The sources are often more in number than it is feasible to estimate the parameters of;
3. If the sources are in the near field or the sources move or the receiving mobile turns or accelerates, the model parameters change significantly;
4. Whenever there is a large scale change in the scattering environment, such as the mobile moving behind a large building, the channel parameters may change drastically and abruptly.

2.7.1 Source Extent

If an object is conducting, large and smooth, it forms a reflector of the signal. If the source is distant and small, the mirror image of the signal may then be considered as a point source. If the size of the object is smaller than the first *Fresnel zone* for a particular path, a significant amount of diffraction occurs. This first Fresnel zone (see Section 2.7.1) is contained by an ellipsoid defined by the locus of points having a difference between the direct and reflected paths of $\lambda/2$. At typical frequencies used in mobile communications (900MHz, 1800MHz, 2400MHz for example) many objects such as cars and even buildings (in an outdoor context) are often too small to act as reflectors [41, p30]. On the other hand, for the objects to be considered as points they must be either distant or small so that the radiation intercepted by the receiver is effectively isotropic. In a typical scenario, objects frequently do not satisfy this criterion either.

Material	Reflection Coefficient
Floor	0.2-0.3
Ceiling	0-0.1
External wall	0.025
Internal wall	0.15-0.5
Soft partition	0.25
Elevator wall	0.35

Table 2.1: Reflection coefficients for electromagnetic waves.

2.7.2 Number of Paths

As was discussed in Section 2.3, it appears that despite some assertions to the contrary, there are often a large number of paths between the transmitter and receiver, and the distribution of the powers of these paths is such that there is no small subset of the paths which delivers most of the energy. For the indoor situation a rough analogy may be made with room acoustics; a 900MHz radio carrier has the same wavelength as a 1KHz tone. Acoustic impulse responses can be simulated using ray tracing techniques, but only when very high order reflections are included. Since the number of paths rises exponentially with the order of reflections considered, it is clear that in the audio situation at least, a realistic approximation must model a large number of paths. In many cases the data which is available is insufficient to allow estimation of the large number of parameters corresponding to these paths.

The reflection coefficients used in [88] for modelling of electromagnetic behaviour of rooms are shown in Table 2.1. These values were obtained through a process of empirical determination of likely values, and refinement of those values with reference to the results of the measurement experiments. They are assumed to be an "average" over all polarisations and angles of incidence. Some of the coefficients used in acoustic design [16] are shown in Table 2.2. These have been converted from energy absorption coefficients to amplitude reflection coefficients, assuming that the energy not absorbed is reflected. Note that this assumption will not always be valid, since transmitted energy may not be neglected if the material is not on the perimeter of a room, or if there is a reflective material behind it. Even allowing for this roughness in the analogy, it is clear from a comparison of these tables that electromagnetic multipath in an office environment corresponds in the analogy to acoustic reverberation in a heavily carpeted and curtained room. Even in such a room however, there is still a large amount of multipath present.

2.7.3 Near field sources

If sources near the trajectory of a mobile receiver are modelled using a far field model, it is clear that the parameters of that model will vary quite rapidly — thus violating

Material	Reflection Coefficient
Brick Wall, unpainted	0.98
Plaster on wood studs	0.98
Wood panelling	0.89
Carpet, with underlay	0.81

Table 2.2: Reflection coefficients for acoustic waves at 500Hz.

the aim of finding a syndetic model. Thus the near field model appears to have an advantage. It is however, a less concise model, requiring four real parameters for every source. Also the non-linear nature of this model causes estimation of these parameters to be much more difficult (see Section 3.2.2).

In the case of near field sources or a large number of sources of any type, there is a conflict between the requirements that the model be both syndetic and parsimonious. The model must be syndetic (having slowly varying parameters) so that the parameters may be extrapolated to allow prediction. A model which included information on the position, velocity size, orientation and nature of every scatterer in the vicinity would have very slowly varying parameters. It would however not be parsimonious — there would of course be far more parameters than could be estimated from the data, even if the maximum likelihood solution were instantly available.

2.7.4 Changes in the Parameters

If the receiver moves into or out of the "shadow" of a building, the scattering scenario can change significantly and rapidly. Clearly such catastrophic (for the predictor) events form "horizons" beyond which prediction cannot be achieved. The percentage of the time occupied by such events would make a worthwhile study.

2.8 Summary and Contributions

The underlying model of discrete point sources used in the majority of this thesis has been presented. The stochastic channel models which are used for "off-line" channel characterisations have been reviewed, and the need for a different type of channel model discussed. Spatial correlation and diversity have been discussed.

We itemise some specific contributions made in this chapter:

i. A new method has been presented for calculating the correlation of a narrowband channel at two spatially separated points for general distributions of scatterers.

ii. It has been explained that reliable prediction requires real time characterisation, rather than simply ensemble characterisation.

iii. The fact that many properties of a channel may be reliably simulated with a small number of discrete sources has led many researchers to believe that in a typical scenario there *are* only a small number. The fallacy of this reasoning has been explained.

iv. The fact has been discovered that the property of a channel that it can be well predicted (even using linear methods) is not well indicated by spatial correlation.

v. New measurements have been presented indicating the degree of effective diversity available from multiple receiving antennas on a handset.

Chapter 3

Prediction Algorithms

ANY mobile radio receiver moving through a series of points in space in a multipath environment will receive a different signal at each of these points. The fundamental concept on which the possibility of prediction is based is the use of these points as the elements of a *synthetic array*. The samples along a known spatial trajectory can be used to estimate the parameters of the channel model. The channel model can then be used to extrapolate beyond the region of the measurements (see Fig. 3.1).

The channel models used for the majority of this thesis have been discussed in detail in Chapter 2. The algorithms by which the parameters of the channel may be estimated and then extrapolated are presented in this chapter. Some other models are also introduced which are used by way of comparison in Chapter 4.

3.1 Deterministic Point Sources Channel Model

In this section the deterministic channel models introduced in Chapter 2 are further developed by deriving the likelihood function. The derivative of the likelihood function is also determined, as this is used for optimisation of a solution to the likelihood function. Other methods which may be used for estimating the parameters of this model are also developed.

3.1.1 Likelihood Function

For solving any estimation problem, the beginning point is deriving the likelihood function for the parameters. This is derived in this section and then simplified. Equation 2.73 is used as a basis.

In some situations the signal is known, but usually $s_n(t)$ is unknown. It can be modelled as a deterministic unknown or as a stochastic process. The noise $\boldsymbol{\eta}_m(t)$ is assumed to be a zero mean, *complex* normal process. If the signals are deterministic, the term $\mathbf{AZ}$ in (2.73) can be regarded as the mean of the complex random process

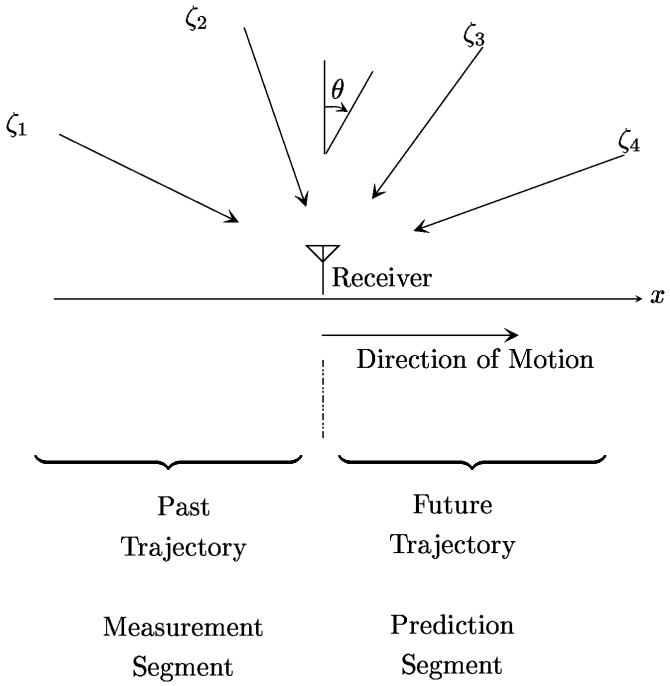


Figure 3.1: The prediction concept

$\mathbf{Y}(\omega)$ where

$$\mathbf{Y}(\omega_l) \sim \mathcal{NC}\left(\mathbf{A}(\omega_l, \mathbf{\Theta})\mathbf{Z}\left(\omega_l\right), \mathbf{C}_{\boldsymbol{\eta}}(\omega_l)\right), \tag{3.1}$$

and $\mathbf{C}_{\boldsymbol{\eta}}(\omega)$ is the covariance matrix of the noise $\boldsymbol{\eta}$ at frequency ω.

If the noise is uncorrelated in frequency, the conditional probability density functions corresponding to different frequencies can be multiplied. The likelihood function of $\mathbf{Y}$ is then given by

$$p(\mathbf{Y}) = \prod_l \left[\frac{1}{\pi^M \left|\mathbf{C}_{\boldsymbol{\eta}}(\omega_l)\right|} e^{-(\mathbf{Y}(\omega_l)-\mathbf{A}(\omega_l,\mathbf{\Theta})\mathbf{Z}(\omega_l))^H \mathbf{C}_{\boldsymbol{\eta}}(\omega_l)^{-1}(\mathbf{Y}(\omega_l)-\mathbf{A}(\omega_l,\mathbf{\Theta})\mathbf{Z}(\omega_l))}\right]. \tag{3.2}$$

This can be maximised by minimising the negative log likelihood, or equivalently, the expression

$$L(\mathbf{\Theta}) = \sum_l \Big(\mathbf{Y}(\omega_l) - \mathbf{A}(\omega_l, \mathbf{\Theta})\mathbf{Z}(\omega_l)\Big)^H \mathbf{C}_{\boldsymbol{\eta}}(\omega_l)^{-1}\Big(\mathbf{Y}(\omega_l) - \mathbf{A}(\omega_l, \mathbf{\Theta})\mathbf{Z}(\omega_l)\Big). \tag{3.3}$$

The noise at each sensor is assumed to be uncorrelated. Thus the spatial noise covariance matrix $\mathbf{C}_{\boldsymbol{\eta}}(\omega)$ is a diagonal matrix. When the noise at each sensor has equal variance, the noise covariance matrix is simply a scaled identity matrix $\sigma_\eta^2\mathbf{I}$. In reality these are reasonable assumptions if the noise is primarily thermal noise in the receiver, and the receivers are identical, and do not have nonlinearities such as Automatic Gain

Control (AGC).

With these assumptions, and treating all of the frequencies ω_l separately, the value of $\mathbf{Z}(\omega)$ which minimises $L(\mathbf{\Theta})$ is [80, 97]

$$\begin{aligned}\mathbf{Z}(\omega) &= \left(\mathbf{A}^H(\omega,\mathbf{\Theta})\,\mathbf{A}(\omega,\mathbf{\Theta})\right)^{-1}\mathbf{A}^H(\omega,\mathbf{\Theta})\,\mathbf{Y}(\omega)\\ &\quad\text{or}\;\; \mathbf{A}^+(\omega,\mathbf{\Theta})\,\mathbf{Y}(\omega)\,.\end{aligned} \tag{3.4}$$

where $\mathbf{A}^+$ denotes the Moore-Penrose generalised inverse of the matrix $\mathbf{A}$. The first expression is valid if $\mathbf{A}^H\mathbf{A}$ is full rank which will be the case whenever $M \geq N$ and the parameters $\mathbf{\Theta}$ are unique.

These values of $\mathbf{Z}(\omega)$ are substituted into equation (3.3), and the assumption made that $\mathbf{C}_{\boldsymbol{\eta}}(\omega_l)^{-1} = \mathbf{I}/\sigma_l^2$. A *projection matrix* is defined as

$$\begin{aligned}\mathbf{P}(\omega_l,\mathbf{\Theta}) &= \mathbf{A}(\omega_l,\mathbf{\Theta})\left(\mathbf{A}^H(\omega_l,\mathbf{\Theta})\,\mathbf{A}(\omega_l,\mathbf{\Theta})\right)^{-1}\mathbf{A}^H(\omega_l,\mathbf{\Theta})\\ &\quad\text{or}\;\; \mathbf{A}(\omega_l,\mathbf{\Theta})\,\mathbf{A}^+(\omega_l,\mathbf{\Theta})\,,\end{aligned} \tag{3.5}$$

so that

$$\begin{aligned}L(\mathbf{\Theta}) &= \sum_l \frac{1}{\sigma_l^2}\Big(\mathbf{Y}(\omega_l) - \mathbf{P}(\omega_l,\mathbf{\Theta})\mathbf{Y}(\omega_l)\Big)^H\Big(\mathbf{Y}(\omega_l) - \mathbf{P}(\omega_l,\mathbf{\Theta})\mathbf{Y}(\omega_l)\Big)\\ &= \sum_l \frac{1}{\sigma_l^2}\left(\mathbf{Y}^H(\omega_l)\mathbf{Y}(\omega_l) - \mathbf{Y}^H(\omega_l)\mathbf{P}(\omega_l,\mathbf{\Theta})\,\mathbf{Y}(\omega_l)\right).\end{aligned} \tag{3.6}$$

Use has been made of the fact that the projection matrix $\mathbf{P}$ is idempotent. Since $\mathbf{Y}^H\mathbf{Y}$ is independent of $\mathbf{\Theta}$, the likelihood of a solution is maximised by maximising

$$L_2(\mathbf{\Theta}) = \sum_l \frac{1}{\sigma_l^2}\mathbf{Y}^H(\omega_l)\mathbf{P}(\omega_l,\mathbf{\Theta})\,\mathbf{Y}(\omega_l). \tag{3.7}$$

This expression can be further simplified in certain situations. For instance, if the system is of narrow bandwidth so that the matrix $\mathbf{P}(\omega_l,\mathbf{\Theta})$, and the noise power σ_l^2 (assumed equal to unity without loss of generality) can be considered to be independent of frequency, then

$$\begin{aligned}L_2(\mathbf{\Theta}) &= \sum_l \mathbf{Y}^H(\omega_l)\mathbf{P}(\mathbf{\Theta})\,\mathbf{Y}(\omega_l)\\ &= \mathrm{tr}\Big(\sum_l \mathbf{Y}^H(\omega_l)\mathbf{P}(\mathbf{\Theta})\,\mathbf{Y}(\omega_l)\Big)\\ &= \sum_l \mathrm{tr}\Big(\mathbf{Y}^H(\omega_l)\mathbf{P}(\mathbf{\Theta})\,\mathbf{Y}(\omega_l)\Big)\\ &= \sum_l \mathrm{tr}\Big(\mathbf{P}(\mathbf{\Theta})\,\mathbf{Y}(\omega_l)\mathbf{Y}^H(\omega_l)\Big)\\ &= L\;\mathrm{tr}\left(\mathbf{P}(\mathbf{\Theta})\,\widehat{\mathbf{R}}_y\right),\end{aligned} \tag{3.8}$$

where $\widehat{\mathbf{R}}_{\mathbf{y}} = \frac{1}{L}\sum_l \mathbf{Y}(\omega_l)\mathbf{Y}^H(\omega_l)$ is an estimate of the covariance of the measurements $\mathbf{Y}(\omega_l)$.

If there is only one source, $\mathbf{A}$ only has one column, $\mathbf{a}$, so $(\mathbf{A}^H\mathbf{A})^{-1}$ is just a scalar, and the expression can be further simplified to

$$L_2(\mathbf{\Theta}) = \frac{|\mathbf{a}(\mathbf{\Theta})\mathbf{y}|^2}{\sum_{m=1}^{M} |\mathbf{a}_m(\mathbf{\Theta})|^2}. \tag{3.9}$$

3.1.2 First Derivative

The first derivative of the likelihood cost function is useful when using gradient methods for finding the optimal solution.

First the derivative of

$$\begin{aligned} L(\theta) &= \mathbf{y}^H\mathbf{A}(\theta)\left(\mathbf{A}(\theta)\mathbf{A}^H(\theta)\right)^{-1}\mathbf{A}^H(\theta)\mathbf{y} \\ &= \mathbf{y}^H\mathbf{P}(\theta)\mathbf{y} \end{aligned} \tag{3.10}$$

is determined, where $\mathbf{A}$ is a simple function of a scalar θ.

Differentiating both sides of $\mathbf{C}^{-1}\mathbf{C} = \mathbf{I}$ with respect to θ gives

$$\frac{\partial \mathbf{C}^{-1}}{\partial\theta} = -\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial\theta}\mathbf{C}^{-1}, \tag{3.11}$$

and this result can be applied to (3.10) to obtain

$$\frac{\partial \mathbf{P}}{\partial\theta} = \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)\left(\frac{\partial \mathbf{A}}{\partial\theta}\right)^H(\mathbf{I}-\mathbf{P}) \quad + \quad (\mathbf{I}-\mathbf{P})\left(\frac{\partial \mathbf{A}}{\partial\theta}\right)\left(\mathbf{A}^H\mathbf{A}\right)\mathbf{A}^H, \tag{3.12}$$

and so

$$\frac{\partial \mathbf{y}^H\mathbf{P}\mathbf{y}}{\partial\theta} = 2\mathrm{Re}\left(\mathbf{y}^H\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)\left(\frac{\partial \mathbf{A}}{\partial\theta}\right)^H(\mathbf{I}-\mathbf{P})\,\mathbf{y}\right). \tag{3.13}$$

Consider now the case of a vector of parameters. Since the steering matrix $\mathbf{A}(\omega,\mathbf{\Theta}) = [\mathbf{a}(\omega,\boldsymbol{\theta}_1),\mathbf{a}(\omega,\boldsymbol{\theta}_2),\ldots,\mathbf{a}(\omega,\boldsymbol{\theta}_N)]$, each of the columns of $\mathbf{A}(\omega,\mathbf{\Theta})$ depend on only one of $\boldsymbol{\theta}_n$, $n = 1,2,\ldots N$. It is assumed now that $\boldsymbol{\theta}_n$ is a scalar (as in the far field situation, where each source has only one spatial parameter), although the result can be readily extended if this is not the case. Thus

$$\frac{\partial \mathbf{A}}{\partial\boldsymbol{\theta}_n} = \frac{\partial \mathbf{a}_n}{\partial\boldsymbol{\theta}_n}\mathbf{e}_n^T, \tag{3.14}$$

where $\mathbf{e}_n^T = (0, 0, \ldots, 1, \ldots, 0)$ contains a 1 only in the n-th position, and so

$$\left(\frac{\partial \mathbf{A}}{\partial \boldsymbol{\theta}_n}\right)^H = \mathbf{e}_n \left(\frac{\partial \mathbf{a}_n}{\partial \boldsymbol{\theta}_n}\right)^H. \tag{3.15}$$

Inserting this expression into (3.13) gives

$$\frac{\partial \left(\mathbf{y}^H \mathbf{P} \mathbf{y}\right)}{\partial \boldsymbol{\theta}_n} = 2\,\mathrm{Re}\left(\underbrace{\overbrace{\mathbf{y}^H \mathbf{A}}^{1\times N} \overbrace{\left(\mathbf{A}^H \mathbf{A}\right)^{-1}}^{N\times 1} \overbrace{\mathbf{e}_n}^{1\times M} \overbrace{\left(\frac{\partial \mathbf{a}_n}{\partial \boldsymbol{\theta}_n}\right)^H}^{M\times 1} (\mathbf{I} - \mathbf{P})\,\mathbf{y}}_{1\times 1 \qquad\qquad 1\times 1} \right) \tag{3.16}$$

This expression is applied to each of $\boldsymbol{\theta}_n$, and the results arranged as a vector to obtain

$$\begin{aligned} \frac{\partial \left(\mathbf{y}^H \mathbf{P} \mathbf{y}\right)}{\partial \boldsymbol{\Theta}} &= 2\,\mathrm{Re}\begin{bmatrix} \mathbf{y}^H \mathbf{A} \left(\mathbf{A}^H \mathbf{A}\right)^{-1} \mathbf{e}_1 \left(\frac{\partial \mathbf{a}_1}{\partial \boldsymbol{\theta}_1}\right)^H (\mathbf{I} - \mathbf{P})\,\mathbf{y} \\ \mathbf{y}^H \mathbf{A} \left(\mathbf{A}^H \mathbf{A}\right)^{-1} \mathbf{e}_2 \left(\frac{\partial \mathbf{a}_2}{\partial \boldsymbol{\theta}_2}\right)^H (\mathbf{I} - \mathbf{P})\,\mathbf{y} \\ \vdots \\ \mathbf{y}^H \mathbf{A} \left(\mathbf{A}^H \mathbf{A}\right)^{-1} \mathbf{e}_N \left(\frac{\partial \mathbf{a}_L}{\partial \boldsymbol{\theta}_N}\right)^H (\mathbf{I} - \mathbf{P})\,\mathbf{y} \end{bmatrix} \\ &= 2\,\mathrm{Re}\left(\left(\mathbf{y}^H \mathbf{A} \left(\mathbf{A}^H \mathbf{A}\right)^{-1}\right) \odot \left(\mathbf{D}^H (\mathbf{I} - \mathbf{P})\,\mathbf{y}\right) \right), \end{aligned} \tag{3.17}$$

where

$$\mathbf{D} = \left[\frac{\partial \mathbf{a}(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}, \frac{\partial \mathbf{a}(\boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2}, \ldots, \frac{\partial \mathbf{a}(\boldsymbol{\theta}_N)}{\partial \boldsymbol{\theta}_N} \right], \tag{3.18}$$

and $\mathbf{A} \odot \mathbf{B}$ is simply the element by element, or *Hadamard* product of two matrices or vectors.

3.1.3 Far field Sources

For far field sources, it is assumed that the gain of each sensor is the same for all sources, and the only feature which distinguishes source locations is the relative delay to each of the sensors of the array. This is simply a function of the source *angle* and the sensor location.

The "gain" of each element of the steering vector (refer to equation 2.72) is

$$G_n(\boldsymbol{\theta}_n) e^{-j\omega_l \tau_m(\boldsymbol{\theta}_n)} = e^{+j2\pi \frac{D_{mn}}{\lambda_l}}, \tag{3.19}$$

where λ_l is the wavelength of the signal at angular frequency ω_l, and D_{mn} is the distance in the direction travelled by the plane waves between the location of sensor m and the array reference point (refer to Fig. 3.2). D_{mn} can be calculated as the distance from the origin of the intersection of the lines $y = \cot\theta x$ and $y = y_m - \tan\theta(x - x_m)$, given by

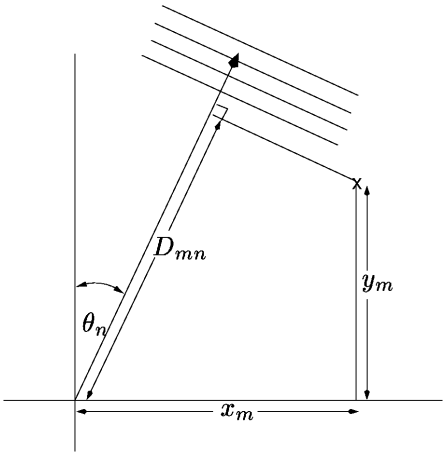


Figure 3.2: Geometry for far field sources.

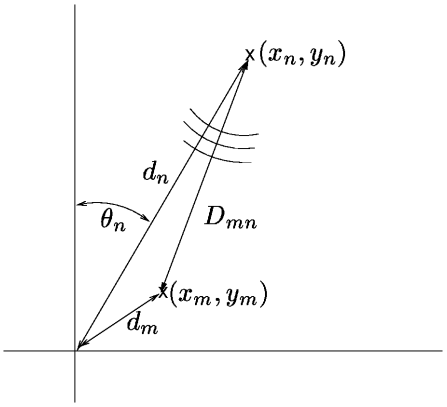


Figure 3.3: Geometry for near field sources.

$D_{mn} = x_m \sin\theta_n + y_m \cos\theta_n$. The angle θ_n of the source is taken from some arbitrary "axis" of the array. For the synthetic array used in this thesis, the angle is from a perpendicular to the direction of travel. The derivative (for use in equation 3.18) is

$$\frac{\partial \mathbf{A}_{mn}}{\partial \theta_n} = j\frac{2\pi}{\lambda_l}\left(x_m \cos\theta_n - y_m \sin\theta_n\right)\mathbf{A}_{mn}. \tag{3.20}$$

3.1.4 Near Field Sources

For near field sources, the delay between a source point and a sensor point is dependent on the absolute position of the source point. In addition, the amplitude of the signal from a particular source is also dependent on the distance.

Thus each element of the steering vector (refer to equation 2.72) is

$$G_n(\boldsymbol{\theta}_n)e^{-j\omega_l \tau_m(\boldsymbol{\theta}_n)} = \frac{1}{D_{mn}} e^{-j2\pi\frac{D_{mn}}{\lambda_l}}, \tag{3.21}$$

where D_{mn} (referring to Fig. 3.3) is given by

$$\begin{aligned} D_{mn}^2 &= (x_m - x_n)^2 + (y_m - y_n)^2 \\ &= d_n^2 + d_m^2 - 2d_n(x_m \sin\theta_n + y_m \cos\theta_m). \end{aligned} \tag{3.22}$$

The derivatives (for use in equation 3.18) are

$$\begin{aligned} \frac{\partial \mathbf{A}_{mn}}{\partial x_n} &= \mathbf{A}_{mn}\left(\frac{x_m - x_n}{D_{mn}}\right)\left(j\frac{2\pi}{\lambda_l} + \frac{1}{D_{mn}}\right) \\ \frac{\partial \mathbf{A}_{mn}}{\partial y_n} &= \mathbf{A}_{mn}\left(\frac{y_m - y_n}{D_{mn}}\right)\left(j\frac{2\pi}{\lambda_l} + \frac{1}{D_{mn}}\right). \end{aligned} \tag{3.23}$$

Note that the standard half wavelength sensor spacing rule, which guarantees that no aliasing will occur in the operation of a far field array, is not sufficient to prevent aliasing in the near field [3] (this is discussed in more detail in Section 3.2.2).

3.1.5 Parameter Estimation

In principle, (3.8) can be used to find the maximum likelihood estimate of the spatial parameters $\mathbf{\Theta}$, and then (3.4) provides a simple linear estimate of the signal parameters $\mathbf{Z}$. If there are more than say two spatial parameters (two far field sources or one near field source) the likelihood function has too many dimensions for a global search to be feasible. The derivative of the likelihood function (from Section 3.1.2) is useful for optimising a solution only if the starting point for the optimisation is close to the global maximum because in many situations the likelihood function has a large number of local maxima. Optimisation using the derivative has been used in this thesis, but only after a good initial estimate has been obtained using some other method. The primary means of obtaining this initial estimate was some form of *subspace spectral estimation*.

3.2 Subspace Spectral Estimation

Equation 2.66, applicable for sources in the far field of the synthetic array, can be seen to be simply an expression for the sum of complex sinusoids with additive white noise. Provided the SNR is sufficient, standard spectral estimation techniques can thus be applied to estimate the signal Doppler frequencies and amplitudes of each of the sources. Once these parameters have been estimated the signal prediction is performed by extrapolation of the sinusoids.

All of the subspace spectral estimation methods commence with an estimate $\widehat{\mathbf{R}}_{\mathbf{y}}$ of the signal correlation matrix introduced in (3.8), and eigendecomposition of this matrix or some sub-matrix of it.

In conventional array processing, it is relatively simple to obtain a good estimate

of this correlation matrix by taking the time average of the correlation between each pair of array elements (averaging over so called *snapshots*):

$$\widehat{\mathbf{R}}_{\mathbf{y}} = \frac{1}{L}\sum_{k=1}^{L} \mathbf{y}_l \mathbf{y}_l^H. \tag{3.24}$$

For a synthetic array, the receiver is normally only at each location once, so this option is not available. Some averaging can be obtained using different frequencies, but as (3.8) is only valid under the condition that the frequencies are all close, the samples will not be independent, and so a good estimate of $\mathbf{R}_{\mathbf{y}}$ will not be obtained. For far field sources, the correlation between the signals at two array elements is

$$\begin{aligned} E\{y_{m_1} y_{m_2}^*\} &= E\left\{ \left(\sum_{n=1}^{N} \zeta_n e^{jm_1\varpi_n} \right) \left(\sum_{n=1}^{N} \zeta_n e^{jm_2\varpi_n} \right)^* \right\} \\ &= E\left\{ \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \zeta_{n_1} \zeta_{n_2}^* e^{j(m_1\varpi_{n_1} - m_2\varpi_{n_2})} \right\} \\ &= E\left\{ \sum_{n=1}^{N} |\zeta_n|^2 \, e^{j\varpi_n(m_1 - m_2)} \right\}, \end{aligned} \tag{3.25}$$

where the last line follows from uncorrelatedness of the signal complex amplitudes. This expression is clearly dependent on m only through the difference $m_1 - m_2$. Thus a good estimate $\widehat{\mathbf{R}}_{\mathbf{y}}$ of the correlation matrix can be obtained by averaging the correlation over different pairs of equally spaced sensors. This leads to what is known as the "covariance method" of estimation the correlation matrix [154, p538]. Unfortunately this does not hold for near field sources (this is discussed in more detail in Section 3.2.2 where the MUSIC method is applied to near field sources).

It can be seen from (3.25) that $E\{y_{m_2} y_{m_1}^*\} = (E\{y_{m_1} y_{m_2}^*\})^*$, and hence that further averaging of the correlation between terms m_1 and m_2 can be obtained by including the complex conjugate of correlation between terms m_2 and m_1. This leads to the "modified covariance" estimate [154, p543]. This increased averaging was found to be useful in reducing the variance of the parameter estimates.

If the signal samples (array elements) are much more closely spaced than the Nyquist sampling rate, a smaller correlation matrix can be calculated. If the distance between array elements is Δ_x (and the moving receiver itself is not part of any multiple-reflection paths) the spatial frequencies are band limited to the range $(-2\pi\Delta_x/\lambda, 2\pi\Delta_x/\lambda]$. The correlation matrix may be constructed based on overlapping sets of samples L times further apart than the original data, where $L < \lambda/(2\Delta_x)$. The $(M - PL + L) \times P$ data

matrix is

$$\mathbf{X} = \begin{bmatrix} x_{(P-1)L+1} & \cdots & x_{L+1} & x_1 \\ x_{(P-1)L+2} & \cdots & x_{L+2} & x_2 \\ \vdots & & \vdots & \vdots \\ x_M & \cdots & & x_{M-(P-1)L} \end{bmatrix}. \tag{3.26}$$

The $P \times P$ correlation matrix estimate is then formed as $\widehat{\mathbf{R}}_{\mathbf{y}1} = (M - PL + L)^{-1}\mathbf{X}^H\mathbf{X}$, or for the modified covariance method, using both forward and backward terms, $\widehat{\mathbf{R}}_{\mathbf{y}} = (\widehat{\mathbf{R}}_{\mathbf{y}1} + \mathbf{K}\widehat{\mathbf{R}}^*_{\mathbf{y}1}\mathbf{K})/2$ where $\mathbf{K}$ is a $P \times P$ matrix with elements $[\mathbf{K}]_{ij} = \delta_{i,P+1-j}$. An integer L larger than unity was found to give some improvement in estimation accuracy, and a considerable decrease in computation time. Using this approach Δ_x is redefined in subsequent calculations as L times the distance between adjacent samples. This decreases the correlation matrix size, while still using all the data to reduce the effect of noise. The improvement in prediction accuracy associated with a lower model order, while making use of higher rate data would appear to be functionally equivalent to that observed by Eyceöz et al. [46] for what is called there "adaptive tracking". The model used there is for an autoregressive (AR) model, implemented in a recursive fashion. The improved averaging resulting from this approach removes the need for the low pass filtering, and reduces the clustering of frequencies around zero reported in [74].

The following few paragraphs are adapted from [154]. From (2.67) the estimated correlation matrix of the received signals $\mathbf{r}$ can written as

$$\begin{aligned} \mathbf{R}_{\mathbf{r}} &= E\left\{\mathbf{r}\mathbf{r}^H\right\} \\ &= E\left\{\mathbf{A}\boldsymbol{\zeta}\boldsymbol{\zeta}^H\mathbf{A}^H\right\} + E\left\{\boldsymbol{\eta}\boldsymbol{\eta}^H\right\} \\ &= \mathbf{A}\mathbf{R}_{\boldsymbol{\zeta}}\mathbf{A}^H + \sigma_\eta^2\mathbf{I} \\ &= \sum_{n=1}^{N}\left(P_n\mathbf{a}_n\mathbf{a}_n^H\right) + \sigma_\eta^2\mathbf{I}, \end{aligned} \tag{3.27}$$

where $P_n = E\{|\zeta_n|^2\}$, $\mathbf{R}_{\boldsymbol{\zeta}} = \mathrm{diag}(P_1, \ldots, P_N)$ and $\mathbf{a}_n$ is the nth column of $\mathbf{A}$.

Now consider the eigenvectors of the $P \times P$ matrix $\mathbf{R}_{\mathbf{r}}$. If the columns of $\mathbf{A}$ are linearly independent, then it is possible to find exactly $P - N$ eigenvectors in the P-dimensional space that are orthogonal to *all* the columns of $\mathbf{A}$. If $\mathbf{v}_i$ denotes one of these eigenvectors, then

$$\mathbf{R}_{\mathbf{r}}\mathbf{v}_i = \sum_{n=1}^{N}\left(P_n\mathbf{a}_n\mathbf{a}_n^H\mathbf{v}_i\right) + \sigma_\eta^2\mathbf{v}_i = \sigma_\eta^2\mathbf{v}_i \tag{3.28}$$

since $\mathbf{a}_n\mathbf{v}_i = 0 \quad \forall n$. Thus these eigenvectors all correspond to eigenvalues $\lambda_i = \sigma_\eta^2$. The

remaining N eigenvectors cannot be orthogonal to the columns of $\mathbf{A}$. Therefore they lie in the subspace defined by $\mathbf{A}$ and have eigenvalues greater than σ_η^2. This subspace is called the *signal subspace*. The complementary subspace, which is orthogonal to the signal subspace, is called the *noise subspace*.

3.2.1 MUSIC

The Multiple Signal Classification (MUSIC) [129] subspace method has the great advantage over the other methods to be introduced later, that it is readily adapted to models featuring near field sources, irregular sampling intervals and samples that are not colinear [20, 129]. This is applicable if the trajectory of a moving receiver is known and is not a straight line traversed at constant speed (detected by some inertial sensor for instance).

Given some estimate $\widehat{N}$ of the number of sources present, the eigendecomposition of $\widehat{\mathbf{R}}_\mathbf{r}$ can be represented as

$$\widehat{\mathbf{R}}_\mathbf{r} = \widehat{\mathbf{V}}\widehat{\mathbf{\Lambda}}\widehat{\mathbf{V}}^H, \tag{3.29}$$

where

$$\widehat{\mathbf{V}} = \left[\begin{array}{cc} \widehat{\mathbf{V}}_{\text{sig}} & \widehat{\mathbf{V}}_{\text{noise}} \end{array} \right], \tag{3.30}$$

$\widehat{\mathbf{\Lambda}}$ is diagonal and the orthonormal eigenvectors comprising $\widehat{\mathbf{V}}_{\text{sig}}$ correspond to the $\widehat{N}$ largest eigenvalues. A projection matrix for the noise subspace can be formed as

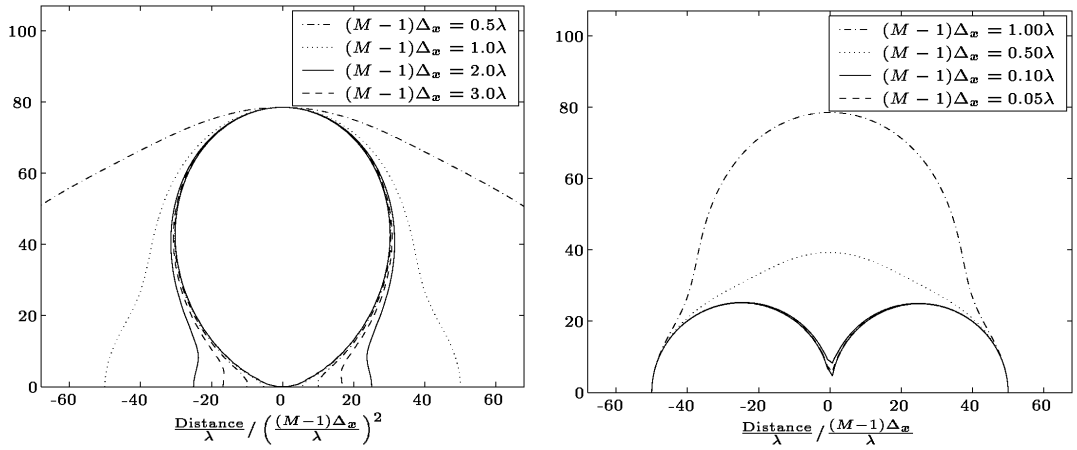
$$\mathbf{P}_{\text{noise}} = \mathbf{V}_{\text{noise}} \left(\mathbf{V}_{\text{noise}}^H \mathbf{V}_{\text{noise}} \right)^{-1} \mathbf{V}_{\text{noise}}^H = \mathbf{V}_{\text{noise}} \mathbf{V}_{\text{noise}}^H. \tag{3.31}$$

If a vector lies in the signal subspace, then its projection onto the noise subspace will be small. The set of possible spatial parameters which can produce vectors lying in the signal space can be searched for those having such a small projection. For complex sinusoidal signals (analogous to far field sources) the search is one dimensional, and the inverse square of the projection magnitude forms what is known as a *pseudo-spectrum* since peaks indicate the presence of signal at a given frequency. For near field sources, the search is two dimensional, but in some situations can still be practical to perform.

The squared magnitude of the projection $\mathbf{p}$ of a vector $\mathbf{a}(\varpi)$ onto the noise subspace is

$$\|\mathbf{p}\|^2 = \mathbf{a}\mathbf{V}\mathbf{V}^H\mathbf{V}\mathbf{V}^H\mathbf{a} = \|\mathbf{V}^H\mathbf{a}\|^2 \tag{3.32}$$

and the inverse of this forms the pseudo-spectrum, the location of the $\widehat{N}$ largest peaks of which are taken as the solution to the estimation problem.



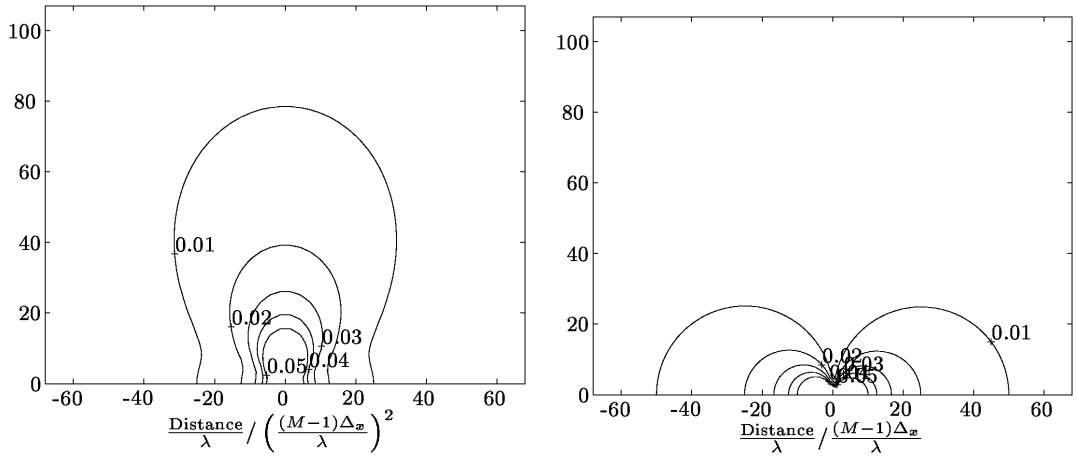
(a) (b)

Figure 3.4: Contours of the relative error (3.34) which results from calculations of the channel using the far field approximation for long (a) and short (b) array lengths. The contour value is 0.01. In both cases the horizontal axis represents distance *along* the array axis, and the vertical axis represents distance *from* the array axis. For short arrays the dimensions of the "near field region" varies linearly with array length, whereas for long arrays the dimensions of the near field region varies with the *square* of the array length. For this reason, in (a), the axes have been scaled by the square of the array length, and in (b) by the array length.

3.2.2 Near Field MUSIC

There have been numerous methods proposed to allow the localisation of near field sources [2, 89, 139]. These were evaluated in the course of the research reported here. It was found that the approximations which are key to the operation of these methods are frequently not valid for sources which are truly *near* (not more than a few wavelengths from) an array. By contrast MUSIC, (given a good correlation estimate) was found to perform robust estimation even for very near sources. There are some additional considerations with using MUSIC for the near field situation. One is the issue, already mentioned, of estimating the correlation matrix when only one snapshot is available at any one location in space.

If measurements of the channel are available at more than one frequency, averaging over different frequencies can be used to overcome this problem. It was found that useful results could be obtained if the frequency variation allowed at least one 2π phase rotation in the term $e^{-j2\pi R_{mn}/\lambda}$, which means the frequency range must be at least of the order of $\Delta_f/f > \lambda/(2\pi R_{mn})$. There is a conflicting requirement that the percentage frequency variation be small so that the scattering scenario be essentially the same for each frequency. This is turn places a limitation on how close the scatterers can be to the receiving antenna. For instance a 2% frequency variation requires the sources be at least 8 wavelengths distant. In most situations this is not an unreasonable requirement, but



(a) (b)

Figure 3.5: Contours of the relative error (3.34) which results from calculations of the channel using the far field approximation for an array length of 2 wavelengths (a) and 0.01 wavelengths (b). In both cases the horizontal axis represents distance *along* the array axis, and the vertical axis represents distance *from* the array axis. In (a), the axes have been scaled by the square of the array length, and in (b) by the array length.

the amount of frequency variation over which the scattering scenario can be considered essentially unchanged is unknown.

The other issue is the two dimensional search. Unlike the far field case there is no finite search range. There is however, a region beyond which a near field source becomes indistinguishable from a far field source. Referring to (3.22) when $x_m \ll d_l$, and the array is linear, so that $y_m = 0$,

$$\begin{aligned} D_{mn} &= d_l \sqrt{1 + \left(\frac{x_m}{d_l}\right)^2 - 2\frac{x_m}{d_l}\sin\theta_n} \\ &\approx d_l - x_m \sin\theta_n. \end{aligned} \tag{3.33}$$

When this expression is used for D_{ml} in the complex exponent of (3.19), and d_l is used for the multiplicative factor in that equation, the expression for the elements of the steering vector simplify to the far field case (with a multiplicative factor of $e^{-j2\pi d_l/\lambda}/d_l$ which is the same for all elements of the array).

The relative error between the near field and far field expressions for a single source point is thus

$$\varepsilon = \frac{\left|\frac{1}{D_{ml}} e^{-j\frac{2\pi}{\lambda}D_{ml}} - \frac{1}{d_l} e^{-j\frac{2\pi}{\lambda}(d_l - x_m \sin\theta)}\right|}{\left|\frac{1}{D_{ml}} e^{-j\frac{2\pi}{\lambda}D_{ml}}\right|}. \tag{3.34}$$

Contours of this quantity will contain all the points for which the relative error is larger

than a certain quantity. This is called here the *near field region*. The contours are shown in Fig. 3.4 and Fig. 3.5. For very short arrays the linear dimensions of the near field region are approximately proportional to the length of the array. For arrays longer than about a wavelength (subfigure (a) in both figures), which includes most arrays of practical interest, the linear dimensions of the near field region are proportional to the *square* of the array length.

In order to ensure that the error does not exceed 1% it is sufficient for the "pseudo-spectrum" to be evaluated at all points inside the 1% contour. Although the region is clearly not semicircular in shape, in practice the "pseudo-spectrum" was evaluated at all the points in a semicircular region based on a polar grid.

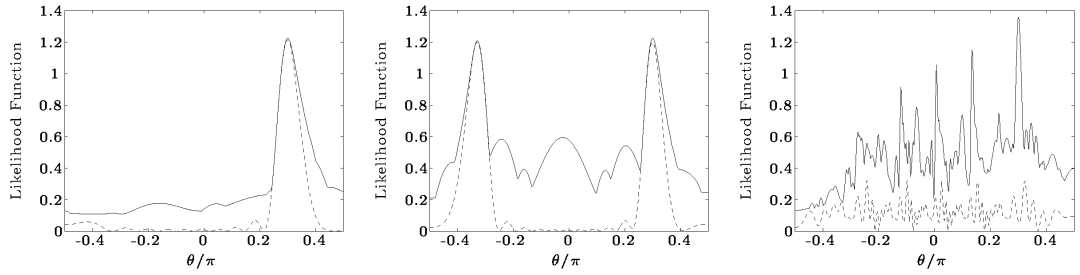
3.2.3 Aliasing

It is well known that if information is not to be lost in sampling a band-limited signal, the sampling rate must exceed twice the highest frequency in the signal [e.g.,103]. In far field array processing this equates to the requirement that elements be placed closer than half of one wavelength. If elements are placed further apart than this, an ambiguity arises in determining the location of sources. For the purposes of prediction, if the sampling continues at *exactly* the same spacing, the ambiguity is not apparent, and does not affect prediction accuracy. In practice, such exact spacing is not likely, and the possibility of aliasing must be considered.

Although it is not considered elsewhere in this thesis, the observation is made here that for some time division multiple access (TDMA) systems, for high speeds of travel and high transmission frequencies it is quite possible that the spacing between consecutive samples may exceed half of one wavelength. Even for far field sources then, spatial aliasing may occur. This factor must of course be considered in any real system design.

For near field sources the situation is considerably more complicated. The Fourier transform with respect to x_m of (3.21) (using (3.876 1&2) of [64]) shows that the "spatial spectrum" is not band-limited [3]. To reliably localise several sources, the element spacing may thus need to be considerably smaller than a half wavelength.

For a very few sources which are not near each other however, the nearness of sources to an array actually helps to resolve the ambiguity. This is demonstrated in Fig. 3.6 which shows the likelihood function (3.7) of bearing angle θ for a source located 30λ from the centre of an 11 element array. The solid line represents the likelihood function for the near field model, and the broken line that for the far field model (for the near field function the dependence on range has been eliminated by taking the maximum value for *any* range). The true value of θ is 0.3π radians, and the SNR is 30 dB. In subfigure (a), the element spacing is 0.49λ, and no aliasing occurs. In subfigure (b) the element spacing is 0.60λ and the location of the source is ambiguous. In subfigure



(a) (b) (c)

Figure 3.6: Likelihood Functions (3.7) for near field (solid line) and far field (broken line) models for array spacings of (a) 0.49λ, (b) 0.60λ and (c) 2.60λ.

(c) the element spacing is 2.60λ. There are now even more aliases present, which the far field model cannot resolve. The near field model however has no difficulty resolving the true location. The profile of the likelihood as a function of the range (not shown here) reveals that the range can also be determined unambiguously. If there are many sources however, the likelihood function may again become ambiguous.

3.2.4 Minimum Norm Spectral Estimation

For far field sources, when the spatial frequencies are fixed, the projection onto the noise subspace in (2.68) of a vector $\mathbf{a}$ of the form defined in (3.32)

$$\mathbf{a}(\varpi) = (e^{j\varpi 0}, e^{j\varpi 1}, \ldots, e^{j\varpi(M-1)})^T$$

can be expressed as the evaluation of a polynomial in $e^{j\varpi}$. The zeros of the projection (peaks of the pseudo-spectrum) thus correspond to zeros of this polynomial which lie on the unit circle. The degree of the polynomial is M, and the number of roots which correspond to sources is $N < M$, hence there are also extraneous zeros.

The noise subspace can be characterised by a single vector $\mathbf{d} = [d_0, d_1, \ldots, d_{P-1}]^T$ in the noise subspace. The corresponding polynomial is thus $D(z) = \sum_{p=0}^{P-1} d_p z^{-p}$. It has been shown [84, 85] that if this vector is chosen so that the norm $\|\mathbf{d}\|^2$ is minimised subject to the constraint that one of the components is fixed (e.g., the first component is fixed to unity) then the extraneous roots are all *inside* the unit circle, and are approximately uniformly distributed around the unit circle. This is a desirable condition in that it minimises the possible confusion of the true roots and the extraneous roots.

The constraints that the vector lies in the noise subspace and that first term be unity can be expressed as $\mathbf{d} = \mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^H\mathbf{d}$ and $\mathbf{d}^H\mathbf{e}_1 = 1$ where $\mathbf{e}_1 = (1, 0, \ldots, 0)^T$. These constraints may be combined as $\mathbf{d}^H\mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^H\mathbf{e}_1 = 1$. The Lagrangian associated

with finding the minimum norm may thus be expressed as

$$\mathcal{L} = \mathbf{d}^H\mathbf{d} + \nu(1 - \mathbf{d}^H\mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^H\mathbf{e}_1). \tag{3.35}$$

Equating the derivative to zero gives $\mathbf{d} - \nu\mathbf{V}_{\text{noise}}\mathbf{V}_{\text{noise}}^H\mathbf{e}_1 = \mathbf{0}$. The matrix of noise eigenvectors can be partitioned as

$$\widehat{\mathbf{V}}_{\text{noise}} = \begin{bmatrix} \mathbf{c}^H \\ \mathbf{V}_{\text{noise}}^{\downarrow} \end{bmatrix} \tag{3.36}$$

so that $\mathbf{c} = \mathbf{V}_{\text{noise}}^H\mathbf{e}_1$, and so $d_0 = 1 = \nu\mathbf{c}^H\mathbf{c}$, which implies that the Lagrange multiplier $\nu = 1/(\mathbf{c}^H\mathbf{c})$ and so the vector $\mathbf{d}$ is then given by

$$\mathbf{d} = \begin{bmatrix} 1 \\ \mathbf{V}_{\text{noise}}^{\downarrow}\mathbf{c}/(\mathbf{c}^H\mathbf{c}) \end{bmatrix}. \tag{3.37}$$

The frequency estimates are formed by finding the $\widehat{N}$ roots of the prediction filter

$$G(z) = 1 + \sum_{p=1}^{P-1} g_p z^{-p} \tag{3.38}$$

with the largest absolute value (i.e., closest to the unit circle), and projecting these along radii onto the unit circle (i.e., taking only the complex argument and replacing the complex magnitude by unity).

3.2.5 Principal Components Linear Prediction (PCLP)

The Wold decomposition theorem [169] (see Section 2.6.3) states that a general random process can be written as a sum of a *regular process*, having a continuous spectrum and a *predictable process*, having a discrete spectrum. Linear prediction of a *noiseless* process is thus functionally equivalent to estimating the parameters of the discrete frequency components. This is utilised in the principal components linear prediction method of spectrum estimation [157, 158] .

The linear prediction coefficients are the solution to the normal equations $\mathbf{R_r d} = \mathbf{0}$. For $P > N$, the solution is not unique, but for reasons similar to those in Section 3.2.4, the minimum-norm solution to the equation is desirable. The correlation matrix estimate and the vector $\mathbf{d}$ can be partitioned so that the normal equations can be expressed as

$$\begin{bmatrix} f & -\mathbf{r}^H \\ -\mathbf{r} & \mathbf{R}^{\searrow} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{d}^{\downarrow} \end{bmatrix} = \mathbf{0}. \tag{3.39}$$

Since the first component d_0 is constrained to be equal to one, minimisation of $\|\mathbf{d}\|^2$

is equivalent to minimisation of $\|\mathbf{d}^{\downarrow}\|^2$. When noise *is* present, the eigenvectors corresponding to the noise subspace are effectively eliminated by using the rank $\widehat{N}$ Moore-Penrose generalised inverse. If the $\widehat{N}$ largest eigenvalues and the corresponding eigenvectors of $\mathbf{R}^{\searrow}$ are $\lambda_1, \lambda_2, \ldots, \lambda_{\widehat{N}}$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{\widehat{N}}$ then

$$\mathbf{d}^{\downarrow} = \mathbf{R}^{\searrow+(\widehat{N})}\mathbf{r} = -\sum_{n=1}^{\widehat{N}} \left(\frac{\mathbf{v}_n^H r}{\lambda_n} \right) \mathbf{v}_n. \tag{3.40}$$

As with the minimum norm procedure, the frequency estimates are obtained by projecting the zeros of the filter along a radius to the unit circle.

3.2.6 ESPRIT Algorithm

In the Estimation of Signal Parameters via Rotational Invariance (ESPRIT) technique [125] use is made of the structure of the signal matrix $\mathbf{A}$ (see equation (2.68)). Sub-matrices of the matrix $\mathbf{A}$ are taken; deleting the top row to obtain $\mathbf{A}^{\downarrow}$ and the bottom row to obtain $\mathbf{A}^{\uparrow}$. $\mathbf{\Xi}$ is defined as $\mathbf{\Xi} = \text{diag}\left(e^{-j\varpi_1}, e^{-j\varpi_2}, \ldots, e^{-j\varpi_{\widehat{N}}}\right)$, so $\mathbf{A}^{\downarrow} = \mathbf{A}^{\uparrow}\mathbf{\Xi}$.

The signal matrix $\mathbf{A}$ and the signal eigenvectors span the same subspace, and hence can be related by the expression $\widehat{\mathbf{V}}_{\text{sig}}\mathbf{\Psi} = \mathbf{A}$. Sub-matrices of the matrix of eigenvectors spanning the signal subspace ($\widehat{\mathbf{V}}_{\text{sig}}$) are obtained by deleting the top row to obtain $\mathbf{V}_{\text{sig}}^{\downarrow}$ and the bottom row to obtain $\mathbf{V}_{\text{sig}}^{\uparrow}$.

It is simple to show that $\mathbf{V}_{\text{sig}}^{\uparrow}\mathbf{\Psi} = \mathbf{A}^{\uparrow}$, and $\mathbf{V}_{\text{sig}}^{\downarrow}\mathbf{\Psi} = \mathbf{A}^{\downarrow}$, and hence that $\mathbf{V}_{\text{sig}}^{\downarrow} = \mathbf{V}_{\text{sig}}^{\uparrow}\left(\mathbf{\Psi\Xi\Psi}^{-1}\right)$. This can be rearranged as $\mathbf{V}_{\text{sig}}^{\downarrow} = \mathbf{V}_{\text{sig}}^{\uparrow}\mathbf{\Phi}$ where $\mathbf{\Phi} = \mathbf{\Psi\Xi\Psi}^{-1}$. $\mathbf{\Phi}$ can be calculated from the first relationship, and from the second, the eigenvalues of $\mathbf{\Phi}$ are the diagonal elements of $\mathbf{\Xi}$.

$\mathbf{\Phi}$ can be estimated using the least squares approach as

$$\mathbf{\Phi}_{\text{LS}} = \left(\mathbf{V}_{\text{sig}}^{\uparrow H}\mathbf{V}_{\text{sig}}^{\uparrow}\right)^{-1}\mathbf{V}_{\text{sig}}^{\uparrow H}\mathbf{V}_{\text{sig}}^{\downarrow} \tag{3.41}$$

or the total least squares approach; if the singular value decomposition of the matrix $\left[\mathbf{V}_{\text{sig}}^{\uparrow}\mathbf{V}_{\text{sig}}^{\downarrow}\right]$ can be written as

$$\left[\mathbf{E}_{\text{sig}}^{\uparrow}\mathbf{E}_{\text{sig}}^{\downarrow}\right] = \mathbf{U\Upsilon}\begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}, \tag{3.42}$$

then the total least squares solution is $\mathbf{\Phi}_{\text{TLS}} = \mathbf{V}_{12}\mathbf{V}_{22}^{-1}$ [62, pp595*ff*].

In [6] the magnitude of the roots is retained; in effect, the roots are *not* projected onto the unit circle. While this may increase the likelihood of a solution (by decreasing the MSE), whether the freedom to include *decaying* or *growing* sinusoids is appropriate or not depends on the validity of the far field model to the particular situation — data based on the Doppler frequencies of far field sources will only contain complex sinusoids

of constant amplitude.

3.3 Gaussian Model

Although the focus in this thesis is on the discrete point sources channel model, other models have also been considered, and are explained in the following sections.

If there is a continuum of far field scatterers (e.g., [30]) surrounding a receiving antenna, the summation of signals such as (2.1) in (2.66) for example may be replaced by an integral;

$$r_m = s_m \int_{-\pi}^{\pi} \zeta(\theta, t_m) e^{-j\omega_c \tau_m(\theta)} d\theta + \eta_m, \tag{3.43}$$

or equivalently, assuming as before that the signal $s_m = 1 \, \forall m$, and arranging the samples as a vector

$$\mathbf{r} = \int_{-\pi}^{\pi} \zeta(\theta)\mathbf{a}(\theta) d\theta + \boldsymbol{\eta}_m \tag{3.44}$$

where $\mathbf{a}(\theta) = [e^{-j\omega_c\tau_1(\theta)}, e^{-j\omega_c\tau_2(\theta)}, \ldots, e^{-j\omega_c\tau_M(\theta)}]^T$ and other terms are as defined in (2.68). The following assumptions are made about the complex attenuation function $\zeta(\theta)$:

A-1 $\zeta(\theta)$ is a random function.

A-2 $E\{\zeta(\theta)\} = 0$.

A-3 $\zeta(\theta)$ is independent for different angles, so the probability density functions have the property:

$$p(\zeta(\theta_1), \zeta(\theta_2)) = p(\zeta(\theta_1))p(\zeta(\theta_2)). \tag{3.45}$$

Using assumption A-1 and A-3 and appealing to the central limit theorem, $\mathbf{r}$ can be assumed complex normal. Using assumption A-2, the mean of $\mathbf{r}$ is zero. The correlation matrix $\mathbf{R}_\mathbf{r}$ of $\mathbf{r}$ is given by

$$\begin{aligned} E\left\{\mathbf{r}\mathbf{r}^H\right\} &= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathbf{a}(\theta_1)\mathbf{a}(\theta_2)^H E\left\{\zeta(\theta_1)\zeta(\theta_2)^*\right\} d\theta_1 d\theta_2 \\ &= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathbf{a}(\theta_1)\mathbf{a}(\theta_2)^H r_\zeta(\theta_1)\delta(\theta_1 - \theta_2) d\theta_1 d\theta_2 \\ &= \int_{-\pi}^{\pi} \mathbf{a}(\theta)\mathbf{a}(\theta)^H r_\zeta(\theta) d\theta, \end{aligned} \tag{3.46}$$

where $r_\zeta(\theta)$ is the variance of ζ at direction θ, and the first equality holds due to assumptions A-2 and A-3. The model can thus be expressed as a sequence of complex normal samples:

$$\mathbf{r} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_r). \tag{3.47}$$

The prediction algorithm corresponding to this method is now presented. The vector $\mathbf{y}$ here represents the *measured* complex signal envelope, and the vector $\mathbf{x}$ represents the *predicted* signal. If $\mathbf{x}$ and $\mathbf{y}$ are jointly complex normal vectors with zero mean, and $\mathbf{z} = [\mathbf{x}^T \mathbf{y}^T]^T$, then

$$\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C_z}), \quad \mathbf{C_z} = \begin{bmatrix} \mathbf{C_x} & \mathbf{C_{xy}} \\ \mathbf{C_{xy}^H} & \mathbf{C_y} \end{bmatrix}. \tag{3.48}$$

The minimum mean square error estimator (MMSE) for $\mathbf{x}$ given $\mathbf{y}$ when $\mathbf{C_y}$ is invertible is given by the conditional expectation

$$\widehat{\mathbf{x}} = E\{\mathbf{x}|\mathbf{y}\} = \mathbf{C_{xy}}\mathbf{C_y^{-1}}\mathbf{y}. \tag{3.49}$$

When $\mathbf{C_y}$ is singular, (i.e., with rank of $r < N$) inversion is undefined and the Moore-Penrose generalised inverse of $\mathbf{C}$, $\mathbf{C}^+$ is used. This is shown as follows. Define the zero mean complex normal $r \times 1$ vector, $\mathbf{g}$, with a correlation matrix, $\mathbf{Cg} = \mathbf{I}_r$ such that:

$$\mathbf{y} = \mathbf{D}\mathbf{g}, \tag{3.50}$$

where $\mathbf{D}$ is an $N \times r$ matrix given by

$$\mathbf{D} = \mathbf{U}_1\sqrt{\boldsymbol{\lambda}}_1, \tag{3.51}$$

and where $\mathbf{U}_1$ and $\boldsymbol{\lambda}_1$ are given in the following eigendecomposition of $\mathbf{C_y}$

$$\mathbf{C_y} = [\mathbf{U}_1\mathbf{U}_2] \begin{bmatrix} \boldsymbol{\lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{U}_1\mathbf{U}_2]^H. \tag{3.52}$$

Since $\mathbf{y}$ is completely defined by $\mathbf{g}$, $\widehat{\mathbf{x}} = E[\mathbf{x}|\mathbf{y}] = E[\mathbf{x}|\mathbf{g}] = \mathbf{C_{xg}}\mathbf{C_g^{-1}}\mathbf{g} = \mathbf{C_{xg}}\mathbf{g}$. Looking for a solution in a similar form to that of (3.49), ($\widehat{\mathbf{x}} = \mathbf{C_{xy}}\widetilde{\mathbf{C}}_\mathbf{y}\mathbf{y}$) will lead to the choice of $\widetilde{\mathbf{C}}_\mathbf{y}$ given by

$$\widetilde{\mathbf{C}}_\mathbf{y} = \mathbf{D}\left(\mathbf{D}^H\mathbf{D}\right)^{-2}\mathbf{D}^H = \mathbf{U}_1\boldsymbol{\lambda}_1^{-1}\mathbf{U}_1^H, \tag{3.53}$$

which is the Moore-Penrose generalised inverse of $\mathbf{C_y}$.

The estimation error of $\mathbf{x}$ given by $\mathbf{e} = \widehat{\mathbf{x}} - \mathbf{x}$ is a zero mean complex normal vector with correlation matrix $\mathbf{C_{ee}} = \mathbf{C_x} - \mathbf{C_{xy}}\widetilde{\mathbf{C}}_\mathbf{y}\mathbf{C_{yx}}$. When $r = N$ the estimation error

correlation matrix is given by

$$\mathbf{C_{ee}} = \mathbf{C_x} - \mathbf{C_{xy}}\mathbf{C_y^{-1}}\mathbf{C_{yx}}. \tag{3.54}$$

In the eigendecomposition of $\mathbf{C_y}$, eigenvalues smaller than those required to include a certain percentage (the majority) of the energy in the signal are taken to be zero.

The correlation between the signal at any two points in space may be modelled, for example, as the Bessel function J_0 [76], so the $ij-$th element of the matrices $\mathbf{C_{xy}}$ and $\mathbf{C_y}$ are given by $J_0\left((i-j)2\pi\frac{\Delta_x}{\lambda}\right)$. As shown in Section 2.3, this correlation function occurs when the scatterers are distributed uniformly in a plane surrounding the receiver.

Rather than assume that the correlation function is fixed, an alternative technique is to attempt to estimate the correlation function from the data. This is essentially the same as estimating a high order autoregressive (AR) model, and using this for prediction.

3.4 Polynomial Prediction

A simple polynomial prediction model may also be used. A complex polynomial of order $\widehat{N}$ is fitted to the data using the subset $\mathbf{B}$ of a Vandermonde matrix

$$\mathbf{B}(N) = \begin{bmatrix} 1 & 1^1 & 1^2 & \dots & 1 \\ 2^0 & 2^1 & 2^2 & \dots & 2^{(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ M^0 & M^1 & M^2 & \dots & M^{(N-1)} \end{bmatrix}, \tag{3.55}$$

and the array of polynomial coefficients $\mathbf{b} = [b_0, b_1, \dots, b_{\widehat{N}-1}]^T$ so that $\mathbf{r} = \mathbf{B}(\widehat{N})\mathbf{b}$. The least squares solution to this equation found in order to perform prediction is $\widehat{\mathbf{b}}(\widehat{N}) = \left(\mathbf{B}(\widehat{N})^H\mathbf{B}(\widehat{N})\right)^{-1}\mathbf{B}(\widehat{N})^H\mathbf{r}$.

The order $\widehat{N}$ of the polynomial used in this method was chosen using a form of the MDL criterion (see section 3.6) appropriate to the polynomial model:

$$\widehat{N} = \arg\min_N M\log\left(\frac{1}{M}\left|\mathrm{Re}\left(\mathbf{r}^H\mathbf{r} - \mathbf{r}^H\mathbf{B}(N)\widehat{\mathbf{b}}(N)\right)\right|\right) + N\log(2M). \tag{3.56}$$

3.5 Fixed Sector Prediction

The performance of prediction based on point source models depends critically on the accuracy of the spatial frequencies involved, since error in frequency becomes more and more apparent as the attempted prediction range increases. This will be shown in more detail in Section 4.4.

A method of prediction which has been independently suggested in discussion by

various researchers [163] is to assume that all the sources within a certain sector surrounding the receiver can be approximated by a single source at the centre of the sector. This approach turns out to be unworkable, for reasons which will be explained in this section. The approach does however provide an insight into the importance of accurate spatial frequency estimation.

Suppose that the sources in a sector of size $2\Delta_\theta$ are modelled by a single source at the centre of the sector at angle θ_0. The model is most in error if the *actual* source location is at the edge of the sector at $\theta_0 \pm \Delta_\theta$ (if the sector is large and if there are a large number of sources present in the sector, it is actually possible to find an arrangement of sources which will produce a model with a larger error than a single source at the edge of the sector, but this possibility is ignored here). The Doppler frequencies measured in radians per sample interval, ϖ and $\widehat{\varpi}$, correspond to the source locations θ_0 and $\theta_0 \pm \Delta_\theta$ respectively (using the relationship $\varpi = (2\pi\Delta_x \sin\theta)/\lambda$). If the signal samples are $\mathbf{r} = (r_{m_1}, r_{m_2}, \ldots r_{m_M})^T$, the signal amplitude is ζ and the samples are indexed $\mathbf{m} = (m_1, m_2, \ldots, m_M)^T$, then the received signal is given by $\mathbf{r} = \mathbf{a}\zeta = \zeta e^{j\varpi\mathbf{m}}$. Using (3.4), the source amplitude will be estimated as $\widehat{\zeta} = \mathbf{a}^H\mathbf{r}/M = (e^{-j\widehat{\varpi}\mathbf{m}^T})\mathbf{r}/M$, and the predicted signal as $\widehat{\mathbf{r}} = \widehat{\zeta}e^{j\widehat{\varpi}\mathbf{m}}$. In this section and in Section 4.4 $e^{j\cdot}$ where $\cdot$ is a matrix or vector means exponentiation of the *elements* of a vector or matrix, rather than the matrix exponential. The squared error at sample position m_f can be calculated as

$$\begin{aligned}
|\epsilon(m_f)|^2 &= \left|\zeta e^{j\varpi m_f} - \widehat{\zeta} e^{j\widehat{\varpi} m_f}\right|^2 \\
&= |\zeta|^2 \left| e^{j\varpi m_f} - \frac{1}{M}\sum_m e^{j(\varpi-\widehat{\varpi})m} e^{j\widehat{\varpi} m_f}\right|^2 \\
&= |\zeta|^2 \left(1 + \left|\frac{1}{M}\sum_m e^{j(\varpi-\widehat{\varpi})m}\right|^2 - 2\,\mathrm{Re}\left(e^{j(\varpi-\widehat{\varpi})m_f}\frac{1}{M}\sum_m e^{j(\varpi-\widehat{\varpi})m}\right)\right),
\end{aligned} \tag{3.57}$$

which can be seen to depend on the Doppler frequencies only through the frequency difference $\Delta_\varpi = \varpi - \widehat{\varpi}$. An implication of this is that to achieve about the same error for each sector, the sectors must be arranged so that each occupies an equal range of Doppler frequencies. When θ is small, $\partial\varpi/\partial\theta$ is large, so the sectors must be small, whereas when θ is large (i.e., the source direction is orthogonal to the direction of travel), the sectors can be larger.

For a fixed error $\epsilon(m_f)$, this expression can be rearranged to obtain an allowed error Δ_ϖ as a function of the prediction point index m_f. If the prediction point is expressed as a percentage of the measurement segment length, the allowed error multiplied by $M-1$ is independent of the actual value of M. This function is shown in Fig. 3.7

If the distance between samples is Δ_x, then the range of possible Doppler frequencies is from $-2\pi\Delta_x/\lambda$ to $2\pi\Delta_x/\lambda$. If the frequency range covered by each sector is $2\Delta_\varpi$,

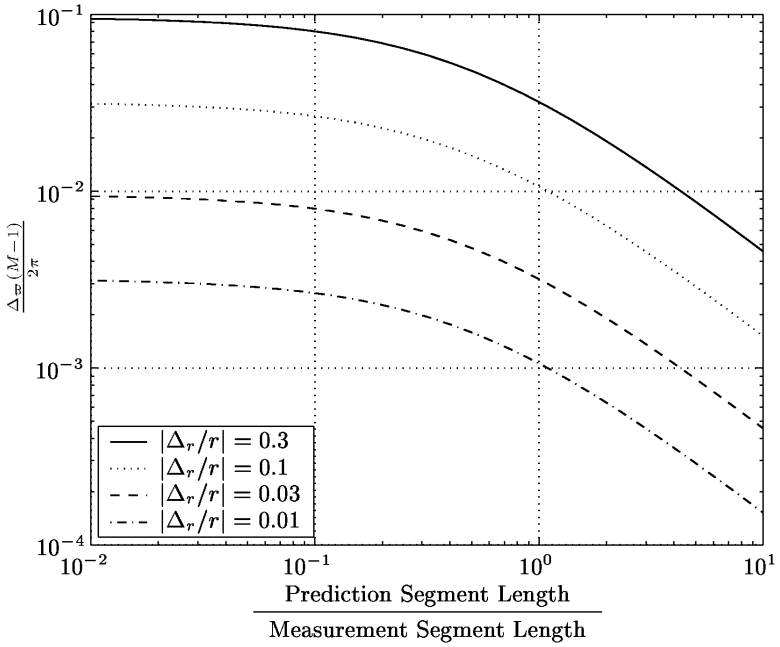


Figure 3.7: The allowed error Δ_ϖ in ϖ if a given prediction range is required. The allowed channel error Δ_r is expressed as a fraction of the actual channel value. Note that when multiplied by $M-1$, the error Δ_ϖ is independent of the actual number of samples M.

then the number of sectors required and thus the number of complex parameters in the model is $2\pi\Delta_x/(\lambda\Delta_\varpi)$. Since the synthetic array spans a distance of $(M-1)\Delta_x$, the number of samples needed to reconstruct the signal, and therefore the number of *independent* samples available, is approximately $2(M-1)\Delta_x/\lambda$. For this number to be greater than the number of parameters requiring estimation, it is then required that

$$\Delta_\varpi \frac{M-1}{2\pi} > \frac{1}{2}. \tag{3.58}$$

It is obvious from Fig. 3.7, that for reasonable error levels and prediction ranges, this requirement is not met. For data which has any reasonable level of noise, estimating a larger number of parameters than the number of uncorrelated data samples available is not a robust procedure.

If the number of sectors N and the sector locations are fixed, the steering matrix for these sector locations can be calculated *a priori*, as well as its generalised inverse. If $\mathbf{m} = (m_1, m_2, \dots, m_M)^T$ and $\boldsymbol{\varpi} = (\varpi_1, \varpi_2, \dots \varpi_N)^T$, then $\mathbf{A} = e^{j\mathbf{m}\boldsymbol{\varpi}^T}$.

Similarly if $\mathbf{m}_2 = (m_1, m_2, \dots, m_M, \dots, m_{M_2})^T$ represents the indices of the synthetic array extended to include the predicted data points, then $\mathbf{A}_2 = e^{j\mathbf{m}_2\boldsymbol{\varpi}^T}$ is the steering vector for the extended array. Since $\mathbf{r} = \mathbf{A}\boldsymbol{\zeta} + \boldsymbol{\eta}$, the estimated source amplitudes are given by $\widehat{\boldsymbol{\zeta}} = \mathbf{A}^+\mathbf{r}$ and so the estimated (predicted) signal is given by $\widehat{\mathbf{r}} = \mathbf{A}_2\widehat{\boldsymbol{\zeta}} = \mathbf{A}_2\mathbf{A}^+\mathbf{r}$.

If $\Delta_x = \lambda/2$, the frequencies $\boldsymbol{\varpi}$ define an orthogonal set of sinusoids, and so estimating the source complex amplitudes $\boldsymbol{\zeta}$ amounts to performing a discrete time Fourier transform (DTFT). The inverse of the DTFT is periodic, and so extrapolation based on the DTFT is simply repeating the data that has already been received, possibly with a period depending on the number of sectors, and thus possibly different to the data collection interval. If the number of sectors is smaller than the number of data samples ($N < M$), the repeated data sets overlap, and if $N > M$, the repeated data sets are separated by zeros.

If $\Delta_x < \lambda/2$, the sinusoids in the set defined by $\boldsymbol{\varpi}$ are no longer orthogonal. Although the data inside the measurement segment is accurately represented, the estimated amplitudes are very large. For data outside the measurement segment, these large amplitudes no longer cancel each other, and so very large signals are predicted.

The behaviour in each of these cases is demonstrated by the following examples of $\mathbf{A}_2\mathbf{A}^+$. For the left matrix, $\Delta_x/\lambda = 0.5$ and for the right matrix $\Delta_x/\lambda = 0.1$. In both cases $N = 5$, $M = 4$, and $M_2 = 8$.

$$\left(\mathbf{A}_2\mathbf{A}^+\right)_{0.5} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad \left(\mathbf{A}_2\mathbf{A}^+\right)_{0.1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 3.7 & -5.4 & 3.7 \\ -3.6 & 12.5 & -16.3 & 8.3 \\ -8 & 26.4 & -31.9 & 14.3 \\ -13.6 & 43.5 & -49.9 & 20.8 \end{bmatrix} \tag{3.59}$$

It is reasonable to conclude then, that a method which assigns source bin locations before the data has been received, is not able to provide reliable prediction.

3.6 Determining the Model Order $\hat{N}$

One of the problems in array processing is that of determining the model order, i.e., the number of sources. Determining the number of sources is necessary since all the parameters of the model are required for prediction to be possible. The number of sources cannot be estimated via the maximum likelihood criteria, since allowing a higher degree of freedom will only increase the likelihood. However, the Akaike information criterion [4] and the Minimum Description Length (MDL) [124, 131] criterion both address the problem by using the likelihood function penalised by a function of model order. The MDL criteria is known to be both unbiased and consistent. A specific choice of the parameters to be the eigenvalues ($\lambda_1 > \lambda_2 \cdots > \lambda_P$) and the eigenvectors of the

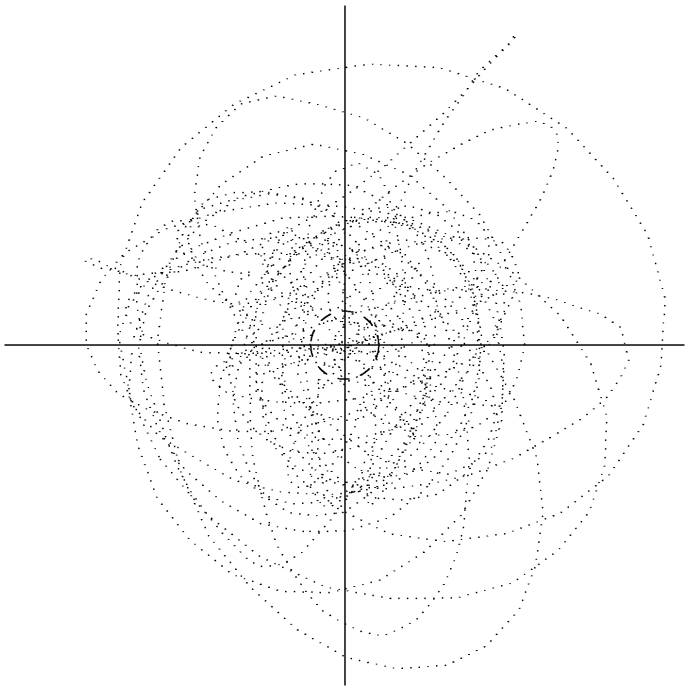


Figure 3.8: An example of a channel locus, showing the boundary that might be used for labelling the sample data for a support vector classification. The axes are the real and imaginary parts of the complex narrowband channel gain. If the locus enters the circular region within the next x wavelengths, the training data may be labelled as "will fade".

correlation matrix yields the version of the MDL by Wax and Kailath [165]. This is the method chosen in this thesis, so $\widehat{N}$ is given by

$$\widehat{N} = \arg\min_{N} \quad -\log\left(\left(\frac{\prod_{i=N+1}^{P}\lambda_i^{\frac{1}{P-N}}}{\frac{1}{P-N}\sum_{i=N+1}^{P}\lambda_i}\right)^{(P-N)(M-PL+L)}\right) + \frac{1}{2}N(2P-N)\log(M-PL+L). \tag{3.60}$$

3.7 Support Vector Machines

Support vector machines [24, 33, 36] are a relatively recent development in the field of machine learning. Briefly, support vector machines provide a method of performing linear classification or regression using a hyperplane in a feature space defined by a kernel function. The loss function to be minimised in finding this hyperplane is usually some form of "ε-insensitive" loss function. In many applications support vector machines are efficient, have good generalisation performance, and produce sparse representations of the classification or regression hypothesis. Since the kernel (the map into feature space) may be a non-linear function, support vector machines are a particularly efficient way

of modelling data in a non-linear way. Once a regression hypothesis has been formed on the basis of a sequential data set, it may be used for extrapolation or prediction of that data set.

This approach has been used on some of the data measured for this thesis in [69]. The prediction performance does not appear to improve on the linear methods presented here. However, the possibilities of using support vector machines in channel modelling have not yet been fully explored. For instance, if narrowband predicted channel information was to be used simply to provide advanced warning of fades, the problem could be simplified to a classification — a sequential set of channel measurements could be classified on the basis of whether or not the channel power became less than some threshold within the next x wavelengths. An example threshold is shown graphically in Fig. 3.8.

3.8 Summary and Contributions

The subspace methods of estimating the parameters of several channel models have been outlined in this chapter. Also several other channel models, together with their estimation algorithms have been presented. The algorithms will be used in Chapter 4 to assess the effectiveness of the models in allowing channel prediction. Most of the algorithms are well known in the literature. However, we itemise some specific contributions made in this chapter:

i. A concise expression has been derived for the first derivative of the likelihood function of the spatial parameters of discrete point sources models.

ii. A simple but novel method has been presented for estimating the correlation matrix which considerably enhances the effectiveness of subspace methods of spectral estimation.

iii. The notion that sources in a particular region (e.g., angular sector) can be represented by a single "effective" source is prevalent. New insights have been presented into this concept which show that for prediction purposes, such a model will generally be inadequate.

iv. The difficulty of performing estimation of the parameters of a near field model from a synthetic array has been highlighted.

Chapter 4

Simulations, Measurements and Bounds

THIS chapter presents the results of applying the models and algorithms of the previous chapters to several simulated and measured channels. Several different parameters, both of the channel and the mobile receiver, are adjusted to evaluate the effect of these on the reliability of channel prediction. These parameters include the SNR, the length of the mobile trajectory over which data is collected, and the number of discrete point sources present in the environment.

Several criteria are proposed for the evaluation of prediction accuracy, and one of these is chosen for the more extensive investigation of the remainder of the chapter.

The Cramer Rao bound for prediction accuracy is derived, which confirms the conclusions resulting from the simulations. A bound based on consideration of multipath dimensionality is also derived.

Some other factors are discussed which may limit prediction in some situations. One of these is rough surface scattering, and this is investigated in detail.

4.1 Performance Measures

The performance criterion chosen here is the distance (in wavelengths) for which the predicted and actual signal envelopes are within 20% of the root mean square (RMS) value of the envelope in the measurement segment. As can be seen in the example of prediction behaviour presented in Fig. 4.1, this may be a pessimistic criterion in many occasions, since the error even in the measurement segment may exceed 20%, especially when the SNR of the data is small.

The performance measure used in [6] (distance for which the predicted and actual signal envelopes are within 5% of the maximum amplitude value in the measurement segment) is not feasible for the SNR values considered here.

In order to obtain a performance measure for a particular set environment the

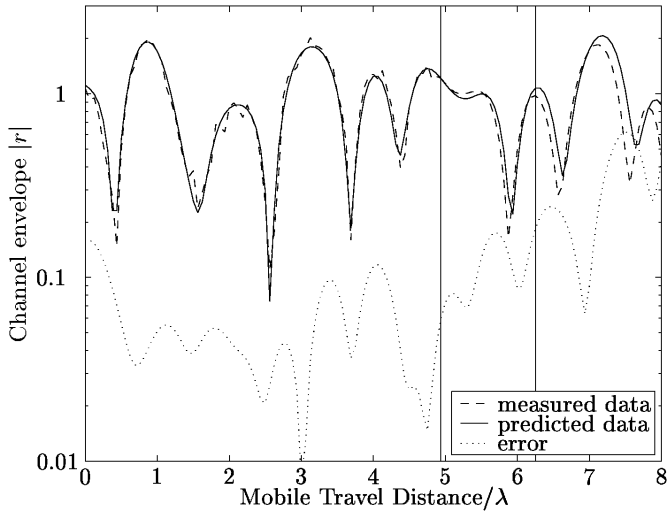


Figure 4.1: Example of Prediction Behaviour. The graph shows the magnitude of the measured and predicted channel and the prediction error. Only the measured data to the left of the first vertical line is used for prediction. The region between the two vertical lines is where the predicted envelope is within 20% of the actual envelope. The RMS level of the *measured* envelope has been scaled to 1, so prediction is said to "fail" when the error first exceeds 0.2.

mean of the prediction length for individual scenarios is used. An alternative measure which has some merit is to consider the expected error in the prediction segment. The *prediction gain* is defined in [17] as

$$G(t) = \frac{E\{|(r(t) - E\{r(t)\}|^2\}}{E\{|r(t) - \hat{r}(t)|^2\}}, \tag{4.1}$$

sometimes expressed in decibels. In order to obtain a distance-independent performance measure, the prediction length can then be defined as the distance at which the prediction gain falls below a certain threshold. In an experimental context this measure is only defined when there is sufficient data for an expectation to be calculated. This measure is in some ways similar to the previous one, but is not so suitable for small data sets where the expectations are harder to estimate. It does have the advantage however that the Cramer Rao bound of this measure is readily calculable. This is done in Section 4.4

Another performance measure which was considered, but not presented here, is the distance for which the *cumulative* error between the predicted and actual envelopes is less than some ratio of the RMS value in the measurement segment. This has the advantage of not "disqualifying" a prediction for error of large magnitude but very short duration. This measure could be appropriate if the prediction information were to be used to negotiate a change in power level, modulation or coding, for example.

Another performance measure could be devised involving the accuracy in time of

the estimates of envelope magnitude minima (i.e., fades), especially those fades of more than a certain depth. This measure would be particularly appropriate if the use to be made of the prediction information was to avoid serious fades, for example, by changing channels.

4.2 Simulations

The performance of each of the algorithms was investigated using simulation. In this section the simulations are described and the results are presented.

4.2.1 Simulation Setup

Simulations are used to provide a comparison of the performance of each algorithm. The simulations used a set of N far field sources with amplitudes taken from a complex normal distribution of variance $1/N$. The expected power of the simulated signal is thus unity (similar experiments with all sources of equal amplitude give almost identical results). Unless specified otherwise, the number of sources N is 5, 10, 20 and 100, the number of measurement points M is 40, the SNR is 20 dB, the measurement segment length $(M-1)\Delta$ is $\lambda/3$ or 10λ, and $\widehat{N}$ is chosen by the MDL criterion. The number of independent scenarios used to find the mean performance is 3000.

4.2.2 Simulation Results

In this section are some general comments which apply to all of the results. Specific comments about each study are in the following sections.

Some of the results presented in this section use a combination of the subspace methods, where the subspace method yielding the lowest MSE in the measurement segment is chosen for any particular scenario. In some cases the gradient calculated in Section 3.1.2 was used to improve the parameter estimates. In almost all cases an improvement was found, but the improvement was not great, and this is not discussed further.

Many of the graphs presented show a decrease in performance as the measurement segment length increases above about 1λ. This decrease is partly an artefact of the spacing between samples increasing; the inter-sample spacing is the same in the prediction segment as in the measurement segment and so the prediction length is observed as zero if it is less than the inter-sample distance. The decrease is also due to the increase in model order chosen using MDL for longer measurement segments. It may be more appropriate in some situations with long measurement segments to limit the model order to achieve more robust, if not always optimal prediction.

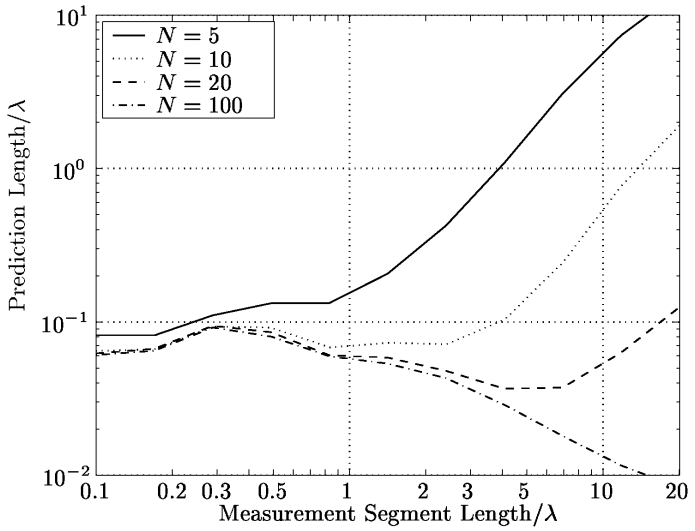


Figure 4.2: Prediction Performance versus Measurement Segment Length using a subspace method (PCLP) for different numbers N of far field sources.

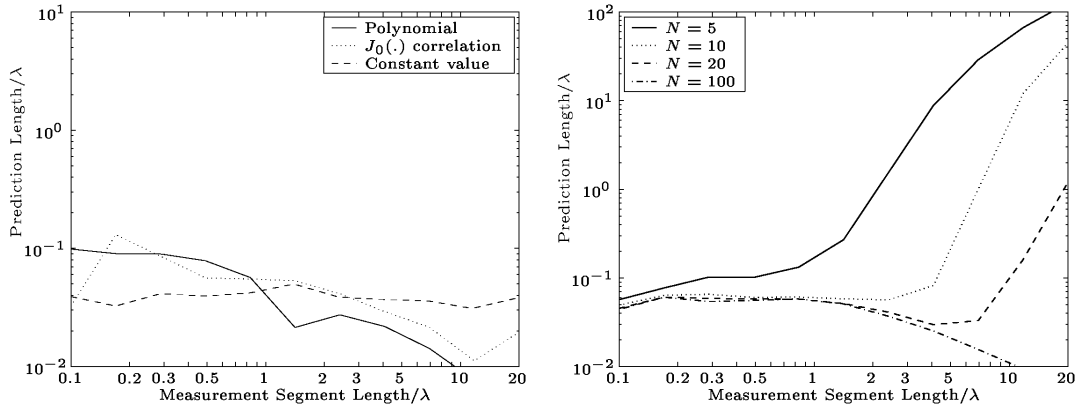
4.2.3 Prediction Length versus Measurement Segment Length

In Fig. 4.2 and Fig. 4.3 the prediction performance is presented as a function of the length (in wavelengths) of the measurement segment. The prediction lengths of interest are those well over 0.1λ and the graphs appear highly variable below this range because of the small distances on the logarithmic scale.

Fig. 4.2 shows the performance of the PCLP algorithm based on the deterministic far field point sources model. Most subspace methods have very similar (though slightly inferior) performance to the PCLP algorithm and so are not shown.

Fig. 4.3 presents the results for some other prediction algorithms — polynomial prediction, linear prediction for the stochastic channel model based on an assumed $J_0(2\pi(x/\lambda))$ correlation function, and the trivial estimator which is to continue the channel at the last measured complex value. It can be seen that none of these methods is capable of reliable prediction to any significant range even if the actual number of sources is very small.

None of the methods achieved prediction beyond about 0.2λ until the length of the measurement segment exceeds 2λ (this is true even when the actual source locations are known perfectly as shown in Fig. 4.3(b)). For longer measurement segments, the subspace methods showed significant prediction performance, the prediction length rising rapidly with increasing measurement segment length (between 1.2 and $1.7th$ power). Where the number of actual sources is very large however, the prediction performance does not increase. A practical limitation for prediction appears to be the number of sources being large, which means that a long measurement segment is required. In practice long measurement segments become impractical because it is



(a) (b)

Figure 4.3: Prediction Performance versus Measurement Segment Length using some other algorithms: (a) Polynomial Prediction, Linear Prediction based on an assumed $J_0(.)$ correlation , and the trivial "constant value" predictor. The number of sources N is 5, but other values of N give very similar results. (b) the Discrete Point Sources model where the locations of the sources are known perfectly, but the amplitudes and phases must be estimated.

unlikely that sources remain unmoving over trajectories of many wavelengths, except perhaps at very high carrier frequencies, where a given trajectory traverses a larger number of wavelengths.

The polynomial and stochastic prediction methods seemed to show no increase in prediction performance as the length of the measurement segment increased. As the number of sources increases one would expect that the assumptions used in deriving the stochastic model (Section 3.3) become progressively more applicable. The performance of *all* the prediction methods then becomes severely limited.

The significant conclusion from this experiment is that *if the number of sources is large, prediction over distances more than a few tenths of a wavelength may not be achievable.* This is true even in Fig. 4.3(b) where the source locations are known perfectly, and it is only the complex amplitudes which must be estimated. It is widely asserted in the literature that the number of *significant* sources in a typical mobile communications environment is small [58, 71], possibly because many of the statistical measures (correlation, probability density etc.) of a real channel can be simulated with only a small number of sources [76], and within a small region it is impossible to distinguish a small number of paths from a large number (see Section 4.7). Though there may be many small sources, a *significant* source in this context is a single source which contributes a large proportion of the total received energy. The experimental justification that these are usually few is rather sparse or may be interpreted as suggesting that there are usually many sources [38, 47]. For the relatively new application of "long range" channel prediction, the number of sources can be seen to be *pivotal.* Simply

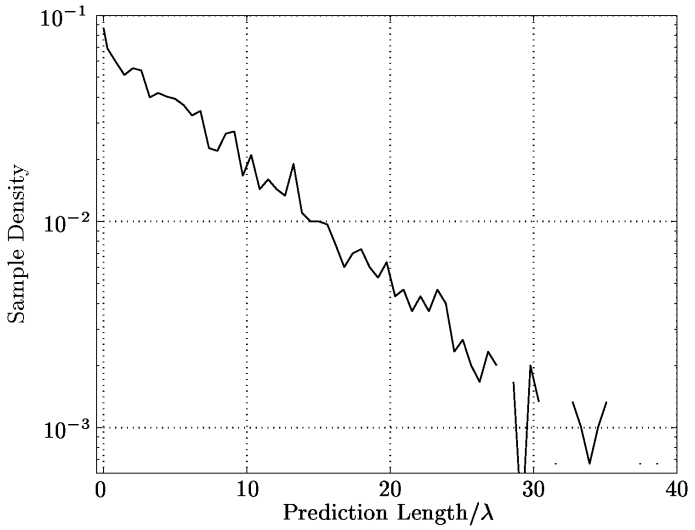


Figure 4.4: An example of the sample probability density function. The Measurement Segment Length is 11.8λ and the number of far field sources $N = 5$.

because a channel model constructed with few sources is realistic in *some* respects, does not imply that it is realistic in *all* respects (see Section 2.3.5). More definitive *bounds* on prediction length are derived in Section 4.4.

The remainder of the results of this section are all based on prediction using the subspace methods of estimating the parameters of the far field point sources model.

4.2.4 Distribution of Prediction Lengths

Fig. 4.4 presents an example distribution of the prediction lengths. In most of the experiments it was found that a large number of scenario samples (typically between 6% and 16%) result in zero successful prediction length. Fig. 4.4 is presented with a logarithmic axis, and appears to suggest that the prediction lengths are exponentially distributed, and this appearance is often found. Although the Kolmogorov Smirnov test [5] rejects the exponential distribution hypothesis, the mean prediction length is achieved or exceeded in about 37% of cases, as an exponential distribution would suggest.

The distribution of prediction lengths is of interest because it influences the usefulness of the prediction information. For instance, a large mean prediction length resulting from a few very "successful" predictions is not very useful if most of the time the prediction length is very short.

4.2.5 Prediction Length versus SNR

Fig. 4.5 presents the prediction performance as a function of SNR for a long measurement segment (10λ) and a short measurement segment ($\lambda/3$). These results show that

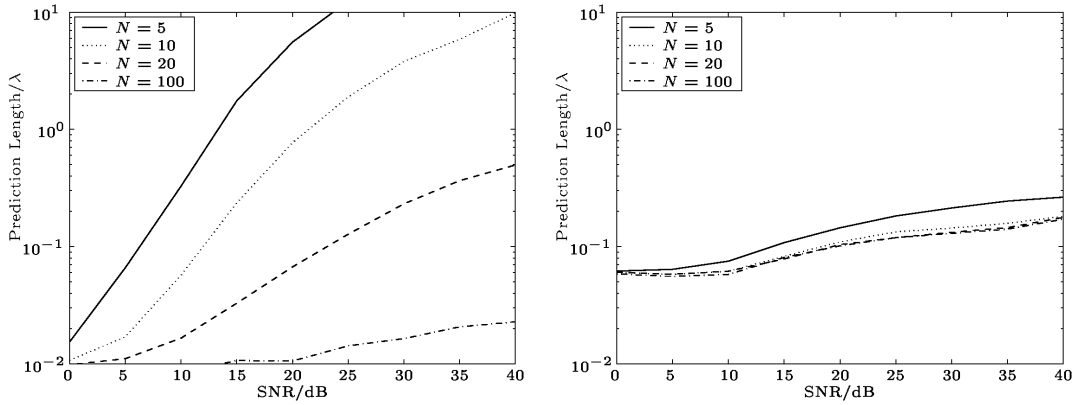


Figure 4.5: Prediction Length versus SNR for (a) a Long Measurement Segment 10λ (b) a Short Measurement Segment($\lambda/3$). The lowest MSE of the subspace algorithms was used for both experiments.

if the measurement segment is long enough for prediction to exceed about 0.2λ, then a significant increase in SNR is accompanied by a significant increase in prediction performance. If the measurement segment is short however, SNR does not improve the prediction significantly (at *very* high SNR, prediction will improve, but such SNR could not be expected in practice). Once again the conclusion is reached that the measurement segment length must be at least several wavelengths for prediction to become feasible.

4.2.6 Prediction Length versus Number of Samples

Fig. 4.6(a) presents the prediction length when the number of measurement samples is varied, while the measurement segment length remains fixed. The SNR *of each sample* is made to vary as the number of sensors varies, so as to keep the overall collected SNR constant, using the expression $\text{SNR} = \text{SNR}_0 - 10\log_{10}(M/M_0)$. The figure shows that once there are sufficient sensing points to ensure sampling every half wavelength, there is no advantage in denser sampling, other than that the overall SNR is increased (though denser sampling *is* an advantage if the sources are in the *near* field of the synthetic array [3] which comprises the measurement segment). This "average" SNR increase supports forming the covariance matrix using the method proposed in Section 3.2, since this allows all the measurement data to be used without increasing the size of the matrix or the order of model used.

4.2.7 Prediction Length versus Model Order

Although in other simulations the model order is selected by using the MDL criterion, graphs of prediction performance versus model order show that under some circum-

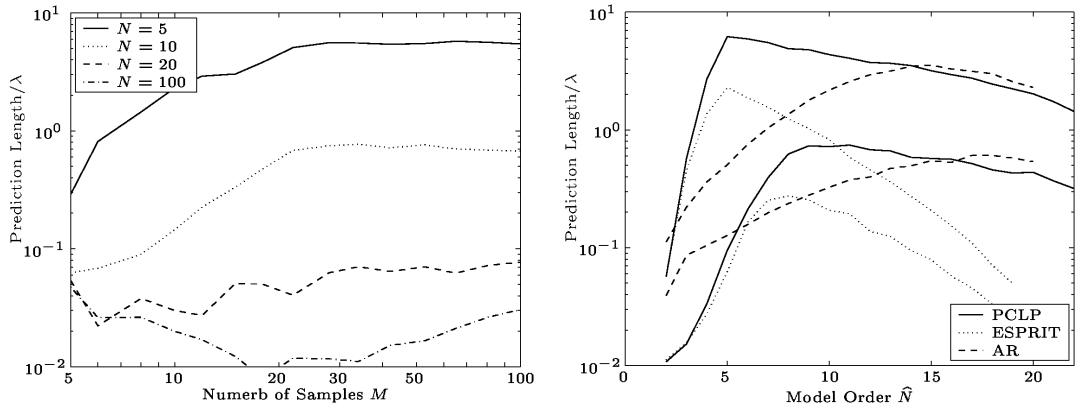


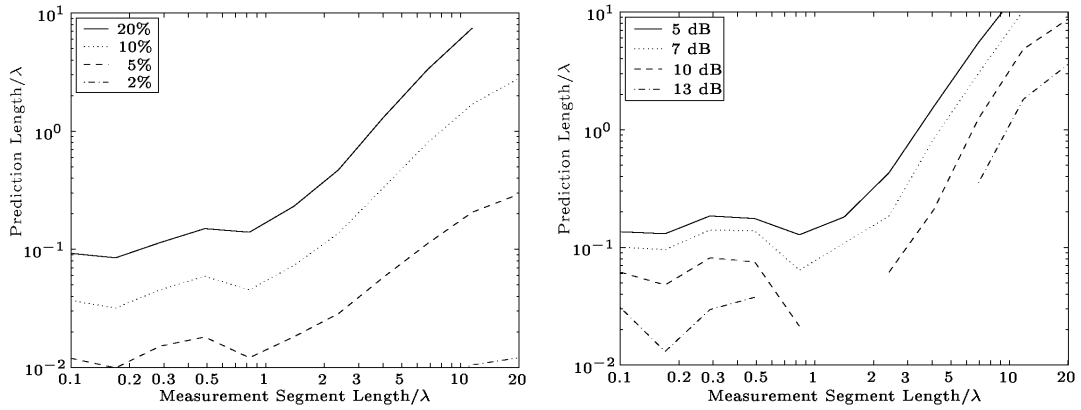
Figure 4.6: (a) Prediction Length versus Number of Samples with the Measurement Segment Length fixed at 10λ. (b) Prediction Performance versus Order of the Model used for Prediction $\widehat{N}$ for PCLP, ESPRIT and Linear Prediction based on an AR model. The upper curve of each type is for $N = 5$, and the lower curve for $N = 10$.

stances, superior prediction can be obtained via a different order selection criterion (Fig. 4.6(b)). Selecting a model with slightly more sources than are actually present in the data does not seriously degrade performance, as perhaps one would expect; modelling some noise as a signal of small amplitude will not change the overall prediction performance. On the other hand, gross over-estimation of the model order can lead to predictions which depart from the actual channel data very rapidly (see Section 4.2.2). The ESPRIT algorithm has greater sensitivity to over-estimation of the model order than the other subspace estimation methods.

Results based on an AR model (see Section 3.3) similar to that used in [46] are also shown in Fig. 4.6(b). In this case, the PCLP method is used, but the Doppler frequencies are not estimated by finding the roots of the prediction filter of (3.38). Instead the filter described by (3.38) is simply used to predict the future values linearly. It can be seen that the performance is very similar to that of the original subspace method, although the optimal model order is not identical. The AR method has the considerable advantage of being relatively simple to convert to an adaptive (Kalman filter) form. This technique has been explored in [41].

4.2.8 Prediction Length versus Performance Threshold

Fig. 4.7(a) shows once again the prediction length as a function of the measurement segment length. The number of sources N is 5 in each case. The different curves are for different values of the threshold used to define the prediction length. As one would expect, if greater accuracy is required, the available prediction length is much smaller. Although noise is added to the data in the measurement segment (SNR of 20 dB), the



(a) (b)

Figure 4.7: Prediction Length presented for various different thresholds used to define the Prediction Length. Greater required accuracy drastically reduces the available prediction length, but the overall trend is very similar.

comparison between the actual and predicted data in the prediction segment is made on the basis of *noiseless* data. Though they are vertically offset, the curves have a similar shape for different performance criteria, which is true of all the results presented in this section.

Fig. 4.7(b) shows the same scenario with the prediction length defined as the distance at which the prediction gain (4.1) falls below some threshold. Once again, the overall trend can be seen to be quite similar, although the curves tend to be more broken since the prediction gain may not achieve the required threshold for *any* range inside the prediction interval. It can be seen then that the conclusions from the simulations are not greatly influenced by the performance threshold chosen.

4.3 Measurements

In this section are presented the results of applying the prediction algorithms to *measured* channel data. The measurements were taken at several frequencies, mostly 1.92 and 5.9 GHz. There were three different locations used — inside a laboratory, outdoors, and inside a large workshop.

A HP8753 network analyser was used to measure the channel attenuation. The transmit antennas used were a Rohde & Schwarz crossed log-periodic antenna and a Huber & Suhner patch PCS antenna. The receiving antenna was a tuned monopole on a circular ground plane of 350mm diameter. The receiving antenna was moved along a 3 metre linear trajectory using a Parker Hann. Corp. Compumotor linear servo. A Mini-Circuits ZHL-1042J or ZHL-1724HLN Low Noise amplifier was used to increase the transmitter and/or receiver power when the SNR was low.

The subspace algorithms were applied to a series of data points measured in the lab-

Figure 4.8: The transmitting antenna outdoors showing the amplifier and the polarisation control unit.

Figure 4.9: The synthetic receiving array outdoors. The monopole is in the centre of a circular ground plane. The network analyser and servo control apparatus are located under the table.

Figure 4.10: The synthetic receiving array in a workshop.

oratory (an irregularly shaped room of dimensions approximately $19 \times 10 \times 3$ metres) at 1.92 GHz (λ=0.156 m). The measured SNR was 40 dB. The segment of data presented in this analysis is typical for measurements made in this and several other indoor and outdoor locations. The channel was measured at 999 points spaced evenly over the 3 metre distance. Averaging of 50 scenarios was obtained by starting the measurement segment at different points in the data set, and by using different but relatively close frequencies. The likelihood of each frequency estimate $\widehat{\boldsymbol{\varpi}}$ was increased by use of a gradient method before estimating the amplitudes $\widehat{\boldsymbol{\zeta}}$.

The results are shown in Fig. 4.11, with upper and lower 95% confidence limits. Prediction does not appear to improve significantly with increasing measurement segment length. There are several likely explanations for this. The first explanation is that the sources are not sufficiently static. The sources may be too close to the receiver for the far field model to be valid, or the sources may themselves move (the former is more likely in this case as the measurements were performed at night with the laboratory unoccupied). The second explanation is that the model is valid, but the number of sources is large. As the simulations show, if the sources are many, prediction is limited even for long measurement segments. It would be expected that if there are many sources, the field becomes equivalent to a diffuse field, the narrow-band channel data becomes effectively a normal process, and prediction is confined to the correlation distance. If the scatterers are uniformly distributed in angle, this correlation distance is approximately $\lambda/5$ (see Fig. 2.2). A third possibility is that rough surface effects (Section 4.5) may be significant.

Fig. 4.12 presents the evolution of spatial frequencies throughout this typical sce-

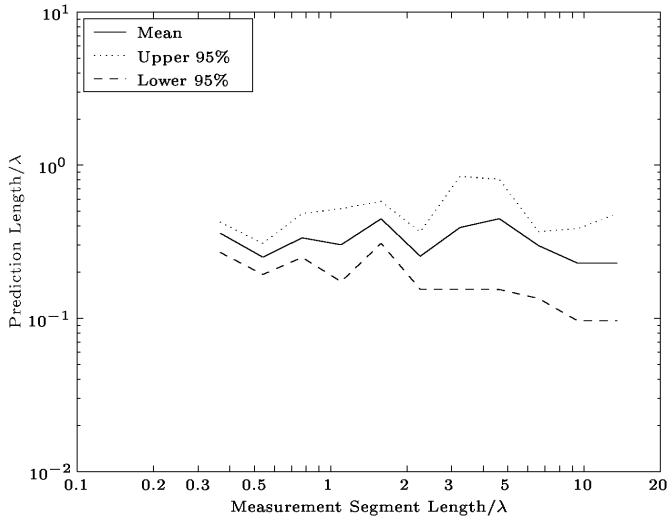


Figure 4.11: Measured prediction range for measurements in laboratory.

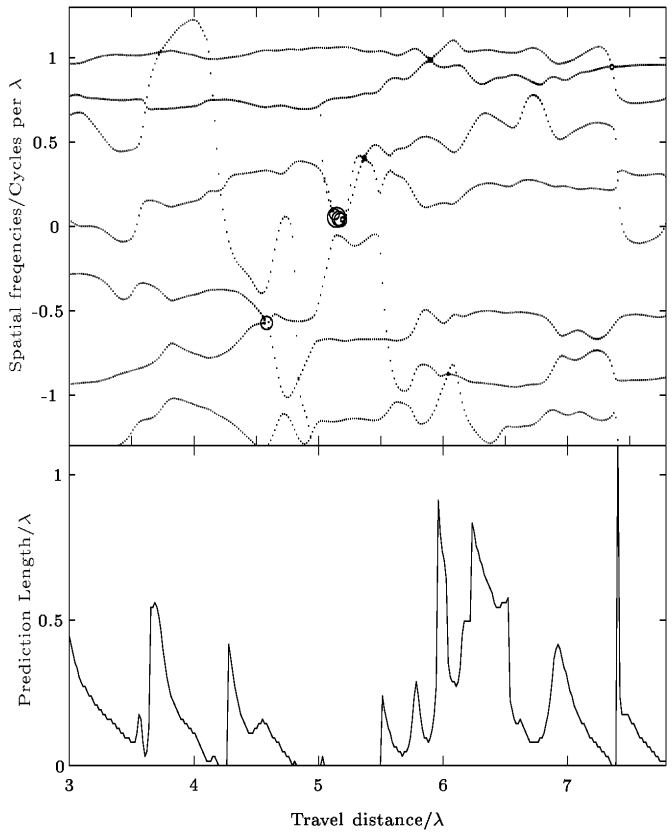


Figure 4.12: Spatial Doppler frequency and Prediction Length for a measured channel.

nario. To produce the graph the number of parameters is fixed with $\widehat{N} = 7$, $P = 8$, $L = 18$, and the prediction segment length is 3 wavelengths. The mean prediction range for this example is 0.185 wavelengths. The symbols representing the spatial frequencies are circles, and the radius of each is proportional to the estimated magnitude of the source. The frequencies outside the range ± 1 cycles per wavelength may result from moving scatterers or from multiple reflections between the receiving antenna assembly and nearby reflectors. The spatial frequencies appear to evolve fairly slowly, and when this is the case, good prediction is achievable. It is interesting to compare the "saw-tooth" appearance to the prediction range in some regions of the lower plot with the areas of rapid change in frequency in the upper plot. Whenever there is a rapid change which cannot be anticipated by the previous data, prediction beyond the point of rapid change cannot be achieved.

The results for the measurements made in the workshop and outdoors are not shown here. However the following observations are made. Most of the measurements made in the workshop were at 5.9GHz. The higher frequency has the effect of "contracting" the near field region of the synthetic array, thus making the far field assumption more valid. However, at this higher frequency rough surface scattering effects are much more significant (see Section 4.5). Typical prediction lengths were observed to be lower than those presented above.

The outdoors measurements unfortunately suffered from significant amounts of noise, even when amplifiers were used, because of the low powers and large distances involved. The power loss was both propagation loss and cable loss. The latter could be overcome by the use of rubidium clocks. These were not available however. From the data which was available it appears that the prediction range is quite limited. Since there were many trees in the vicinity, rough surface effects may once again be significant.

The measurements used for [6] (one of the very first implementations of real-time prediction) were obtained. It was found that it was possible to reproduce the results of that paper only by performing an interpolation step as described by the authors. This interpolation however appears to produce misleading results. The interpolation filter actually incorporates information from future values into each of the interpolated values. The prediction filter estimated from the data is actually recovering the parameters of the interpolation filter, rather than parameters of the original measurements. Since the interpolation filter parameters do not change, good prediction is obtained. This however does not indicate reliable prediction of the original channel.

4.4 Cramer Rao Bounds

In this section a bound is derived for the variance of the estimate of a predicted channel based on the estimation of the component sinusoids. This is done by first finding the

Fisher information matrix for the model parameters, and thus a bound on the variance of their estimates. This variance is then used to calculate a bound on the actual channel estimate.

The model of far field narrowband sources is equivalent to complex sinusoids in noise. The parameters to be estimated are the amplitudes ς_n and phases ψ_n (or the complex amplitudes $\zeta_n = \varsigma_n e^{j\psi_n}$) and the frequencies ϖ_n of N complex sinusoids in noise from M regularly spaced samples.

The measured samples are given by

$$r[m] = \sum_{n=1}^{N} \varsigma_n e^{j\psi_n} e^{j\varpi_n m} + \eta[m], \quad m = 0, \ldots, M-1, \tag{4.2}$$

where $\eta[m]$ is a white complex normal noise with variance σ_η^2. Let $\mathbf{m} = (0, 1, \ldots, M-1)^T$ and $\mathbf{r} = (r_0, r_1, \ldots, r_{M-1})$. The signal $\mathbf{r}$ is thus a complex normal random vector with mean

$$\boldsymbol{\mu} = \sum_{n=1}^{N} \mathbf{r}_n = \sum_{n=1}^{N} \varsigma_n e^{j\psi_n} e^{j\varpi_n \mathbf{m}} \tag{4.3}$$

and variance $\mathbf{C} = \sigma_\eta^2 \mathbf{I}$. In this section as in Section 3.5, $e^{j\cdot}$ where $\cdot$ is a matrix or vector means exponentiation of the *elements* of a vector or matrix, rather than the matrix exponential. The parameter vector $\boldsymbol{\xi} = (\varsigma_1, \psi_1, \varpi_1, \ldots, \varsigma_N, \psi_N, \varpi_N)^T$. Using Bangs' formula [9, 80, 141], the elements of the Fisher information matrix are given by

$$\begin{aligned}
[\mathbf{J}(\boldsymbol{\xi})]_{ij} &= \operatorname{tr}\left[\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \xi_i}\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \xi_j}\right] + 2\operatorname{Re}\left[\frac{\partial \boldsymbol{\mu}^H}{\partial \xi_i}\mathbf{C}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \xi_j}\right] \\
&= \frac{1}{\sigma_\eta^2} 2\operatorname{Re}\left[\frac{\partial \boldsymbol{\mu}^H}{\partial \xi_i}\frac{\partial \boldsymbol{\mu}}{\partial \xi_j}\right].
\end{aligned} \tag{4.4}$$

Differentiating with respect to each of the parameters,

$$\begin{aligned}
\frac{\partial \boldsymbol{\mu}}{\partial \varsigma_n} &= \frac{\mathbf{r}_n}{\varsigma_n} \\
\frac{\partial \boldsymbol{\mu}}{\partial \psi_n} &= j\mathbf{r}_n \\
\frac{\partial \boldsymbol{\mu}}{\partial \varpi_n} &= j\mathbf{m} \odot \mathbf{r}_n.
\end{aligned} \tag{4.5}$$

Thus the Fisher information matrix consists of 3×3 blocks of the form:

$$\frac{2}{\sigma_\eta^2} \operatorname{Re} \left\{ e^{j(\psi_{n_1} - \psi_{n_2})} \begin{bmatrix} 1 & j\varsigma_{n_2} & j\varsigma_{n_2} \\ -j\varsigma_{n_1}^* & \varsigma_{n_1}^* \varsigma_{n_2} & \varsigma_{n_1}^* \varsigma_{n_2} \\ -j\varsigma_{n_1}^* & \varsigma_{n_1}^* \varsigma_{n_2} & \varsigma_{n_1}^* \varsigma_{n_2} \end{bmatrix} \odot \sum_m \left(e^{jm(\varpi_{n_2} - \varpi_{n_1})} \begin{bmatrix} 1 & 1 & m \\ 1 & 1 & m \\ m & m & m^2 \end{bmatrix} \right) \right\} \tag{4.6}$$

where n_1 and n_2 are the row and column indices for the 3×3 blocks. This expression simplifies for the diagonal blocks (where $n_1 = n_2 = n$) to

$$\frac{2}{\sigma_\eta^2} \begin{bmatrix} M & 0 & 0 \\ 0 & \varsigma_n^2 M & \varsigma_n^2 \sum_m m \\ 0 & \varsigma_n^2 \sum_m m & \varsigma_n^2 \sum_m m^2 \end{bmatrix}. \tag{4.7}$$

If m were to range from $(1 - M)/2$ to $(M - 1)/2$ this further simplifies to

$$\frac{2}{\sigma_\eta^2} \operatorname{diag} \left(M, \varsigma_n^2 M, \varsigma_n^2 \sum_m m^2 \right) = \frac{2}{\sigma_\eta^2} \operatorname{diag} \left(M, \varsigma_n^2 M, \varsigma_n^2 \frac{(M-1)M(M+1)}{12} \right) \tag{4.8}$$

(and when $N = 1$, the entire $\mathbf{J}$ is given by equation 4.8).

Suppose now that each time the estimation is performed, the true parameter values being estimated are different, drawn randomly from the following distributions. The amplitudes ς_n are iid Rayleigh variables and ψ_n are iid uniform over $(-\pi, \pi]$ (this means that the *complex* amplitude $\zeta_n = \varsigma_n e^{j\psi_n}$ is zero-mean complex normal with variance of say σ_ς^2). The frequencies $\varpi_n = (2\pi\Delta_x \sin\theta_n)/\lambda$, where θ_n are iid uniform over $(-\pi/2, \pi/2]$ and $0 < 2\pi\Delta_x/\lambda < \pi$.

The prediction error for sample m is

$$|\epsilon[m]|^2 = \left| \sum_{n=1}^{N} s_n e^{j\psi_n} e^{j\varpi_n m} - \sum_{n=1}^{N} \hat{s}_n e^{j\hat{\psi}_n} e^{j\hat{\varpi}_n m} \right|^2. \tag{4.9}$$

Taking a first order approximation for $\epsilon[m]$ (this idea is from [10])

$$\begin{aligned} \epsilon[m] &\approx \sum_{n=1}^{N} \frac{\partial r[m]}{\partial \varsigma_n} \epsilon_{\varsigma_n} + \frac{\partial r[m]}{\partial \psi_n} \epsilon_{\psi_n} + \frac{\partial r[m]}{\partial \varpi_n} \epsilon_{\varpi_n} \\ &= \sum_{n=1}^{N} \varsigma_n e^{j\psi_n} e^{j\varpi_n m} \left[\frac{\epsilon_{\varsigma_n}}{s_n} + j\epsilon_{\psi_n} + jm\epsilon_{\varpi_n} \right] \end{aligned} \tag{4.10}$$

so

$$|\epsilon[m]|^2 = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} \varsigma_{n_1}\varsigma_{n_2} e^{j(\psi_{n_1}-\psi_{n_2})} e^{jm(\varpi_{n_1}-\varpi_{n_2})} \left(\frac{\epsilon_{\varsigma_{n_1}}}{\varsigma_{n_1}} + j\epsilon_{\psi_{n_1}} + jm\epsilon_{\varpi_n}\right)\left(\frac{\epsilon_{\varsigma_{n_2}}}{\varsigma_{n_2}} - j\epsilon_{\psi_{n_2}} - jm\epsilon_{\varpi_{n_2}}\right). \tag{4.11}$$

Taking the expectation, one obtains

$$E\left(|\epsilon[m]|^2\right) = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} \varsigma_{n_1}\varsigma_{n_2} e^{j(\psi_{n_1}-\psi_{n_2})} e^{jm(\varpi_{n_1}-\varpi_{n_2})} h_{n_1 n_2}, \tag{4.12}$$

where $h_{n_1 n_2}$ is the sum of the elements of the matrix

$$\begin{bmatrix} \frac{1}{\varsigma_{n_1}\varsigma_{n_2}} & \frac{-j}{\varsigma_{n_1}} & \frac{-jm}{\varsigma_{n_2}} \\ \frac{j}{\varsigma_{n_2}} & 1 & m \\ \frac{jm}{\varsigma_{n_2}} & m & m^2 \end{bmatrix} \odot \left(\mathbf{J}^{-1}\right)_{n_1 n_2} \tag{4.13}$$

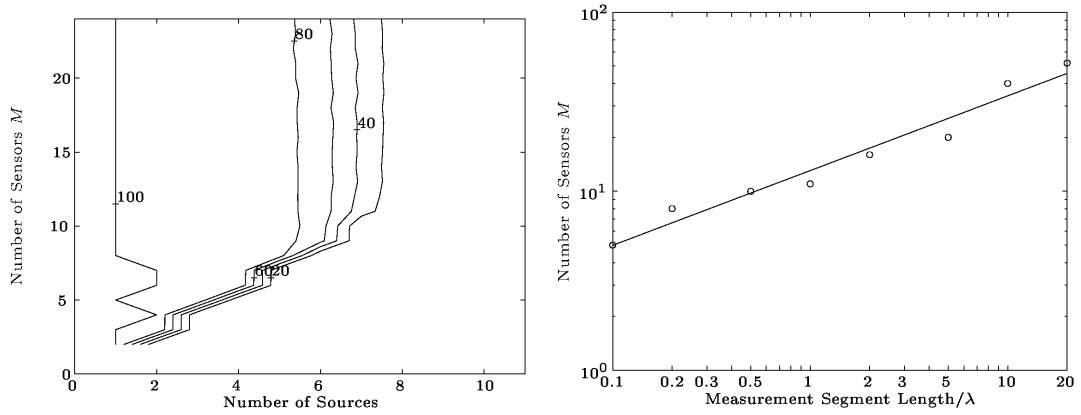
and $\left(\mathbf{J}^{-1}\right)_{n_1 n_2}$ is the 3×3 block of the Fisher inverse with n_1 and n_2 being the row and column indices of the 3×3 blocks in this inverse. This assumes of course that the bound is nearly achieved so that $E\{(\xi - \widehat{\xi})(\xi - \widehat{\xi})^H\} - \mathbf{J}^{-1}$ is not merely a positive semi-definite matrix, but close to the zero matrix.

For a large prediction range (large m), the most critical parameter is the frequency ϖ, since the variance of this estimate is multiplied in (4.13) by m^2 to calculate the overall prediction error.

4.4.1 Invertibility of the Fisher Information Matrix

Numerical calculation of the bound described above leads to the discovery that a large proportion of the Fisher information matrices are very poorly conditioned. It is conjectured in [123] that the Fisher information matrix is singular only if two or more of the tone frequencies are equal, modulo 2π, assuming M is large enough. In any estimation problem, the Fisher information matrix being singular indicates that there is some redundancy in the parameterisation — in this case two sinusoids having the same frequency. Conditions affecting the likelihood of the Fisher information matrix being close to singular are the number of sources present N, the number of elements in the virtual array M, and the length of the virtual array $(M-1)\Delta_x$. The symbol p is now defined as the proportion of well conditioned Fisher information matrices (given the distributions of the parameters described above).

It was found that provided the number of samples M is sufficient, M has no further effect on p. Contours of p expressed as a percentage are shown in Fig. 4.13(a), for $(M-1)\Delta_x = 2\lambda$.



(a) (b)

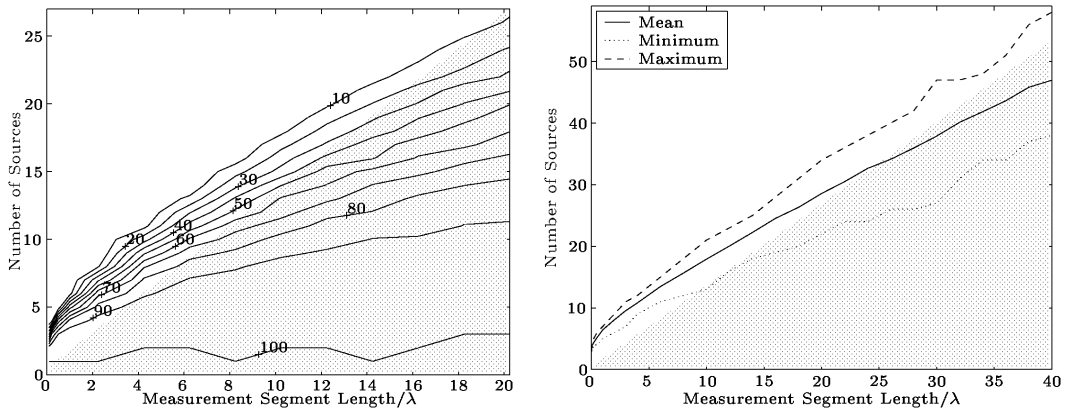
Figure 4.13: (a) Contours of the percentage of well conditioned Fisher information matrices for a measurement segment length of 2λ, showing the effect of the number of sensors in the array. (b) The number of sensors M required to achieve independence from M for different measurement segment lengths. The least squares line gives $M > 13x^{0.4}$ where x is the measurement segment length expressed in wavelengths. The criterion for "well conditioned" is that the LINPACK reciprocal condition estimate [168] be larger than 10^{-15}.

A similar graph was produced for various values of the measurement segment length, and the required number of samples M required to reach independence of M was noted. The results of this are shown in Fig. 4.13(b).

After ensuring that the number of samples was not a limiting factor, the effect of the measurement segment length was investigated, and this is shown in Fig. 4.14(a). The shaded region represents the region where the number of real parameters being estimated is less than the number of real data values supplied, assuming that there are two independent real values per half wavelength of measurement segment.

The graphs in Fig. 4.13 and Fig. 4.14 clearly indicate that if the number of parameters is too large, the Fisher information matrix is singular, and the variance of the parameter estimates is unbounded. In an actual system however, the number of sources being modelled would actually be smaller. To evaluate a realistic bound then, the number of parameters being estimated must be kept to about the same as would be the case in a real system. Hence the parameters of the largest N sources out of N' are estimated. If there are more sources, the remaining $N' - N$ are treated as noise.

The problem now is to evaluate the effective "noise power" of these remaining sources. If the complex amplitudes $\zeta_n = \varsigma_n e^{j\psi_n}$ are zero-mean complex normally distributed, then so also will be $\zeta_n e^{jm\varpi_n}$. The power of all the sources is normalised to 1. The power of un-modelled sources is the variance of the sum of the N smallest out of N' iid zero-mean, complex normal variables. If for each of these variables $z_n = x_n + jy_n$,



(a) (b)

Figure 4.14: (a) Contours of the percentage of Fisher information matrices showing the effect of variation of the measurement segment length. (b) The value to which the number of sources must be reduced by combining nearby sources in order to obtain well conditioned Fisher information matrices.

and $z = \sum_n z_n$,

$$\text{var}(z) = E\left(\sum_n z_n \sum_n z_n^*\right) \tag{4.14}$$

$$= E\left(\sum_{n_1}\sum_{n_2} z_{n_1} z_{n_2}^*\right) \tag{4.15}$$

$$= \sum_n \left(x_n^2 + y_n^2\right) \tag{4.16}$$

which is simply the mean of the sum of *exponentially* distributed variables, each with pdf $f_a(a) = N' e^{-N'a} U(a)$. The cdf of each is thus $F_a(a) = (1 - e^{-N'a})U(a)$ ($U(a)$ denotes the unit step function).

The density of the n-th smallest of a set of such random variables, (the n-th order statistic a_n) is given by [81, 107]

$$f_{a_n}(a) = \frac{N'!}{(n-1)!(N'-n)!}(F_a(a))^{n-1}(1 - F_a(a))^{N'-n} f_a(a)$$

$$= \frac{N'!N'}{(n-1)!(N'-n)!}(1 - e^{-N'a})^{n-1}(e^{-N'a})^{N'-n+1}U(a). \tag{4.17}$$

The mean of this distribution can be calculated as

$$E(a_n) = \int_0^\infty \frac{N'!N'}{(N'-n)!(n-1)!}(1 - e^{-N'a})^{n-1}(e^{-N'a})^{N'-n+1}\, a\, da, \tag{4.18}$$

which can be arranged using a binomial expansion of $1 - e^{-N'a}$, and (3.381-4) of [64]

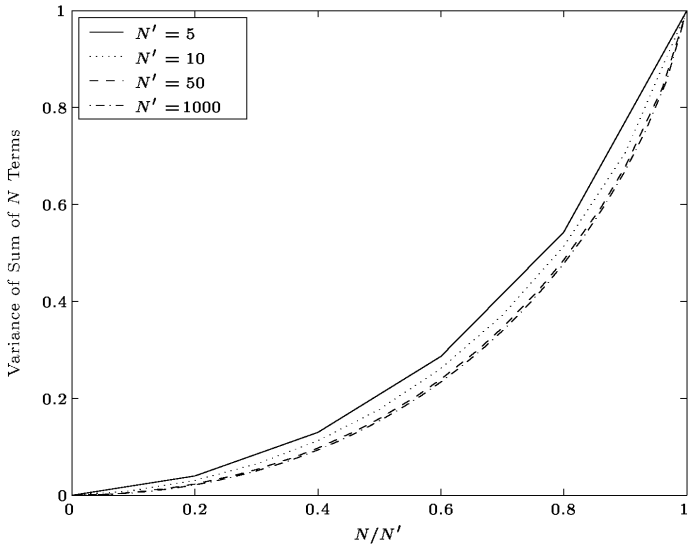


Figure 4.15: The variance of the sum of the N smallest out of N' zero mean complex normal iid random variables, where the variance of the sum of N' is 1.

as

$$E(a_n) = \frac{(N'-1)!}{(N'-n)!}\sum_{i=0}^{n-1}\frac{1}{i!(n-1-i)!}(-1)^i\int_0^\infty (e^{-u})^{N'-n+1+i}\,u\,du$$
$$= \frac{(N'-1)!}{(N'-n)!}\sum_{i=1}^{n}\frac{(-1)^{i-1}}{(i-1)!(n-i)!}\frac{1}{(N'-n-i)^2}. \tag{4.19}$$

The mean of the sum of N of these variables (being the sum of the means) is

$$E\left(\sum_{n=1}^{N} a_n\right) = \sum_{n=1}^{N}\sum_{i=1}^{n}\frac{(N'-1)!(-1)^{i-1}}{(N'-k)!(i-1)!(k-i)!(N'-k-i)^2}. \tag{4.20}$$

Although this expression is exact, in practice it is not usable for large N' because the terms being added become extremely large, while their sum is in the range $[0,1]$. It was found that quadrature of the expression (obtained by the substitution $v = e^{-N'a}$ into equation 4.18 and then summation over n)

$$\int_0^1 \log v \sum_{n=1}^{N}\frac{(N'-1)!}{(n-1)!(N'-n)!}(1-v)^{n-1}v^{N'-n}\,dv \tag{4.21}$$

is much more robust (the integrand is well behaved, even near $v = 0$) and can be rapidly computed. It appears in fact that the variance of this sum is relatively insensitive to (though not independent of) the actual value of N' provided the ratio N/N' is kept constant, as shown in Fig. 4.15.

The power of the sources not parameterised has two effects. First it decreases the

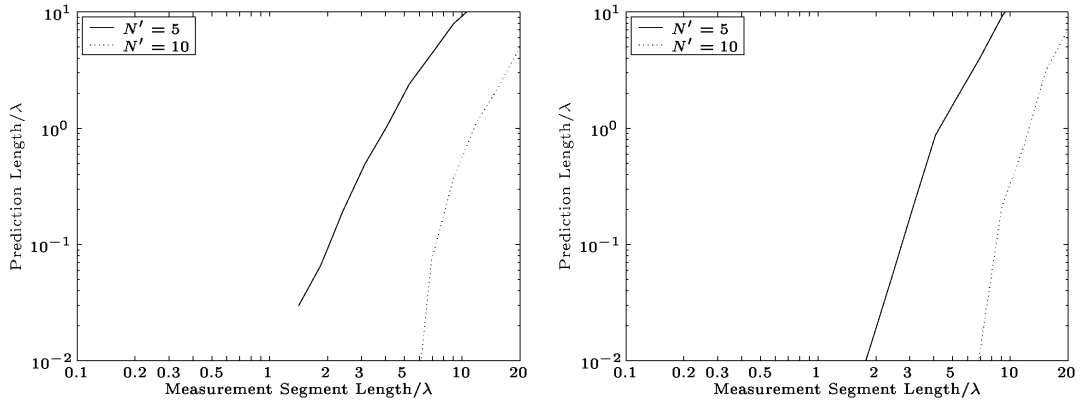


Figure 4.16: A prediction length bound derived from the CRLB on prediction error as a function of Measurement Segment Length for various values of the number N' of sources present. In (a) the number of sources has been limited so that 50% of all scenarios have well conditioned Fisher information matrices. In (b) sources very near other sources have been combined.

overall SNR, and secondly it causes an increase in the error between the predicted and actual signal.

If $\epsilon_1[m]$ is the error that would be expected in the case where $N = N'$ and for unity SNR, and Q is the power of the "ignored" sources where the power of all sources is also unity, the expected error now becomes

$$\begin{aligned} E\left(|\epsilon[m]|^2\right) &= E\left(\left|\sum_{n=1}^{N} \varsigma_n e^{j\psi_n} e^{j\varpi_n m} - \sum_{n=1}^{N} \hat{\varsigma}_n e^{j\hat{\psi}_n} e^{j\hat{\varpi}_n m} + \sum_{n=N}^{N'} \varsigma_n e^{j\psi_n} e^{j\varpi_n m}\right|^2\right) \\ &= E\left(\left|\sum_{n=1}^{N} \varsigma_n e^{j\psi_n} e^{j\varpi_n m} - \sum_{n=1}^{N} \hat{\varsigma}_n e^{j\hat{\psi}_n} e^{j\hat{\varpi}_n m}\right|^2\right) + E\left(\left|\sum_{n=N}^{N'} \varsigma_n e^{j\psi_n} e^{j\varpi_n m}\right|^2\right) \\ &= (\sigma^2 + Q) E\left\{|\epsilon_1[m]|^2\right\} + Q. \end{aligned} \tag{4.22}$$

In the following results N is constrained so that for any given value of the measurement segment length, at least 50% of the Fisher information matrices are invertible, and the remaining $N' - N$ sources are treated as noise in the manner described above.

Recall that in Section 4.1 the performance criterion chosen is the mean of the distances at which the error magnitude first exceeds 20% of the RMS signal level. The criterion used here is a threshold on the expected error or *prediction gain*. It was shown in Section 4.2.8 that these criteria lead to similar if not identical results provided the thresholds are chosen appropriately. The prediction length here is defined as the distance at which the expected error magnitude squared first exceeds 0.04 of the RMS signal level. The number of array elements is $M = 51$ and the SNR is 20dB. The results

are shown in Fig. 4.16(a).

An alternative approach to the problem of singular Fisher information matrices is to combine into a single source any pairs of sources which are located near each other. In the form of the Cramer Rao bound presented so far, co-located sources cause an ambiguity in estimation of the parameters. In a real prediction scenario however, co-located sources can be simply combined into a single source. A trial was conducted to utilise this fact, commencing with a large number of sources. The sources were successively combined on the basis of their proximity (measured as the difference in Doppler frequency) until the Fisher information matrix became well conditioned. The results from using this procedure are presented in Fig. 4.14(a). The shaded region of this figure is the region where the number of real parameters being estimated is less than or equal to the number of real independent data points available. The average curve is quite similar to the edge of the shaded region, which indicates that the approximation based on the number of independent real parameters is a good one. The bound found using this procedure is shown in Fig. 4.16(b). It can be seen by comparison of (a) and (b) of Fig. 4.16 that the two approaches for obtaining a bound produce very similar results.

In can be concluded then, that despite the difficulties presented in deriving a well conditioned Fisher information matrix, the Cramer Rao bound provides a means of obtaining a bound on the performance of prediction based on a narrowband far field point sources model. It appears immediately that if the number of sources is large, the range of reliable prediction is practically negligible. The results of the simulations presented in Section 4.2 (in particular Fig. 4.2) show close agreement with these bounds. Other conclusions are discussed in that section.

4.5 Rough Surface Scattering

This section presents some results on the influence that rough surface scattering may be expected to play on the predictability of a channel.

A model used for prediction of a mobile channel in this thesis and in other works [40, 140, 151] is that of discrete far field point sources. Prediction based on extrapolation of the complex sinusoids corresponding to each of these point sources may be shown (via the Wold decomposition [106], see Section 2.6.3) to be functionally equivalent to linear prediction, except that the roots of the prediction polynomial are fixed at unity magnitude so the sinusoids have fixed amplitude rather than being allowed to decay or grow.

It is of interest to know to just what extent the mobile channel can be predicted, and what are the fundamental limitations on prediction range. It has been shown in sections 4.2.3 and 4.4 that one of the fundamental limitations results from the number of scatterers. If there are many scatterers, the useful prediction range is seriously limited.

Another limitation may arise for models which assume the scatterers are discrete points. Such models assume that each point reflecting the signal to the receiver is either small enough to behave as a point, or large and smooth enough to behave as a reflector of a point. In practice, many objects fall in between these extremes, and produce a non-isotropic reflection. For the surface to be large enough to behave as a reflector it must have dimensions at least the size of the first Fresnel zone (see Section 2.7.1).

Even if a surface is *large* enough to be considered as a reflector, it may not be *smooth* enough. A criterion commonly used for determining whether the surface may be considered smooth is the Rayleigh criterion — that waves reflected from the upper and lower and lower extremes of a surface differ in phase by less than $\pi/2$ [12] (some authors insist the maximum allowed phase difference is only $\pi/8$). This criterion may be expressed as $\mathcal{J} < \pi/2$, where

$$\mathcal{J} \triangleq 4\pi \frac{\sigma_\xi}{\lambda} \sin\psi \tag{4.23}$$

is called the Rayleigh parameter, λ is the radiation wavelength, ψ is the grazing angle of the reflection and σ_ξ is the depth of the surface. In many models of rough surfaces the height of the roughness is considered a normal process, with the height standard deviation equal to σ_ξ.

In this section it is shown that even surfaces classified as smooth according to the Rayleigh criterion produce diffuse radiation components large enough to be a limiting factor in channel prediction.

4.5.1 Scattering formulation

Rough surface scattering is a subject of investigation characterised by the need for many simplifying assumptions in order to produce useful results. However, the assumptions made here are justifiable in many situations. The details of the small perturbation model used in this section may be found in [11] and the reader is referred to this work for details. The barest essentials of the method are presented here. An outline and references of other models may be found in [82].

The surface $\mathcal{S}$ is assumed horizontal, and the height is modelled as a zero mean random process $z = \xi(\boldsymbol{r})$, where $\boldsymbol{r}$ is a vector lying in the plane $z = 0$.

The total field reflected by the surface may be considered to be the sum of a specular and a diffuse component. Define σ_r to be the correlation distance of the surface, and $k = 2\pi/\lambda$ to be the wavenumber. For large scale fluctuations in the surface ($k\sigma_r \gg 1$), the ratio of the power of the diffuse component to the power of the specular component is $\mathcal{J}^2$ (from (10.45) of [11]).

For a random surface, the diffuse component of the field is essentially a random process. Of interest here is the correlation of this component at spatially separated points $\boldsymbol{R}$ and $\boldsymbol{R}'$ (refer to Fig. 4.17).

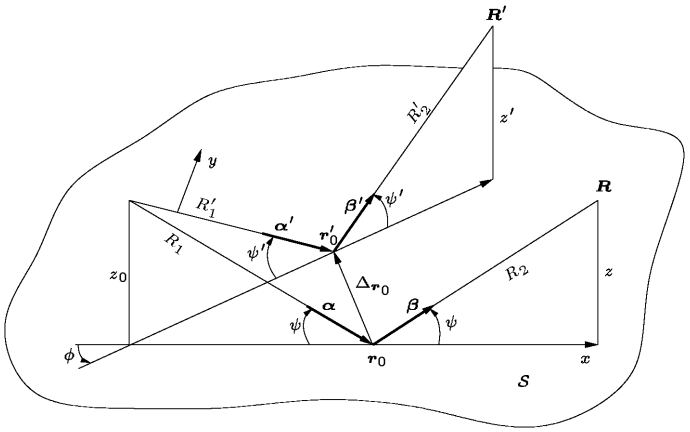


Figure 4.17: Geometry for rough surface scattering to two locations. $\mathcal{S}$ is a horizontal surface, the source is located at point $(0, 0, z_0)$, and measured at points $\boldsymbol{R} = (0, 0, z)$ and $\boldsymbol{R}' = (0, 0, z')$. The unit vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ point in the directions of incidence and specular reflection respectively. The points of specular reflection are separated by $\Delta_{\boldsymbol{r}_0}$

The field scattered by a rough surface may be calculated by using the boundary condition for the electric field $\boldsymbol{E}$ on the surface, which is assumed to be perfectly conducting:

$$(\boldsymbol{N} \times \boldsymbol{E})_{z=\xi(\boldsymbol{r})} = 0, \tag{4.24}$$

where $\boldsymbol{N}$ is the unit vector normal to the surface $z = \xi(\boldsymbol{r})$. The tangential field components fully define the field in all space. If the deviation of the surface from the mean plane is small, this boundary condition may be transferred to the plane $z = 0$, and using perturbation theory, the diffuse component of the field results from the perturbation of the surface. The field may be integrated over the surface $\mathcal{S}$ to obtain the diffuse scattered field as

$$u(\boldsymbol{R}) = \frac{k^2}{\pi} \int_{\mathcal{S}} \frac{e^{jkR_1R_2}}{R_1R_2} \hat{\mathcal{F}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \xi(\boldsymbol{r}) d\boldsymbol{r}, \tag{4.25}$$

where $\hat{\mathcal{F}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is a slowly varying function of the unit vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (not detailed here) and the geometry terms are as illustrated in Fig. 4.17.

If the surface irregularities possess sufficiently large linear (horizontal) dimensions that only a small region of integration is essential near the point of specular reflection $\boldsymbol{r}_0$, so that $k\sigma_r \gg 1$, the argument of the exponential may be expanded as a power series, giving the diffuse component of the field as

$$u(\boldsymbol{R}) = \frac{vk^2 \sin^2\psi}{\pi R_1 R_2} e^{jkR_1R_2} \int_{\mathcal{S}} e^{j\frac{k}{R}(x^2 \sin^2\psi + y^2)} \xi(\boldsymbol{r}_0 + \boldsymbol{r}) d\boldsymbol{r}, \tag{4.26}$$

where v is a constant multiplier which depends on the polarisation of the source antenna

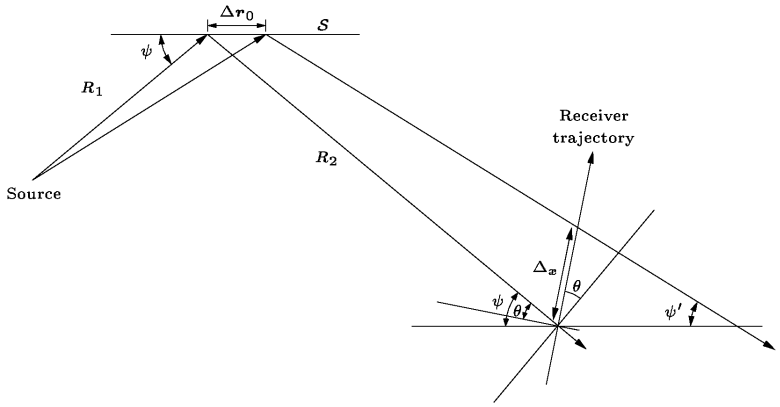


Figure 4.18: Geometry of moving receiver.

and the receiver assembly, and $R = 2R_1R_2/(R_1 + R_2)$. When the assumption (used in this section) that $k\sigma_r \gg 1$ is valid, polarisation does not have a significant impact on the correlation.

The height function $\xi(\boldsymbol{r})$ is random, and hence the field $u(\boldsymbol{R})$ is also random. However, the correlation

$$\mathcal{K}(\boldsymbol{R}, \boldsymbol{R}') = \frac{E\{u(\boldsymbol{R})u^*(\boldsymbol{R}')\}}{E\{|u(\boldsymbol{R})|^2\}} \tag{4.27}$$

can be expressed in terms of the normalised correlation of the surface, which is assumed here to be homogeneous (hence stationary) and isotropic,

$$W(\boldsymbol{\rho}) = \frac{E\{\xi(\boldsymbol{r}_0)\xi(\boldsymbol{r}_0 + \boldsymbol{\rho})\}}{E\{|\xi(\boldsymbol{r}_0)|^2\}}. \tag{4.28}$$

The correlation function of the diffuse field is given by a rather complicated expression. However, for small separations, non-grazing propagation, and assuming the integration is over an area at least as large as the first Fresnel zone, it can be shown (see (16.9) and (16.21) of [11]) that

$$\mathcal{K}(\boldsymbol{R}, \boldsymbol{R}') \approx e^{-jk\Delta_{R_0}} W(\Delta_{\boldsymbol{r}_0}), \tag{4.29}$$

where $\Delta_{\boldsymbol{r}_0}$ is the separation of the points of specular reflection $\boldsymbol{r}_0$ and $\boldsymbol{r}_0'$, and Δ_{R_0} is the difference in the total path length between the source and the two receiving points $\boldsymbol{R}$ and $\boldsymbol{R}'$ (see Fig. 4.17).

4.5.2 Moving Receiver

Consider a receiver moving through the field scattered by a rough surface. The spatial correlation between two points on its trajectory will depend on both the grazing angle

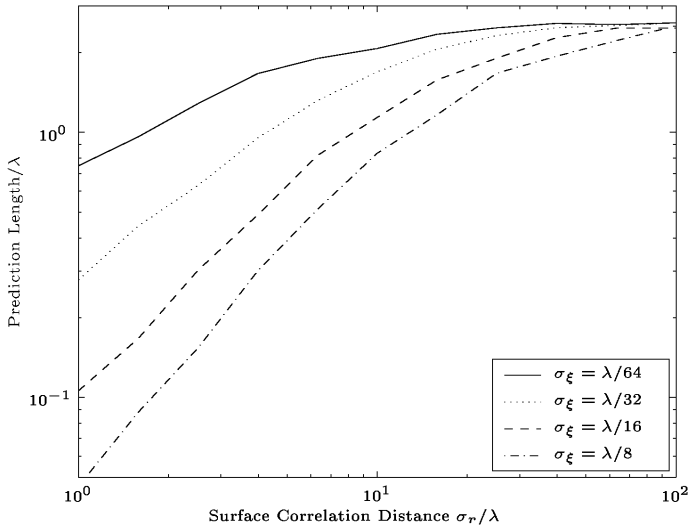


Figure 4.19: The influence of rough surface correlation distance on predictability of the channel, for various values of roughness. In this scenario, the number of scattering surfaces $N = 5$, the received SNR is 20dB, the number of measurements $M = 40$, the trajectory length is 6λ.

of the scattering, and the angle between the scattered radiation and the movement of the receiver. The geometry is shown in Fig. 4.18. The receiver is assumed to be far from the scatterer so that $\psi \approx \psi'$. It can be shown that the distance between the points of specular reflection when the receiver moves distance Δ_x is given by

$$\Delta_{\boldsymbol{r}_0} \approx \Delta_x \frac{\cos\theta}{\sin\psi} \frac{R_1}{R_1 + R_2}. \tag{4.30}$$

It can be seen from (4.30) then, that for small grazing angles ψ the distance between the points of specular reflection is large, and so the correlation is small. However, the amplitude of the diffuse component, which is proportional to $\mathcal{J}$ is also small, so that on the whole as ψ decreases the diffuse field provides less of an impediment to channel prediction.

4.5.3 Numerical Study

In this section the effect of rough surface scattering is investigated for an example scenario. A narrowband signal was reflected by $N = 5$ scatterers, and received with SNR of 20dB at points on a straight mobile trajectory with consecutive measurement points being $\Delta_x = 0.15\lambda$ apart. The surface correlation function was assumed to have the form $W(\boldsymbol{\rho}) = e^{-|\boldsymbol{\rho}|^2/\sigma_r^2}$. Each scatterer had a magnitude $|\zeta_n|$ (ζ_n are complex normal) corresponding to the specular component, which was constant throughout the trajectory, and a diffuse component u_{mn} which was also complex normal, and had magnitude $|\zeta_n \mathcal{J}_n|$, and spatial correlation given by (4.29). The scatterer had bearing θ_n and was

assumed to be sufficiently far from the receiver that the θ_n did not change significantly throughout the trajectory. The angles θ_n and ψ_n were uniformly distributed on $(-\pi/2, \pi/2]$. The ratio of distances $R_1/(R_1 + R_2)$ was uniformly distributed on $(0, 1]$. (In practice multiple scattering leads to a more diffuse field, and the last scatterer is likely to be closer to the receiver, so the assumption here is a conservative one). Each of the M values of the resulting narrowband channel was synthesised according to

$$r_m = \sum_{n=1}^{N} (\zeta_n + u_{mn}) e^{jkm\Delta_x \sin\theta_n}. \tag{4.31}$$

Each source amplitude was scaled so that $E\{|r_m|^2\} = 1$. The prediction technique used was the principal components linear prediction (PCLP) [157] technique as implemented in Section 4.2.2. The channel data for samples after the 40th was then predicted using data from the first 40 (i.e., from the first 6 wavelengths), and the point was recorded at which the error first exceeded 0.2. The mean of such distances over 3000 scenarios is presented in Fig. 4.19 (an alternative graph of the distance at which the mean square error over the scenarios exceeds a certain threshold is very similar to Fig. 4.19). Each line in the figure is for a different value of roughness σ_ξ. For the left-most region of the figure $k\sigma_r = 2\pi$, so for most of the figure the validity condition $k\sigma_r \gg 1$ is satisfied.

Since $|\sin\psi| < 1$, and $\sigma_\xi/\lambda < 1/8$, all of the surface roughness conditions presented in the figure qualify as smooth, according to the Rayleigh criterion. Most of the assumptions made also favour high correlations in the diffuse component of the scattered field. It is obvious however, that even these surfaces can have a serious effect on the predictability of the channel. For instance, at 900MHz, the top trace, ($\sigma_\xi = \lambda/64$) corresponds to surface irregularities of the order of only 5mm. If the correlation distance σ_r is of the order of 3λ (about 1 metre), the mean prediction distance is decreased to one third of the perfectly smooth case. Such roughness is likely to be encountered in many situations in practice, in trees, roads and buildings. A table of typical surface variances and correlation distances which may be encountered in practice is presented in Table 4.1.

4.5.4 Conclusion

The effect of scattering from rough surfaces has been investigated by modifying the discrete point sources model to take account of rough surfaces being involved in signal reflection. It has been shown that even surfaces which are only slightly rough (according to the Rayleigh criterion) produce a diffuse reflection which can provide a serious barrier to real-time channel prediction. This is consistent with the conclusion reached in Section 4.2.3, since a rough surface is equivalent to a large number of scatterers, found there to be a significant limitation.

	900 MHz		2400 MHz	
Surface	σ_ξ/λ	σ_r/λ	σ_ξ/λ	σ_r/λ
Road	0.02	0.02	0.06	0.06
Brick Wall	0.02	0.2	0.06	0.6
Wall with Windows	0.3	3	0.8	8
Trees	0.09	0.09	0.2	0.2
Shelving	0.3	0.9	0.8	2.4

Table 4.1: Typical values of surface variance (σ_ξ) and correlation distance (σ_r).

4.6 Mutual Information Considerations

This section (largely motivated by [42]) provides a brief overview of the approach which may be used to calculate a bound on the range of accurate prediction based on mutual information considerations. These are independent of the method of prediction used, and so in a sense form "absolute" bounds.

A bound on the gain (4.1) of an arbitrary predictor has been proposed in [17]. The predictor predicting L steps forward into the future, with memory of m values (m may be infinite) has gain limited by

$$G(L) \leq 2^{2(M_{r\vec{r}}(L)+\Delta)} \quad \text{or} \quad G_{\text{dB}}(L) \leq 20\log_{10}(2)(M_{r\vec{r}}(L)+\Delta) \tag{4.32}$$

where

$$\Delta = \frac{1}{2}\log_2(2\pi e E\{r(t) - E\{r(t)\}\}) - H(r(t)) \tag{4.33}$$

is the difference between the entropy of $r(t)$ and a normal variable with the same variance, and $M_{r\vec{r}}(L)$ is the mutual information between the past values $\vec{r}(t) = (r(t), r(t-\tau_1), r(t-\tau_2), \ldots, r(t-\tau_m))^T$, and the predicted value $r(t+L)$. The mutual information is this case is defined as

$$M_{r\vec{r}}(L) \triangleq \int p(r(t+L), \vec{r}) \log_2 \left(\frac{p(r(t+L), \vec{r})}{p(r(t+L))p(\vec{r})} \right) dr(t+L)d\vec{r}. \tag{4.34}$$

where $p(\cdot,\cdot)$ and $p(\cdot)$ represent the appropriate joint and marginal density functions.

The mutual information $M_{r\vec{r}}$ and the entropy $H(r(t))$ can be estimated from the data, using the algorithm proposed in [18] or [51]. This requires estimation of multidimensional density functions, and hence is a procedure subject to the "curse of dimensionality"[1] [13]. Without an large amount of data, it becomes infeasible to estimate the density functions. The dimensions are double the order of the predictor since the

[1]The phrase "curse of dimensionality" is used to describe either the problems associated with the feasibility of density estimation in many dimensions, or the complexity of computations connected with certain types of signal processing.

values of the base band channel are complex.

A program to implement the algorithm described in [18] was obtained from the author, but insufficient data was available to implement the algorithm. An area of possible future research however, is to derive bounds on the prediction gain from theoretical consideration of the mutual information rather than measured data. The theoretical analysis could be based on near field situations for example.

4.7 Multipath Dimensionality

This section derives a bound on the extent to which a mobile channel can be predicted based on the properties of the wave equation in a finite region of space. The concept is introduced in Sections 4.7.1 to 4.7.5, included here from [78] with permission. Section 4.7.6 is an original extension of this work.

A theoretical model is presented which can be applied to *any* narrowband multipath environment regardless of the number or nature of the multipath sources. It is shown that there is an intrinsic dimensionality to a narrowband multipath field in a region of space of a given size. This is used to show that there is an upper limit on the degree of multipath richness which can exist within that area.

It is shown that the field in a given area of space can be represented by relatively few terms of a functional expansion, the number of which represent the dimensionality. Furthermore, it is possible to synthesise an arbitrary field as accurately as desired using an appropriate combination of arbitrary sources, near field or far field, discrete or continuous.

This result is then used to provide an indication of the accuracy with which the knowledge of the field inside a region can be used to deduce knowledge of the field *outside* the region.

4.7.1 General Two Dimensional Multipath Field

Consider two dimensional (2D) narrowband multipath interference in a region of some given size. This models the situation in three dimensions where the multipath is restricted to the horizontal plane, having no components arriving at large elevations. As such, the multipath field is height invariant. The multipath signals may have sources which are near field or far field, specular or diffuse. Polar co-ordinates are used to represent a point in space $\boldsymbol{x} \equiv (\|\boldsymbol{x}\|, \phi_x)$. The field, $F(\boldsymbol{x}; k)$, is a function of the position and the wave number, $k = 2\pi/\lambda$, and is a solution to the Helmholtz wave equation in polar co-ordinates [35], the most general solution of which is

$$F(\boldsymbol{x}; k) = \sum_{n=-\infty}^{\infty} \alpha_n \, J_n(k\|\boldsymbol{x}\|) \, e^{jn\phi_x} \tag{4.35}$$

where α_n are complex constants independent of position and $J_n(\cdot)$ is the order n Bessel function [94]. That is, in (4.35) the field strength at a point $\boldsymbol{x}$ is represented as a weighted sum of orthogonal basis functions.

Plane and circular waves are examples of 2D waves which can be expressed in the form (4.35). A single plane wave with complex amplitude a_p and propagation direction ϕ_p has $\alpha_n = a_p j^n e^{-jn\phi_p}$ [31, p.66]. A single circular wave with source position vector, $\boldsymbol{y}_p \equiv (\|\boldsymbol{y}_p\|, \phi_p)$, has $\alpha_n = a_p e^{-jn\phi_p} H_n^{(1)}(k\|\boldsymbol{y}_p\|)$ [31, p.66] where $H_n^{(1)}(\cdot)$ is the order n Hankel function of the first kind.

The application of (4.35) is now illustrated with fields generated by a superposition of P plane waves. Let plane wave of index p have complex amplitude a_p and propagation direction ϕ_p with normalised direction $\hat{\boldsymbol{\eta}}_p \equiv (\cos\phi_p, \sin\phi_p)$. The field strength at $\boldsymbol{x}$ is then given by

$$F(\boldsymbol{x}; k) = \sum_{p=1}^{P} a_p e^{-jk\boldsymbol{x}\cdot\hat{\boldsymbol{\eta}}_p} \tag{4.36}$$

$$= \sum_{p=1}^{P} a_p \sum_{n=-\infty}^{\infty} j^n J_n(k\|\boldsymbol{x}\|) e^{jn(\phi_x - \phi_p)} \tag{4.37}$$

where (4.36) is a Cartesian form and (4.37) is the polar equivalent. Note (4.37) is in the form of (4.35) with

$$\alpha_n = \sum_{p=1}^{P} a_p j^n e^{-jn\phi_p} = \sum_{p=1}^{P} a_p e^{-jn(\phi_p - \pi/2)}. \tag{4.38}$$

Similarly, a field of P circular waves, where $\boldsymbol{y}_p \equiv (\|\boldsymbol{y}_p\|, \phi_p)$ is the position vector of the source of index p, is given by

$$F(\boldsymbol{x}; k) = \sum_{p=1}^{P} a_p \frac{e^{jk\|\boldsymbol{x}-\boldsymbol{y}_p\|}}{\|\boldsymbol{x}-\boldsymbol{y}_p\|} = \sum_{p=1}^{P} a_p \sum_{n=-\infty}^{\infty} H_n^{(1)}(k\|\boldsymbol{y}_p\|)\, J_n(k\|\boldsymbol{x}\|) e^{jn(\phi_x - \phi_p)} \tag{4.39}$$

and for the representation (4.35)

$$\alpha_n = \sum_{p=1}^{P} a_p H_n^{(1)}(k\|\boldsymbol{y}_p\|) e^{-jn\phi_p}. \tag{4.40}$$

A field of plane and circular waves could also be formed by linearly combining (4.37) and (4.39) to obtain an equation of the form of (4.35), with α_n the weighted sum of (4.38) and (4.40).

The Bessel functions $J_n(\cdot)$ for $n \geq 1$ in (4.35) have a spatial high pass character ($J_0(\cdot)$ is spatially low pass). That is, as illustrated in Fig. 4.20, $J_n(z)$ starts small increasing monotonically to its maximum at arguments around $O(n)$ before decaying

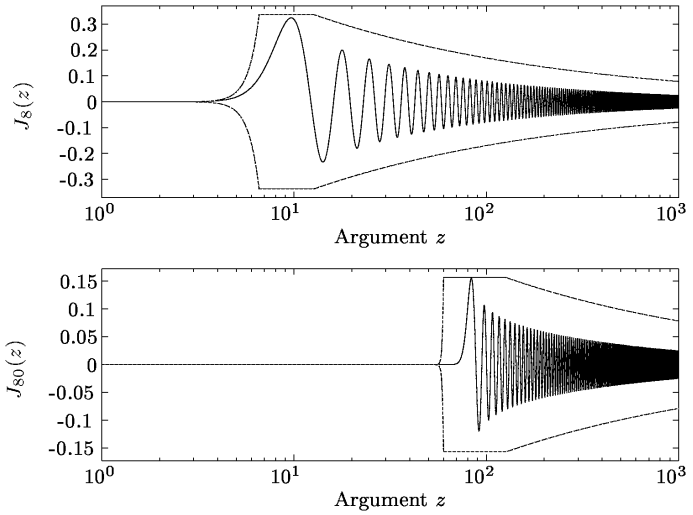


Figure 4.20: High pass character of the Bessel functions $J_8(z)$ and $J_{80}(z)$ versus argument z (horizontal axis is on a logarithmic scale). Also shown are the combination of three bounds.

asymptotically to zero as $z \to \infty$ (oscillating as it does so). Also shown in Fig. 4.20 are limits imposed by three upper bounds on $|J_n(z)|$: $1/n!(z/2)^n$, $0.6748851/n^{1/3}$ and $0.7857468704/z^{1/3}$ [86].

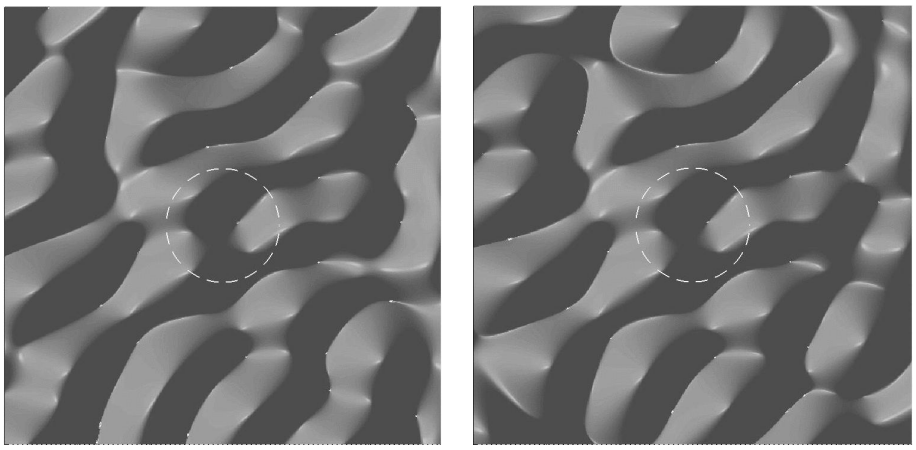
4.7.2 Dimensionality of Multipath

The aim now is to quantify the complexity of an arbitrary multipath field $F(\boldsymbol{x}; k)$ in a circular region of radius R wavelengths by defining the effective *dimensionality* of the field. This is achieved by truncating the series in (4.35) and determining the minimum number of terms, $2N + 1$, for the field $F_N(\boldsymbol{x}; k)$, so generated, to be within a specified error, ϵ, of the actual field $F(\boldsymbol{x}; k)$. Thus, the approximate field strength $F_N(\boldsymbol{x}; k)$ is defined by the finite sum

$$F_N(\boldsymbol{x}; k) = \sum_{n=-N}^{N} \alpha_n J_n(k\|\boldsymbol{x}\|) e^{jn\phi_x}, \tag{4.41}$$

where the approximation is to be sufficiently accurate within the region $\|\boldsymbol{x}\| \leq R$.

Fig. 4.21 shows the actual field strength of the sum of 30 plane waves, as in (4.36), over a $3\lambda \times 3\lambda$ area, compared with the same field represented by the truncated series in (4.41) with α_n given by (4.38) and $N = 7$. Clearly $F_N(\boldsymbol{x}; k)$ can model $F(\boldsymbol{x}; k)$ very well in a finite region about the origin.



(a) (b)

Figure 4.21: Example of accuracy of truncation in (4.41) shown by isosurfaces of field amplitude. The actual field (a) (4.36) has 30 component plane waves in random directions. The approximate field (b) (4.41) has $2N+1 = 15$ plane waves. The approximate field is within 1×10^{-3} of the actual field for $R/\lambda \leq 0.3890$ (the circular region highlighted in each subfigure).

4.7.3 Bounding the Relative Error

Consider the magnitudes of the α_n coefficients for a superposition of a possibly infinite number of plane waves indexed by p with amplitudes a_p. From (4.38)

$$|\alpha_n| \leq \Big| \sum_p a_p j^n e^{-jn(\phi_p - \frac{\pi}{2})} \Big| \leq \sum_p |a_p| \tag{4.42}$$

The RHS of (4.42) is an upper bound on the field strength at any point being the sum of the amplitudes of the plane waves constituting the given wave field. On physical grounds the field is assumed to be bounded at all points in space which implies $\sum_p |a_p| < B$. Without loss of generality take $B = 1$, amounting to a normalisation such that the field strength is bounded by unity. Hence $|\alpha_n| \leq 1 \; \forall n$.

To form a relative error, first bound the peak amplitude of the multipath field to unity. This implies from the foregoing considerations, that $|\alpha_n| \leq 1, \forall n$. Define the error between the actual and approximate fields by $\epsilon_N(\boldsymbol{x})$. Then

$$\begin{aligned} \epsilon_N(\boldsymbol{x}) &= \big|F(\boldsymbol{x};k) - F_N(\boldsymbol{x};k)\big| = \Big| \sum_{|n|>N} \alpha_n \, J_n(k\|\boldsymbol{x}\|) \, e^{jn\phi_x} \Big| \\ &\leq \sum_{|n|>N} \big|J_n(k\|\boldsymbol{x}\|)\big| = 2 \sum_{n>N} \big|J_n(k\|\boldsymbol{x}\|)\big| \end{aligned} \tag{4.43}$$

Now a bound on $|J_n(\cdot)|$ is required for $n > N$. For integer $n \geq 0$, the order n Bessel

function is given by [83]

$$J_n(z) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell (z)^{2\ell+n}}{2^{2\ell+n}\,\ell!\,(\ell+n)!}, \quad n \geq 0. \tag{4.44}$$

From [94, p.192], for $n > -1/2$

$$\begin{aligned} J_n(z) &= \frac{1}{\sqrt{\pi}\,\Gamma(n+\frac{1}{2})} \left(\frac{z}{2}\right)^n \int_{-1}^{+1} e^{j\lambda z}(1-\lambda^2)^{n-\frac{1}{2}}\,d\lambda \\ &\leq \frac{1}{\sqrt{\pi}\,\Gamma(n+\frac{1}{2})} \left(\frac{z}{2}\right)^n \int_{-1}^{+1} (1-\lambda^2)^{n-\frac{1}{2}}\,d\lambda \end{aligned} \tag{4.45}$$

where $\Gamma(\cdot)$ is the Gamma function. From (15.24) of [136], the integral (4.45) is equal to $\sqrt{\pi}\,\Gamma(n+1/2)/\Gamma(n+1)$. Thus,

$$J_n(z) \leq \frac{1}{\Gamma(n+1)} \left(\frac{z}{2}\right)^n, \tag{4.46}$$

and so $J_n(\cdot)$ is always bounded by the first term in (4.44). An example of this upper bound is shown in Fig. 4.20 for $J_8(\cdot)$ and $J_{80}(\cdot)$.

By using the Stirling lower bound on $n!$ [83] $|J_n(k\|\boldsymbol{x}\|)|$ can be bounded as

$$\begin{aligned} \left|J_n(k\|\boldsymbol{x}\|)\right| &\leq \frac{(k\|\boldsymbol{x}\|)^n}{2^n n!} \leq \frac{1}{\sqrt{2\pi n}} \left(\frac{ke\|\boldsymbol{x}\|}{2n}\right)^n \\ &\leq \frac{\rho(N,R)^n}{\sqrt{2\pi(N+1)}}, \quad n > N,\ \|\boldsymbol{x}\| \leq R. \end{aligned} \tag{4.47}$$

where

$$\rho(N,R) = \frac{keR}{2(N+1)} = \frac{\pi eR/\lambda}{(N+1)}. \tag{4.48}$$

Substituting (4.47) and (4.48) into (4.43) and choosing N large enough such that $\rho(N,R) < 1$, the following bound on the error for given values of N and R is obtained:

$$\epsilon_N(\boldsymbol{x}) \leq \sqrt{\frac{2}{(N+1)\pi}} \cdot \frac{\rho(N,R)^{N+1}}{1-\rho(N,R)}, \quad \forall \|\boldsymbol{x}\| \leq R. \tag{4.49}$$

The restriction on $\rho(N,R)$ gives a lower bound on N

$$N > (\pi e)\,R/\lambda - 1. \tag{4.50}$$

With N thus selected $F(\boldsymbol{x};k)$ and $F_N(\boldsymbol{x};k)$ are essentially indistinguishable within $\|\boldsymbol{x}\| \leq R$. Since $2N+1$ basis elements are used, $2N+1$ characterises the dimensionality.

The dimensionality of fields of radius $R/\lambda \leq 3$, for error thresholds between 10^{-1} and 10^{-4} are shown in Fig. 4.22. The number of extra terms required to achieve an

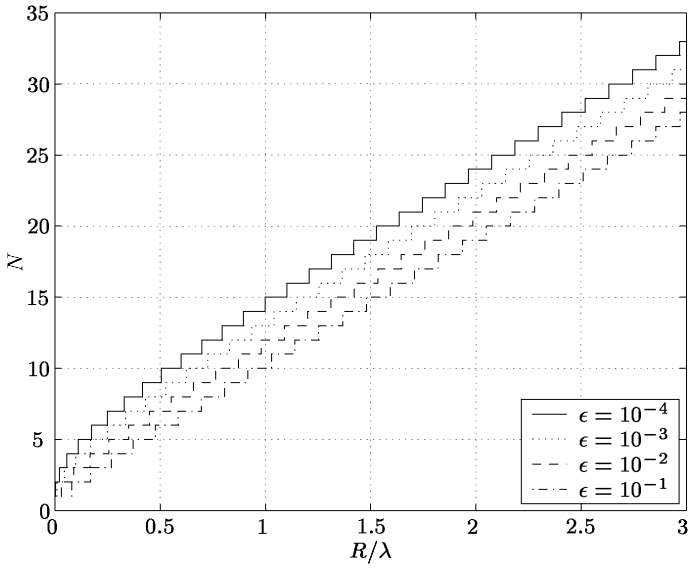


Figure 4.22: Minimum N required for different error thresholds ϵ, as given by theoretical bounds (4.49) and (4.50), for increasing values of radius R/λ.

error of 10^{-4} over an error of 10^{-1} is relatively small. For $R/\lambda = 1$, the minimum values of N required are 10 and 15, respectively, meaning that just 10 extra terms give 3 orders of magnitude improvement in accuracy.

Simulation results for the relationship between N and the error ϵ between actual and approximate fields of radius $R/\lambda = 1$ are shown in Fig. 4.23. Simulated actual fields had from $P = 4$ to 50 plane wave components. The error appears to converge, regardless of the value of P at around $N = 10$.

The following observations may be made

1. Field complexity or dimensionality, $(2N + 1)$, increases linearly with R/λ.

2. The actual number of terms required to represent an arbitrary field to high accuracy is relatively small. The addition of just a few terms can give orders of magnitude improvement in the relative error.

3. Any wave field, including a a diffuse field, consisting of any number of actual wave components may be represented by relatively few parameters.

4.7.4 Dimensionality of 3 Dimensional Field

The dimensionality for a three dimensional (3D) field can be similarly specified. The general equation for a three dimensional wave, using spherical co-ordinates is

$$F(\boldsymbol{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \alpha_n(k) j_n(k\|\boldsymbol{x}\|) Y_{nm}(\hat{\boldsymbol{x}}) Y_{nm}^*(\hat{\boldsymbol{y}}), \tag{4.51}$$

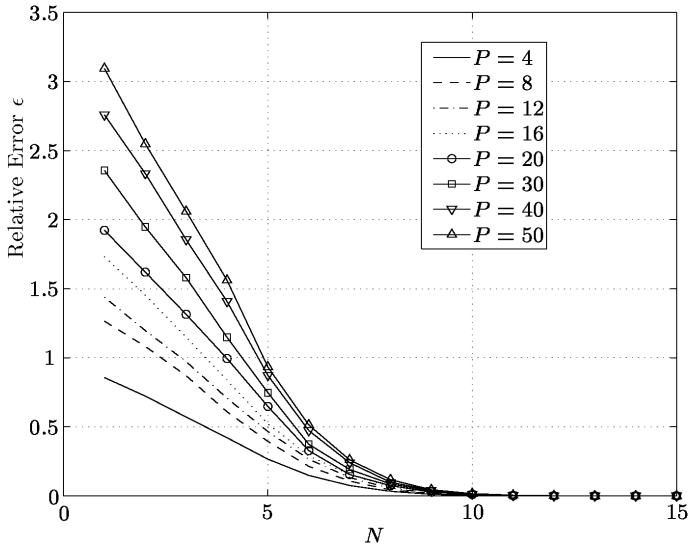


Figure 4.23: Average error between actual and approximate fields, as in (4.36) and (4.41) for up to $P = 50$ plane waves and $N \leq 15$ for $R/\lambda = 1$.

where $Y_{nm}(\cdot)$ is a spherical harmonic function (2.34), $\hat{\boldsymbol{x}}$ is the unit vector in the direction of the position $\boldsymbol{x}$, $\hat{\boldsymbol{y}}$ is the unit vector in the direction of the source at $\boldsymbol{y}$, and $j_n(\cdot)$ is the order n spherical Bessel function, related to the ordinary Bessel function by

$$j_n(r) = \sqrt{\frac{\pi}{2r}} J_{n+\frac{1}{2}}(r). \tag{4.52}$$

The spherical Bessel function is bounded by the first term of the Taylor expansion, in a similar way to the ordinary Bessel function. Thus, using the Sterling bound for the Gamma function [83]

$$\begin{aligned} j_n(k\|\boldsymbol{x}\|) &\leq \sqrt{\frac{\pi}{2k\|\boldsymbol{x}\|}} \frac{\left(\frac{k\|\boldsymbol{x}\|}{2}\right)^{(n+\frac{1}{2})}}{\Gamma(n+\frac{3}{2})} \\ &= \sqrt{\frac{e}{2}} \frac{1}{2n+1} \left(\frac{keR}{2N+1}\right)^n \quad r \leq R, n \leq N. \end{aligned} \tag{4.53}$$

In the 2D case it was assumed that $|\alpha_n| \leq 1$. It may be similarly assumed in this case that $Y_{nm}(.) \leq 1$. Defining $\rho(N,R) \triangleq keR/(2N+1)$,

$$\begin{aligned} \epsilon_n &= |F(\boldsymbol{x};k) - F_N(\boldsymbol{x};k)| \\ &\leq \sum_{n=N+1}^{\infty} (2n+1)|j_n(k\|\boldsymbol{x}\|)| \\ &\leq \sum_{N+1}^{\infty} \sqrt{\frac{e}{2}} \left(\frac{keR}{2N+1}\right)^n \end{aligned}$$

$$\leq \left(\sqrt{\frac{e}{2}}\right) \frac{\rho(N,R)^{N+1}}{1-\rho(N,R)}, \qquad \rho(R,N) < 1. \tag{4.54}$$

The upper bound on $\rho(R,N)$ gives a lower bound for N as

$$N > \frac{keR}{2} - \frac{1}{2} \tag{4.55}$$

This bound is very similar to that for the two dimensional case. Note however, that since (4.51) involves a double summation, the number of independent parameters required increases with the *square* of N.

4.7.5 Plane Wave Synthesis

Equation (4.41) shows that an arbitrary wave field in a given region can be represented by a finite number of terms, regardless of the complexity of the scattering environment. Such a field can also be closely approximated by an appropriate combination of plane waves.

Once the coefficients α_n in (4.41) are determined for the field of interest, one can use (4.38) to define a set of plane waves producing the same α_n's over the critical indices $|n| \leq N$. Let $\boldsymbol{\alpha} = [\alpha_{-N}, \cdots, \alpha_n, \cdots, \alpha_N]^T$ and $\boldsymbol{a} = [a_1, \cdots, a_p, \cdots, a_P]^T$. Define the diagonal matrix $\Theta = \text{diag}[j^N, \cdots, j^n, \cdots, j^{-N}]$ and the Vandermonde matrix

$$\mathbf{V} = \begin{bmatrix} e^{jN\phi_1} & \cdots & e^{jN\phi_p} & \cdots & e^{jN\phi_P} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e^{-jn\phi_1} & \cdots & e^{-jn\phi_p} & \cdots & e^{-jn\phi_P} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e^{-jN\phi_1} & \cdots & e^{-jN\phi_p} & \cdots & e^{-jN\phi_P} \end{bmatrix}. \tag{4.56}$$

Equation 4.38 may be written in matrix form as

$$\Theta\boldsymbol{\alpha} = \mathbf{V}\boldsymbol{a}. \tag{4.57}$$

As each of the ϕ_p's are distinct, it is known that $\mathbf{V}$ is non-singular. Thus, given specific $\boldsymbol{\alpha}$, $\mathbf{V}$ and Θ, (4.57) can always be solved for $\boldsymbol{a}$. Therefore, at most $P = 2N + 1$ plane waves from arbitrary directions can synthesise an arbitrary field over a region $\|\boldsymbol{x}\| \leq R$ whenever $N \approx (\pi e)R/\lambda$.

The following observations can be made

1. There are an infinite number of plane wave combinations which can represent a given field $F(\boldsymbol{x}; k)$. The choice of directions $\{\phi_p\}$ is arbitrary provided the directions are distinct.

2. By choosing $\phi_p = 2p\pi/P$, $\mathbf{V}$ in (4.56) becomes a scaled discrete Fourier transform

matrix and the plane wave weights, a_p, can be efficiently computed using the FFT from the weights α_n.

3. The result is not restricted to plane waves. Any superposition of sources can be used, such as a superposition of near field point sources, etc.

4.7.6 Extrapolation Bounds

Consider now a circular (2D) region of radius R, in which the field $F(\boldsymbol{x};k) = F(r,\phi)$ is known to within some tolerance. The set of α_n may be estimated from this field knowledge by integration around any circle of radius r

$$\alpha_n = \frac{1}{2\pi J_n(kr)} \int_{-\pi}^{\pi} F(r,\phi) e^{-jn\phi}\, d\phi. \tag{4.58}$$

Suppose that

$$\hat{F}(\boldsymbol{x}) = F(\boldsymbol{x}) + \eta(\boldsymbol{x}) \tag{4.59}$$

where $\eta(\boldsymbol{x})$ is some random process representing any inaccuracy in the knowledge of $F(\boldsymbol{x})$. It is assumed that $\eta(\boldsymbol{x}) = \eta(r,\phi)$ is a wide sense stationary process and has an autocorrelation function with respect to ϕ of $\Psi(r,\phi)$, with Fourier transform

$$\tilde{\Psi}(r,n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Psi(r,\phi) e^{-jn\phi}\, d\phi. \tag{4.60}$$

Note that since $\eta(r,\phi)$ is periodic in ϕ, so also is $\Psi(r,\phi)$, and so the appropriate Fourier transform is a Fourier series discrete in n. This may be represented as

$$\Psi(r,\phi) \underset{\phi \leftrightarrow n}{\overset{\mathcal{FS}}{\Longleftrightarrow}} \tilde{\Psi}(r,n) \tag{4.61}$$

It is known that the Fourier transform of $\Psi(r,\phi)$ is also the magnitude squared of the Fourier transform of $\eta(r,\phi)$. From (4.58), it follows that

$$\hat{\alpha}_n - \alpha_n = \frac{1}{2\pi J_n(kr)} \int_{-\pi}^{\pi} \eta(r,\phi) e^{-jn\phi}\, d\phi. \tag{4.62}$$

Taking the expectation of the correlation between two such terms

$$\begin{aligned}
& E\{(\hat{\alpha}_{n_1} - \alpha_{n_1})(\hat{\alpha}_{n_2} - \alpha_{n_2})^*\} \\
&= \frac{1}{(2\pi)^2 J_{n_1}(kr_{n_1}) J_{n_2}(kr_{n_2})} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} E\{\eta(r_{n_1},\phi_1)\eta^*(r_{n_2},\phi_2)\}\, e^{j(n_2\phi_2 - n_1\phi_1)}\, d\phi_1\, d\phi_2 \\
&= \frac{1}{(2\pi)^2 J_{n_1}(kr_{n_1}) J_{n_2}(kr_{n_2})} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \Psi(r_{n_1},\phi_1 - \phi_2)\, e^{j(n_2\phi_2 - n_1\phi_1)}\, d\phi_1\, d\phi_2 \\
&= \frac{1}{(2\pi)^2 J_{n_1}(kr_{n_1}) J_{n_2}(kr_{n_2})} \int_{-\pi}^{\pi} e^{j\phi_2(n_2 - n_1)}\; 2\pi\tilde{\Psi}(r_{n_1},-n_1)\, d\phi_2
\end{aligned}$$

$$= \frac{1}{J_{n_1}(kr_{n_1})J_{n_2}(kr_{n_2})}\tilde{\Psi}(r_{n_1}, -n_1)\delta_{n_1,n_2} \tag{4.63}$$

The notation $\Psi(r_{n_1}, \phi_1 - \phi_2)$ in the above derivation hides the dependence of $\Psi(r, \phi)$ on r_{n_2} as well as r_{n_1}. However, the presence of the Kronecker delta in the last line assures that the final result only depends on one value of r_n.

The expected error in reconstructing the signal $\hat{F}(r, \phi)$ from the calculated values of α_n is now

$$\begin{aligned}
&E\{|\hat{F}(r,\phi) - F(r,\phi)|^2\} \\
&= E\left\{\left|\sum_{n=-\infty}^{\infty}(\hat{\alpha}_n - \alpha_n)J_n(kr)e^{jn\phi}\right|^2\right\} \\
&= \sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty}E\{(\hat{\alpha}_{n_1} - \alpha_{n_1})(\hat{\alpha}_{n_2} - \alpha_{n_2})^*\}J_{n_1}(kr)J_{n_2}(kr)e^{j\phi(n_1-n_2)} \\
&= \sum_{n=-\infty}^{\infty}\frac{J_n^2(kr)}{J_n^2(kr_{n_1})}\tilde{\Psi}(r_{n_1}, -n),
\end{aligned} \tag{4.64}$$

where the notation r_{n_1} has been retained from (4.63) to indicate the radius at which the integration (4.58) is performed to estimate the values of α_n. This distance may or may not be the same as r.

Once the values of α_n are known, (4.35) may be used in estimate the field at *any* radius. Clearly the error in (4.64) can be minimised by choosing r_{n_1} so that $J_n^2(kr_{n_1})$ is a maximum. If the field is only known for areas inside the region $r < R_1$, however, the values of $J_n^2(kr_{n_1})$ which can be achieved are quite small, and so the resulting error is large.

If it is assumed that the correlation function $\Psi(\cdot, \cdot)$ is radially symmetric, it may be expressed as a function $\xi(\rho)$ of the radius, so that, expressing $\eta(r, \phi)$ in Cartesian co-ordinates $\eta_{xy}(x, y)$,

$$E\left\{\eta_{xy}(x,y)\eta_{xy}^*(x - x_0, y - y_0)\right\} = \xi\left(\sqrt{x_0^2 + y_0^2}\right). \tag{4.65}$$

The correlation with respect to the angle ϕ is then

$$\begin{aligned}
\Psi(r,\phi_0) &= E\{\eta(r,\phi)\eta^*(r,\phi-\phi_0)\} \\
&= E\{\eta_{xy}(r\cos\phi_0, r\sin\phi_0)\eta_{xy}^*(r\cos(\phi-\phi_0), r\sin(\phi-\phi_0))\} \\
&= \xi\left(\left((r\cos\phi - r\cos(\phi-\phi_0))^2 + (r\sin\phi - r\sin(\phi-\phi_0))^2\right)^{\frac{1}{2}}\right) \\
&= \xi\left(\sqrt{2}r\sqrt{1-\cos\phi_0}\right).
\end{aligned} \tag{4.66}$$

If, to give a symbolically tractable example, the radially symmetric correlation function

is specified by

$$\xi(\rho) = \kappa \exp(-\rho^2/(2\sigma_\eta^2)), \tag{4.67}$$

then using (512-5b) of [66] the Fourier transform of $\Psi(r, \phi)$ is given by

$$\begin{aligned}\tilde{\Psi}(r, n) &= \frac{\kappa}{2\pi} \int_{-\pi}^{\pi} e^{-r^2(1-\cos\phi_0)/\sigma_\eta^2} e^{jn\phi_0}\, d\phi_0 \\ &= \kappa e^{-r^2/\sigma_\eta^2} I_n\left(\frac{r^2}{\sigma_\eta^2}\right)\end{aligned} \tag{4.68}$$

where $I_n(\cdot)$ is the modified Bessel function of the first kind.

For the field $F(r, \phi)$ to be estimated to within the accuracy ϵ, the values of N for which (4.41) must be evaluated is known to be limited by (4.49). Provided that $r \leq R$, values of $r_{n_1} < R$ can be chosen so that the error is kept relatively small. If $r > R$ the expected error can become very large. This is shown by a numerical experiment in Fig. 4.24. The "normal shaped" example correlation function (4.67) given is used, with $\kappa = 1$, and $\sigma_\eta = \lambda/(2\pi)$ so that $r/\sigma_\eta = rk$. The value of N was chosen according to (4.49) so that $\epsilon < 0.2$. The error is then calculated according to (4.64). It can be seen from Fig. 4.24 that the error remains very low while ever $r < R$, but as soon as extrapolation is attempted outside the region of knowledge, the error rapidly increases. The range of useful extrapolation may be considered to be those values of r for which the error is of the same order as that in the region of known field $r < R$. This is shown in Fig. 4.24 as the values of expected error ≤ 1. It can be seen from the figure that the range of useful extrapolation does increase slowly as R increases.

These results can be derived for the case of a three dimensional field. Similar results can also be derived using any analytic orthonormal basis.

4.7.7 Conclusions

It was shown in [78] that inside a circular or spherical region, any narrowband field (even a diffuse field) can be reconstructed to within a certain tolerance with only a finite number of sources. The number of sources, called the *dimensionality* scales linearly with the radius of the region.

It has been shown in this section that although in principle, extrapolation of the knowledge of the field beyond the region is possible, the error rapidly becomes prohibitive. This situation is somewhat different to the prediction situation presented elsewhere in this thesis. The conclusions however, are entirely consistent with those of Sections 4.2.3 and 4.4, that if the number of paths is large (and so the field may be considered as effectively diffuse) the prediction range is severely limited.

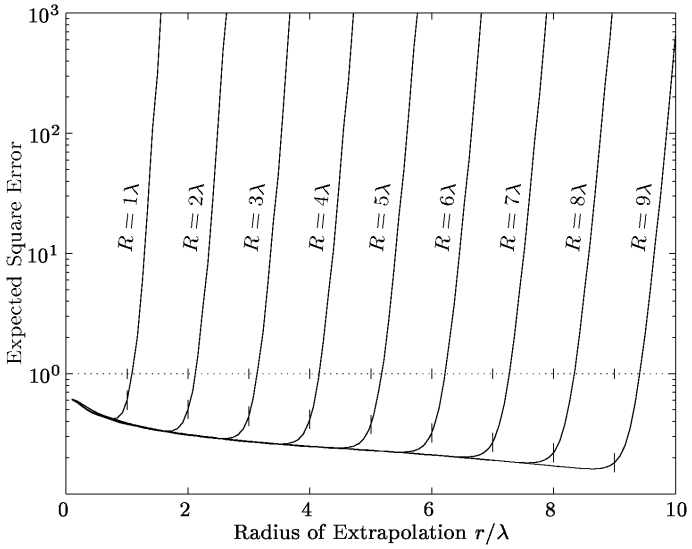


Figure 4.24: The expected error in the reconstructed field as the radius of the extrapolation increases, for various radii of known field. The small marks for the values at which $r = R$ indicate that although the expected error increases rapidly when $r > R$, the error stays within that of the known region for a slightly larger extent $r - R$ as R increases.

4.8 Summary and Contributions

The effectiveness of the proposed channel models in providing a real time syndetic characterisation of both simulated and measured channels has been evaluated. Some factors which may limit real time prediction have been identified. The following conclusions have been drawn:

i. The number of significant scatterers is a vital factor in the viability of long range channel prediction. If there are many scatterers accurate prediction may be limited to only a small fraction of one wavelength. As seen in Chapter 3 this situation may occur in practice more often than is commonly realised.

ii. The parameters associated with sources near the receiver may vary rapidly.

iii. Many surfaces in both indoor and outdoor mobile environments possess sufficient roughness to have a significant impact on the predictability of channels.

iv. Macroscopic changes in the scattering environment, such as shadowing, are in general unpredictable, and the frequency of such changes has a significant effect on the expected prediction range.

We itemise some specific contributions made in this chapter:

i. The limiting factors to long range prediction listed above, have been identified and quantified.

ii. The Cramer Rao bound for the variance of channel prediction error has been derived.

iii. The first application of the theory of rough surface scattering to the problem of channel prediction has been presented.

iv. The error which may be encountered in extrapolation of a field beyond a known circular or spherical region has been presented.

Chapter 5

Channel Performance

PREVIOUS chapters have examined the feasibility and limitations of real time channel prediction. One of the most simple applications of the prediction information is in "avoidance" of channels which are known in advance to offer poor performance. This avoidance could take the form of simply suspending transmission for a short time, changing time slot (in a TDMA system) or changing frequency (in a FDMA system).

However, the relationship between the performance of a communication system and the channel over which it operates is not a simple one. For even a moderately broadband system, the system error rate depends not only on the total signal power available from the channel to the receiver (and thus the SNR), but also on how this power is distributed in the channel impulse response.

For a system to be able to determine when a channel should be avoided, it must have an efficient method of relating a channel impulse response to the system performance.

In this chapter such a method is derived. It is assumed that a *linear* modulation scheme is used, and that Maximum Likelihood Sequence Estimation (MLSE) such as implemented using the Viterbi algorithm is used for equalisation. MLSE has been analysed here because it defines (under the conditions of a static, perfectly known channel) a lower bound on the error probability [48, 56]. Analysis could equally well be performed for different systems, such as those using continuous phase modulations and/or linear or decision feedback equalisation, although such analysis would take quite a different form from that derived here. It is also assumed that inaccuracy in the predicted channel information can be ignored. Section 5.6 examines the effect of channel estimation errors on the error rate prediction accuracy.

Previous analyses of the probability of error of systems using MLSE [7, 87, 93, 116, 132] have focussed on the "worst case" channel. These give an approximation to the channel error probability where the channel is unknown. A *much* better approximation can be obtained when the channel is known. To give one example only, the worst-case length 7 QPSK channel needs transmission power 11.4 dB greater than the best-case

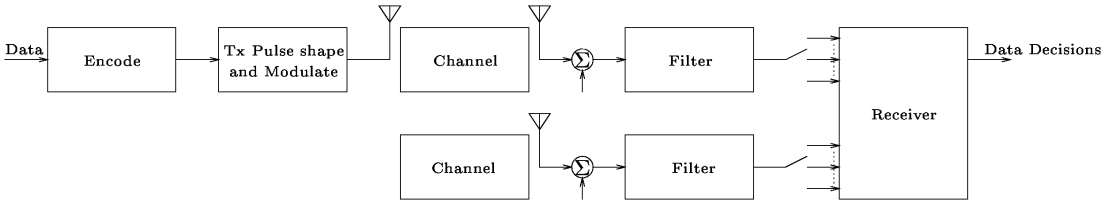


Figure 5.1: Structural representation of the communications system, possibly with multiple receive antennas, and possibly with fractional spaced sampling.

channel of the same length to achieve the same error rate.

In this chapter it is assumed that the channel is varying in time slowly enough for it to be considered unchanged for several symbol periods. From the previous chapters, where the focus has been on the time-varying nature of the mobile channel, it can be seen that this assumption is not always valid. If the channel information used for equalisation is continually updated [19, 25], the long error events discussed in the analysis below would not contribute to the overall error rate as much as the analysis predicts.

5.1 MLSE Error Probability

This section introduces the terminology used in this chapter and derives the probability of error for a particular system operating with a particular channel.

The system used in this chapter is represented in Fig. 5.1. Note that there may be multiple received data points corresponding to each transmitted symbol due to the use of multiple receiving antennas and/or fractionally spaced sampling.

The data is linearly modulated using a constellation such as one of the those represented in the left column of Fig. 5.2. The alphabet of symbols available is denoted by $\mathcal{A}$. The alphabet of possible symbol errors or *differences* is denoted by $\mathcal{D}$, represented in the right column of Fig. 5.2. The symbols s_i obtained are shaped by some transmit filter with impulse response $p(t)$. If a symbol is input to the system at time $t = iT$, the output of the transmit filter will be $p(t) \circledast \delta(t - iT) = p(t - iT)$. If the impulse response of the channel (note the notation from Section 2.1.3) is $h(\tau, t)$, then the received signal will be

$$r(t) = \int_{-\infty}^{\infty} h(\tau, t) p(t - iT - \tau) d\tau, \tag{5.1}$$

and the output of a receive filter of impulse response $p_r(t)$ will then be

$$\begin{aligned} y(t) &= r(t) \circledast p_r(t) \\ &= \int_{-\infty}^{\infty} r(\tau') p_r(t - \tau') d\tau' \end{aligned}$$

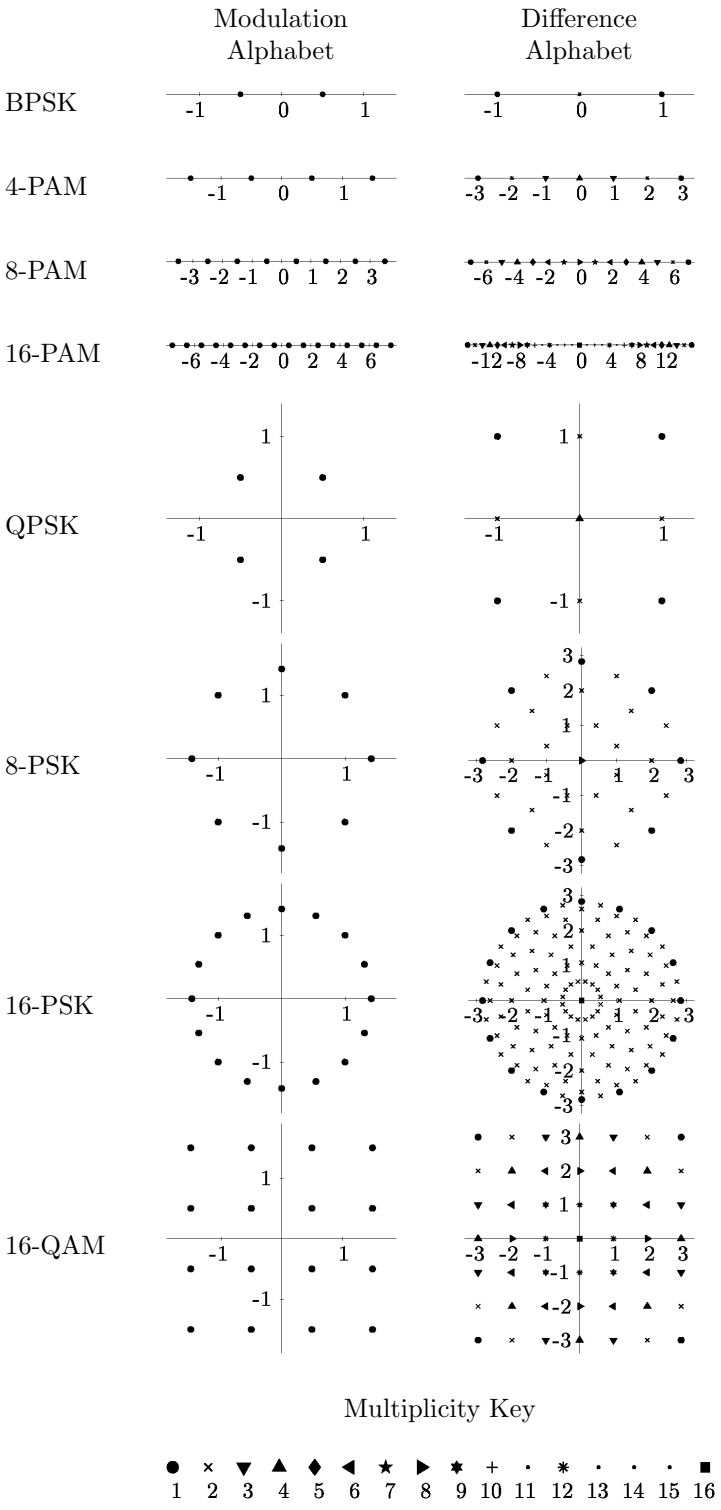


Figure 5.2: Modulation and Difference Alphabets

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_r(t-\tau')p(\tau'-\tau-iT)h(\tau,\tau')d\tau d\tau'. \tag{5.2}$$

Following symbol period sampling, the discrete time response at j to an impulse at i is then

$$f[j-i,j] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_r(jT-\tau')p(\tau'-\tau-iT)h(\tau,\tau')d\tau d\tau'. \tag{5.3}$$

The response of the system to the symbol sequence s_i, for a shift-invariant channel, with a discrete impulse response having L non-zero values, will then be

$$y_i = \sum_{l=1}^{L} f_i s_{i-l+1} + \eta_i \tag{5.4}$$

where η_i is a zero-mean complex normal noise, independent of the data s_i, which is assumed to have been whitened, so that $E\{\eta_j\} = 0$ and $E\{\eta_i\eta_j^*\} = \delta_{ij}\sigma_\eta^2$. (Some of the results to follow allow the noise to have a more general correlation).

Equation 5.4 may be represented in matrix form as

$$\mathbf{y} = \mathbf{F}\mathbf{s} + \boldsymbol{\eta} = \mathbf{S}\mathbf{f} + \boldsymbol{\eta}, \tag{5.5}$$

where

$$\begin{aligned}
\mathbf{y} &= (y_1, y_2, \ldots, y_R\,|, \ldots, y_{R+L-1})^T,\\
\mathbf{s} &= (s_{2-L}, \ldots, |\, s_1, s_2, \ldots, s_R)^T,\\
\boldsymbol{\eta} &= (\eta_1, \eta_2, \ldots, \eta_R\,|, \ldots, \eta_{R+L-1})^T,\\
\mathbf{f} &= (f_1, f_2, \ldots, f_L)^T,
\end{aligned}$$

$$\mathbf{F} = \left[\begin{array}{ccc|cccc}
f_L & f_{L-1} & \ldots & f_1 & 0 & \ldots & 0\\
0 & f_L & \vdots & f_2 & f_1 & \ldots & \vdots\\
\vdots & & \vdots & \vdots & \vdots & \ddots & \\
0 & \ldots & 0 & f_L & \vdots & & \vdots\\
\hline
0 & \ldots & 0 & 0 & f_L & & \vdots\\
\vdots & & \vdots & \vdots & & \ddots & f_{L-1}\\
0 & \ldots & 0 & 0 & 0 & 0 & f_L
\end{array}\right], \text{ and} \tag{5.6}$$

$$\mathbf{S} = \begin{bmatrix} s_1 & s_0 & \cdots & s_{L-2} \\ s_2 & s_1 & & \vdots \\ \vdots & \vdots & \ddots & \\ s_R & \vdots & & \vdots \\ s_{R+1} & s_R & & \vdots \\ \vdots & & \ddots & s_{R-1} \\ s_{R+L-1} & \cdots & \cdots & s_R \end{bmatrix}. \tag{5.7}$$

The vectors and matrices have been partitioned above to separate regions which may be included or omitted depending on whether it is appropriate in the context to assume that symbols preceding and following a block, $(s_{L-2}, \ldots, s_0)$, $(s_{R+1}, \ldots, s_{R+L-1})$, and $(y_{R+1}, \ldots, y_{R+L-1})$ are zero or non-zero.

When $(s_{L-2}, \ldots, s_0)$ and $(s_{R+1}, \ldots, s_{R+L-1})$ are zero, $\mathbf{F}$ is a $(R+L-1) \times R$ matrix given by the right two blocks of (5.6). The maximum likelihood data estimate is given by

$$\widehat{\mathbf{s}} = \arg\min_{\mathbf{s}} \left\{ \|\mathbf{y} - \mathbf{F}\mathbf{s}\| \,:\, \mathbf{s} \in \mathcal{A}^R \right\}. \tag{5.8}$$

Suppose that there exists an interval of R symbols such that the transmitted sequence $\mathbf{s}$ and estimated sequence $\widehat{\mathbf{s}}$ are identical before and after the interval, but differ in the positions 1 and R and some symbols in between. This is what has become known [49] as an *error event*. Define $\mathbf{e} = \widehat{\mathbf{s}} - \mathbf{s}$, $\mathbf{E} = \widehat{\mathbf{S}} - \mathbf{S}$, and $\boldsymbol{\epsilon} = \mathbf{F}\mathbf{e} = \mathbf{E}\mathbf{f}$. Note that $(e_{L-2}, \ldots, e_0)$ and $(e_{R+1}, \ldots, e_{R+L-1})$ are zero. The probability of the event is the probability that the maximum likelihood metric of $\widehat{\mathbf{s}}$ is smaller than that of $\mathbf{s}$:

$$\begin{aligned} P(\mathcal{E}) &= P\left(\|\mathbf{y} - \mathbf{F}\widehat{\mathbf{s}}\| \le \|\mathbf{y} - \mathbf{F}\mathbf{s}\|\right) \\ &= P\left(\|\mathbf{F}\mathbf{s} + \boldsymbol{\eta} - \mathbf{F}\widehat{\mathbf{s}}\| \le \|\mathbf{F}\mathbf{s} + \boldsymbol{\eta} - \mathbf{F}\mathbf{s}\|\right) \\ &= P\left(\|\boldsymbol{\epsilon} + \boldsymbol{\eta}\|^2 \le \|\boldsymbol{\eta}\|^2\right) \\ &= P\left(\|\boldsymbol{\epsilon}\|^2 + 2\,\mathrm{Re}(\boldsymbol{\epsilon}^H\boldsymbol{\eta}) \le 0\right). \end{aligned} \tag{5.9}$$

If the covariance of $\boldsymbol{\eta}$ is $\mathbf{C}_\eta = \sigma_\eta^2\mathbf{I}$, then the variance of $\boldsymbol{\epsilon}^H\boldsymbol{\eta}$ is

$$\mathrm{Var}\left\{\boldsymbol{\epsilon}^H\boldsymbol{\eta}\right\} = E\left\{(\boldsymbol{\eta}^H\boldsymbol{\epsilon})^H(\boldsymbol{\eta}^H\boldsymbol{\epsilon})\right\} = \boldsymbol{\epsilon}^H\mathbf{C}_{\boldsymbol{\eta}}\boldsymbol{\epsilon} = \|\epsilon\|^2\sigma_\eta^2. \tag{5.10}$$

The variance of the real part of this will thus be $\frac{1}{2}\|\epsilon\|^2\sigma_\eta^2$, and so $\mathrm{Var}\left\{2\,\mathrm{Re}(\boldsymbol{\epsilon}^H\boldsymbol{\eta})\right\} = 2\|\epsilon\|^2\sigma_\eta^2$. The probability of a real zero-mean normal random variable of variance σ_η^2 being less than k is $\mathrm{erfc}(-k/(\sqrt{2}\sigma_\eta))/2$, where the complementary error function

$\text{erfc}(x) = 1/(\sqrt{\pi}) \int_x^\infty e^{-t^2} dt$, so

$$P(\mathcal{E}) = \frac{1}{2} \text{erfc} \left(\frac{\|\boldsymbol{\epsilon}\|}{2\sigma_\eta} \right). \tag{5.11}$$

It is convenient to define the *distance* δ of a particular error event $\mathbf{e}$ for a particular channel $\mathbf{f}$ by [116, p 619]

$$\delta^2(\mathbf{e}, \mathbf{f}) = \frac{\|\boldsymbol{\epsilon}\|^2}{\|\mathbf{f}\|^2} = \frac{\|\mathbf{E}\mathbf{f}\|^2}{\|\mathbf{f}\|^2} = \frac{\mathbf{f}^H \mathbf{A} \mathbf{f}}{\mathbf{f}^H \mathbf{f}} \tag{5.12}$$

or by

$$\delta^2(\mathbf{e}, \mathbf{f}) = \frac{\|\mathbf{F}\mathbf{e}\|^2}{\|\mathbf{f}\|^2} = \frac{\mathbf{e}^H \boldsymbol{\Phi} \mathbf{e}}{\mathbf{f}^H \mathbf{f}}, \tag{5.13}$$

where the square Hermitian Toeplitz matrices $\mathbf{A}$ and $\boldsymbol{\Phi}$ are $\mathbf{A} = \mathbf{E}^H \mathbf{E}$, and $\boldsymbol{\Phi} = \mathbf{F}^H \mathbf{F}$. For the purpose of comparing channels of different impulse responses, it is also useful to confine attention to channels having the same power, so frequently it is assumed that the channel is "normalised" so that $\mathbf{f}^H \mathbf{f} = 1$, and consequently $\delta(\mathbf{e}, \mathbf{f}) = \|\boldsymbol{\epsilon}\|$.

The overall probability of a symbol error is bounded and well approximated by the weighted sum over all possible error events of the pair-wise error probabilities given by (5.11):

$$P_e(\mathbf{f}) \leq \sum_{\mathbf{e} \in \mathcal{D}^R} \frac{1}{2} \overline{w}(\mathbf{e}) \, \text{erfc} \left(\frac{\delta(\mathbf{e}, \mathbf{f})}{2\sigma_\eta} \right) \tag{5.14}$$

where $\overline{w}(\mathbf{e})$ is the average *multiplicity* and error weight of $\mathbf{e}$. The multiplicity is the proportion of time that the transmitted sequence makes error event $\mathbf{e}$ possible. The diagrams on the right hand side of Fig. 5.2 show the number of possible ways that a single symbol error can occur i.e., $w(e_r)$. The multiplicity of a longer error event is then $w(\mathbf{e}) = \prod_{r=1}^R w(e_r)$, assuming that each data symbol is equally likely in each symbol period. The error weight is simply the number of errors a particular error event causes. Note that for a given value of R, some error events may occur several times, and this must be taken into account in calculating the multiplicity of each error event.

5.2 Reduced Complexity Error Rate Calculation

5.2.1 Minimum Distance Error Events

Equation 5.14 must be evaluated over all possible error events, which is an infinite sum, since there is in principle no limit to the length of error event R. Even a summation for a moderate finite length R may be impractical, especially for the larger modulation alphabets.

One approximation to the error rate is given by the rate for the *worst case* channel. The pairwise error probability of error of (5.11) is largest when $\delta(\mathbf{e}, \mathbf{f})$ is smallest. From (5.12) this clearly occurs when the channel $\mathbf{f}$ is the eigenvector corresponding to the minimum eigenvalue of $\mathbf{A}(\mathbf{e})$.

The search through possible error events for the *minimum distance* error event, which minimises the minimum eigenvalue of $\mathbf{A}(\mathbf{e})$ has been the subject of some investigation [87, 132]. The trouble with this analysis is that the resulting error rate estimate is a very poor and pessimistic estimate. As has already been mentioned, the difference in power between the worst-case and best-case channels required to achieve the same rate can be very large. For a length $L = 7$ QPSK channel the difference is 11.4 dB. The difference is larger for longer impulse responses and for more spectrally efficient modulations.

5.2.2 Contributing Error Events

The very "short" tail of the normal distribution results in the error probability of (5.14) being well approximated (for medium to high SNR) by the single term of the sum which has the largest likelihood of occurrence. The approximation is even better if the summation is over those *few* terms which have the largest likelihood. This term (or terms) can be found rapidly if it is known to belong to a small set $\mathcal{U}$ of *significant* error events. The set $\mathcal{U}$ is the smallest set of error events which consists of those $\mathbf{e}$ so that for all $\mathbf{e}_N \notin \mathcal{U}$, and for all $\mathbf{f}$, there exists an $\mathbf{e}_U \in \mathcal{U}$ such that

$$\delta\left(\mathbf{e}_U, \mathbf{f}\right) \leqslant \delta\left(\mathbf{e}_N, \mathbf{f}\right). \tag{5.15}$$

Note that the same $\delta\left(\mathbf{e}_U, \mathbf{f}\right)$ does not need to be less than $\delta\left(\mathbf{e}_N, \mathbf{f}\right)$ for every value of $\mathbf{f}$. There needs only to be *for a given* $\mathbf{f}$, some $\mathbf{e}_U \in \mathcal{U}$ such that the condition is satisfied.

Once such a set $\mathcal{U}$ of error events has been found, the error performance of the system can be easily approximated by finding the error rate caused by the most probable element of that set. The method of obtaining the performance estimate for a particular channel $\mathbf{f}$ is thus to evaluate (5.14) for the $\mathbf{e}_1 \in \mathcal{U}$ for which $\delta(\mathbf{e}_1, \mathbf{f})$ is smallest.

An iso-surface of the exact error performance for the modulation Binary Phase Shift Keying (BPSK), and $L = 3$ is shown in Fig. 5.3(a). All points on the surface represent channels with a symbol error probability of 3×10^{-5}. An approximation to this surface based on just three error sequences is shown in Fig. 5.3(b).

Note that for a fixed distance δ such as unity, (5.12) defines an ellipsoid in the space of impulse response coefficients $\mathbf{f}$. The axes of this ellipsoid are aligned with the eigenvectors of $\mathbf{A}$, and the lengths of the semi-axes are the inverse of the square roots of the eigenvalues of $\mathbf{A}$ (i.e., the inverse of the singular values of $\mathbf{E}$). Since $\mathbf{A}$ is Hermitian, its eigenvalues are real. An error event consisting of a single symbol error defines a sphere.

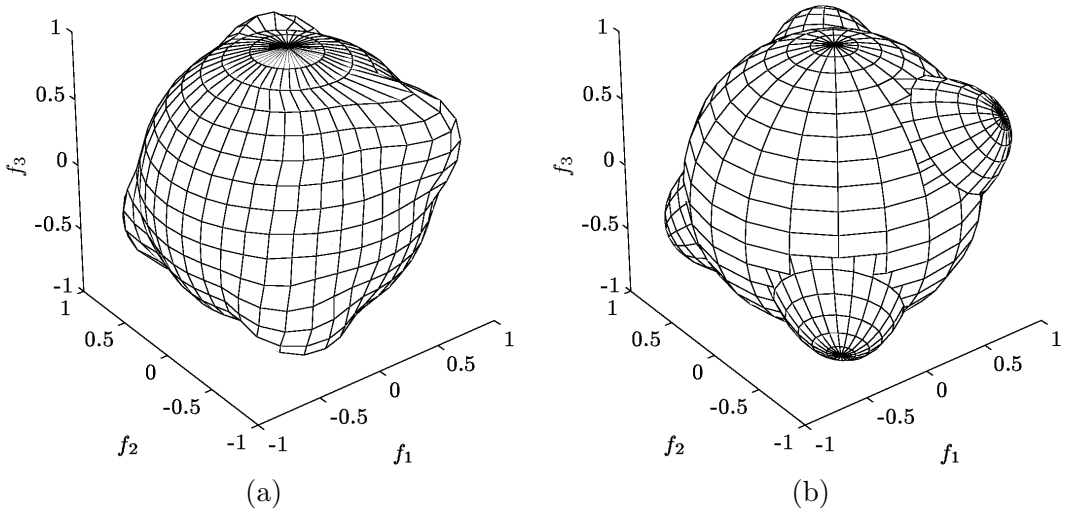


Figure 5.3: Iso-surface of bit error rate for BPSK (left) and an approximation based on the union of ellipsoids (right). The three axes represent the impulse response components of a discrete channel of length three. The error rate probability is 3×10^{-5} and the SNR of the best channel is 11 dB. Each of the ellipsoids corresponds to a particular error event.

The set $\mathcal{U}$ defined in (5.15) can be visualised as the minimum set of ellipsoids required to contain the union of *all* ellipsoids.

5.3 Algorithm for finding the set $\mathcal{U}$

The set of error events required to approximate the error of *all* error events can be found using the algorithm described in this section.

5.3.1 Stage 1

The first stage is, for a given length of error event R, to search through all possible sequences, and eliminate duplication of the matrix $\mathbf{A}$. There are several error sequences which will produce the same $\mathbf{A}$. For large R this is quite computationally intensive, particularly for large difference alphabets (e.g., 16-PSK has 129 elements in the difference alphabet). The computation time rises exponentially with R.

Sequences containing (but not ending with) a sequence of zeros of length $\geq L-1$ need not be included. This is because after such a sequence of zero errors (i.e., correct transmissions), a MLSE equaliser will be in the all zero state, and so the behaviour is already covered by another sequence (although as mentioned in Section 5.1 this duplication must be allowed for in calculating the multiplicity of the error event).

Restricting the search of error events to only those of length R or less may mean that a valid member of the set is omitted. Provided however that R is chosen to be sufficiently large, the inaccuracy in the performance approximation will be small [87,

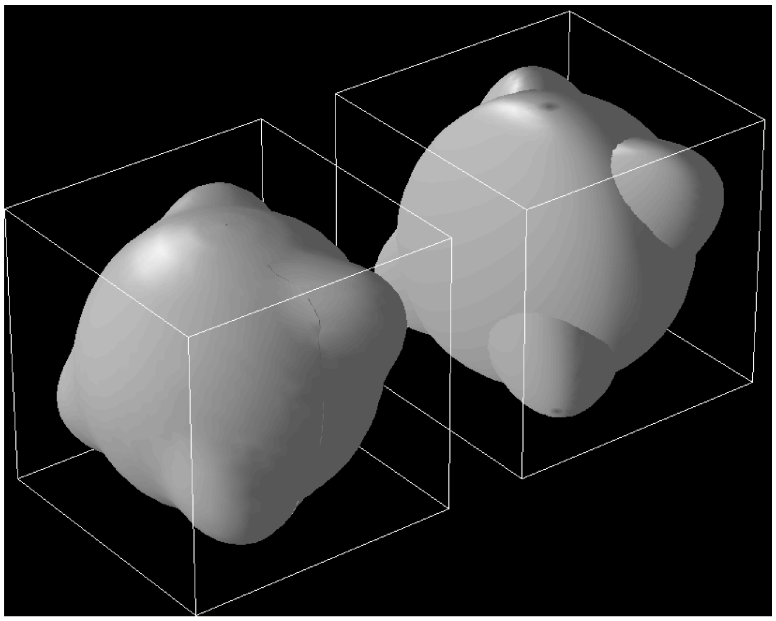


Figure 5.4: Another representation of the resemblance between the actual error rate and the union of ellipsoids approximation.

pp22,122]. In fact it is shown in [87] that very long error events, although having a very small distance δ may only occur under very specific conditions in the transmitted data sequence which have such small probability of occurrence that their contribution to the overall error probability is negligible. The probability of their occurrence becomes even less significant if the channel is not static.

5.3.2 Stage 2

The second stage of the algorithm is to eliminate all of the matrices $\mathbf{A}$ which *dominate* any others.

The following definition of positive semidefinite partial ordering from [73, p469] is used

Definition 1 *Let* $\mathbf{A}_1, \mathbf{A}_2$ *be* $n \times n$ *Hermitian matrices. We write* $\mathbf{A}_1 \succ \mathbf{A}_2$ *if the matrix* $\mathbf{A}_1 - \mathbf{A}_2$ *is positive definite.*

If $\mathbf{A}_1 \succ \mathbf{A}_2$, then for any $n \times m$ matrix $\mathbf{T}$, $\mathbf{T}^H \mathbf{A}_1 \mathbf{T} \succ \mathbf{T}^H \mathbf{A}_2 \mathbf{T}$. If $\mathbf{T}$ is a vector, $m = 1$, and the relation $\succ$ is simply $>$. The positive definiteness of $\mathbf{A}_1 - \mathbf{A}_2$ can be established by checking that all its eigenvalues are positive [73, p402].

The analogy with the ellipsoids is that the ellipsoid defined by $\mathbf{A}_1$ is entirely contained within that defined by $\mathbf{A}_2$, and so it is not a member of that minimum set required to define the union.

Since pairs of sequences are being compared in this stage, the computation time required here rises with the square of the number of matrices $\mathbf{A}(\mathbf{e})$ found in stage 1.

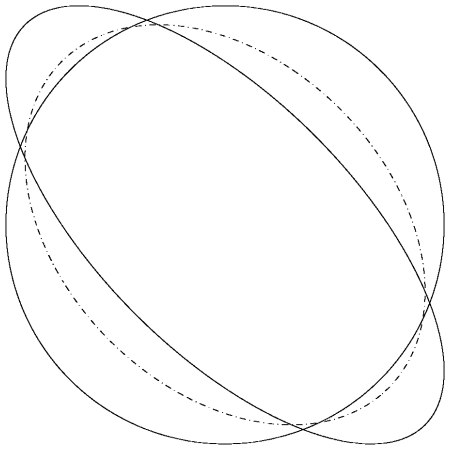


Figure 5.5: An ellipse (broken line) which is not contained by either of two others, and yet does not contribute to the union.

Hence this stage is also quite computationally intensive. However, many of the matrices can be quickly eliminated by comparison with the largest sphere; if the smallest singular value is larger than the radius of the sphere, further search is not required.

5.3.3 Stage 3

Having eliminated all sequences which define ellipsoids which are entirely contained within others still leaves many others, which, while they are not contained by any other *single* ellipsoid, are still entirely contained by the union of all ellipsoids, and do not add to that union. An illustration of this behaviour is shown in Fig. 5.5.

The smallest set which *does* define the union is generally much smaller than the set of matrices which do not dominate any others. Hence the third stage of the algorithm is to perform further elimination so the smallest set required to define the union remains.

One method of performing this elimination is to find a set of points on the unit (hyper-) sphere in the space of $\mathbf{f}$, and perform a search over the surface of the sphere to find the sequence with the smallest distance δ at every point. This will find the smallest set required to define the union provided there isn't too much numerical inaccuracy, and the grid is fine enough. Unfortunately for dimensions greater than $\mathbb{R}^{11}$ or $\mathbb{C}^5$ the operation becomes too time consuming with the computational power available. The search grid consequently becomes rather coarse, and the probability that a potentially significant error event is missed increases.

Another method which will eliminate some sequences, but perhaps not necessarily result in the smallest *possible* set, is to consider triples of error sequences. Consider three matrices $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$ corresponding to three different error events $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ respectively, and $\delta_i^2(\mathbf{f}) = \mathbf{f}^H \mathbf{A}_i \mathbf{f}$ for $i = 1, 2, 3$. The intersections of the pairs $\delta_i^2 = \delta_j^2$ all exist since otherwise one of the matrices would have been eliminated in stage 2. If

however

$\delta_1^2 < \delta_3^2$ for all the points of the intersection $\delta_2^2 = \delta_3^2$, *and*

$\delta_2^2 < \delta_3^2$ for all the points of the intersection $\delta_1^2 = \delta_3^2$

then $\mathbf{e}_3$ cannot belong to the set $\mathcal{U}$.

Each of the requirements above can be established in the following manner. For the first requirement, for example, it must be shown that the minimum of $\mathbf{f}^H(\mathbf{A}_1-\mathbf{A}_3)\mathbf{f} < 0$, subject to the constraint that $\mathbf{f}^H(\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f} = 0$, and the constraint that $\mathbf{f}^H\mathbf{f} = 1$. Using the method of Lagrange multipliers, the solution is points at which the derivative of augmented function

$$\chi = \mathbf{f}^H(\mathbf{A}_1 - \mathbf{A}_3)\mathbf{f} - \nu_1\left(\mathbf{f}^H(\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}\right) - \nu_2\left(\mathbf{f}^H\mathbf{f}\right) \tag{5.16}$$

is zero. Taking the derivative

$$\frac{\partial\chi}{\partial\mathbf{f}} = 2\left[(\mathbf{A}_1 - \mathbf{A}_3) - \nu_1(\mathbf{A}_2 - \mathbf{A}_3) - \nu_2\mathbf{I}\right]\mathbf{f} \tag{5.17}$$

and so at the solution

$$\left[(\mathbf{A}_1 - \mathbf{A}_3) - \nu_1(\mathbf{A}_2 - \mathbf{A}_3)\right]\mathbf{f} = \nu_2\mathbf{f}. \tag{5.18}$$

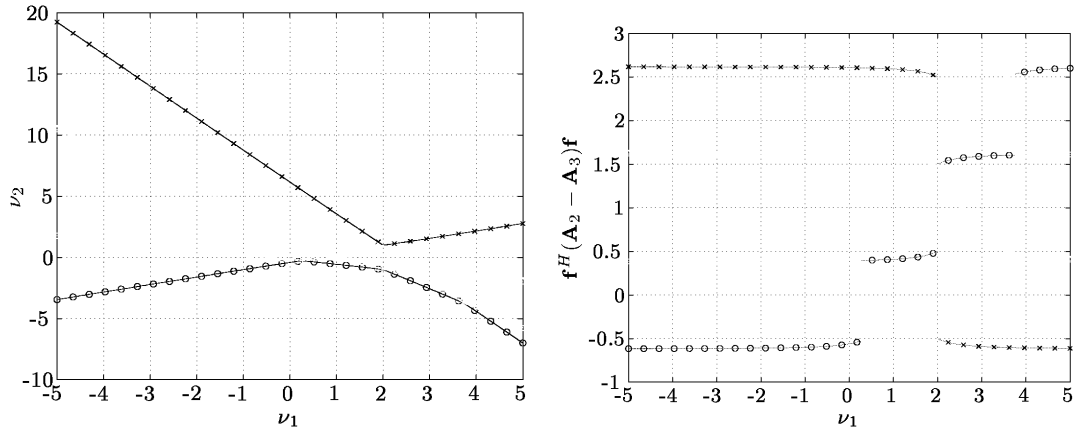
In other words, ν_2 is an eigenvalue and $\mathbf{f}$ an eigenvector of the matrix $(\mathbf{A}_1 - \mathbf{A}_3) - \nu_1(\mathbf{A}_2 - \mathbf{A}_3)$. The solution to (5.16) can thus be obtained by finding the value of ν_1 for which the normalised eigenvector corresponding to the maximum eigenvalue of $(\mathbf{A}_1 - \mathbf{A}_3) - \nu_1(\mathbf{A}_2 - \mathbf{A}_3)$ satisfies $\delta_{23} = \mathbf{f}^H(\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f} = 0$. When this solution is found

$$\begin{aligned}\mathbf{f}^H(\mathbf{A}_1 - \mathbf{A}_3)\mathbf{f} &= \mathbf{f}^H(\mathbf{A}_1 - \mathbf{A}_3)\mathbf{f} - \nu_1\left(\mathbf{f}^H(\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}\right)\\ &= \nu_2\mathbf{f}^H\mathbf{f}\\ &= \nu_2,\end{aligned} \tag{5.19}$$

where the first and second lines follow from the constraint conditions. Thus if the value of the eigenvalue ν_2 at this point is less than zero, the condition is satisfied.

It may happen that there is no value of ν_1 for which an eigenvector is a solution to $\delta_{23} = 0$. It is then sufficient to find *any* solution to $\delta_{23} = 0$, and find the sign of $\mathbf{f}^H(\mathbf{A}_1 - \mathbf{A}_3)\mathbf{f}$.

There will frequently be a discontinuity in $\delta_{23} = \mathbf{f}^H(\mathbf{A}_2-\mathbf{A}_3)\mathbf{f}$ for the value at which the maximum (or minimum) eigenvalue has algebraic multiplicity of two. Since at any point near this discontinuity there are two distinct eigenvectors, it seems reasonable to suppose (and it was found in practice) that the geometric multiplicity of the eigenvalue was also two. In other words [73, p58] there will be two eigenvectors which span a space of dimension two, and hence the value of δ_{23} may assume a range of values. If



(a) (b)

Figure 5.6: Solution to the constrained optimisation using Lagrange multipliers. The Lagrange multiplier ν_2 (a) and the product $\mathbf{f}(\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}$ (b) are presented as functions of the multiplier ν_1. The values corresponding to the maximum eigenvalue are represented by the crosses and to the minimum eigenvalue by circles. Here $L = 4$, $\mathbf{e}_1 = (1, -1, 1, 0, 0, 0)^T$, $\mathbf{e}_2 = (1, -1, 0, 0, 0)^T$, and $\mathbf{e}_3 = (1, 0, 0, 0)^T$.

the values an each side of the discontinuity are opposite in sign (as shown in Fig. 5.6) it may not be obvious that δ_{23} may take the value 0 and hence be a solution to the equation.

In all the situations tried however, the following was found to be true. Let ν_d be the value of ν_1 at which a discontinuity occurs, let ϵ be a small positive real value, let $\mathbf{f}_1$ be the normalised eigenvector corresponding to the minimum eigenvalue for $\nu_1 = \nu_d - \epsilon$, and let $\mathbf{f}_2$ be the normalised eigenvector corresponding to the minimum eigenvalue for $\nu_1 = \nu_d + \epsilon$. It was found that the space spanned by $\mathbf{f}_1$ and $\mathbf{f}_2$ as $\epsilon \to 0$ is the same as the eigenspace of the minimum eigenvalue at $\nu_1 = \nu_d$. A normalised vector in this space has the form

$$\mathbf{v} = \alpha\mathbf{f}_1 \pm \sqrt{1-\alpha^2}\mathbf{f}_2 \tag{5.20}$$

where $\alpha \in [-1, 1]$ and so

$$\begin{aligned}\delta_{23} &= \left(\alpha\mathbf{f}_1 \pm \sqrt{1-\alpha^2}\mathbf{f}_2\right)^H (\mathbf{A}_2 - \mathbf{A}_3)\left(\alpha\mathbf{f}_1 \pm \sqrt{1-\alpha^2}\mathbf{f}_2\right) \\ &= \alpha^2\mathbf{f}_1^H (\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}_1 + (1-\alpha^2)\mathbf{f}_2^H (\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}_2 \\ &\quad \pm \alpha\sqrt{1-\alpha^2}\left(\mathbf{f}_1^H (\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}_2 + \mathbf{f}_2^H (\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}_1\right).\end{aligned} \tag{5.21}$$

Not only are $\mathbf{f}_1$ and $\mathbf{f}_2$ orthogonal, but $\mathbf{f}_1^H (\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}_2 = 0$. Hence at the point of discontinuity δ_{23} may take on the values of $\mathbf{f}_1(\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}_1$ (when $\alpha = \pm 1$) or $\mathbf{f}_2(\mathbf{A}_2 - \mathbf{A}_3)\mathbf{f}_2$ (when $\alpha = 0$) and any value in between. The range of δ_{23} is thus precisely those values between the end points of the discontinuity. If these values are opposite in sign,

the range includes 0.

In principle this technique could be extended to take into account four or more error events. However, the number of Lagrange multipliers was found to prevent simple computation. In practice, the last stage of elimination was performed by extensive multidimensional search. The possibility remains that some sequences have been missed.

5.4 Symmetry of Contributing Error Events

The algorithm described in Section 5.3 was implemented and tried for various channel lengths L, various sequence lengths R and various modulation schemes (see Appendix A). It was then observed that all of the sequences required to estimate the error of a system had a symmetry property. The nature of the symmetry is elaborated in the following conjecture.

Conjecture 1 *Let* $\mathbf{K}$ *be the* $R \times R$ *matrix with elements* $[\mathbf{K}]_{ij} = \delta_{i,R+1-j}$. *All of the length* R *members* $\mathbf{e}$ *of the set* $\mathcal{U}$ *which define the union of ellipsoids have the symmetry property that for some real* θ, $\mathbf{e} = e^{j\theta}\mathbf{K}\mathbf{e}^*$

Premultiplication by $\mathbf{K}$ reverses the order of the elements of the vector $\mathbf{e}$, so that the conjecture in effect says that the second half of a sequence $\mathbf{e}$ which has the minimum distance for a certain channel is the mirror image of the conjugate of the first half, to within some rotation allowed by the set of symbols in the difference alphabet. (The rotational symmetry of the difference alphabet turns out to be the same as that of the modulation alphabet).

The conjecture is made more credible by the fact that the minimiser of (5.13) over a continuous set of complex vectors has this symmetry property. In the continuous case the minimisation is constrained to be over vectors having norm of 1, or alternatively the minimisation is of a Rayleigh quotient

$$\frac{\mathbf{e}^H\mathbf{\Phi}\mathbf{e}}{\mathbf{e}^H\mathbf{e}}. \tag{5.22}$$

The matrix $\mathbf{\Phi} = \mathbf{F}^H\mathbf{F}$ is clearly Hermitian and so has real eigenvalues [73, p170]. The Toeplitz structure of $\mathbf{F}$ results in $\mathbf{\Phi}$ being also Toeplitz. Matrices which are both Toeplitz and Hermitian are members of a larger class of matrices symmetric about both main diagonals, called persymmetric matrices [62, p193] for which

$$\mathbf{K}\mathbf{\Phi} = \mathbf{\Phi}^*\mathbf{K}. \tag{5.23}$$

where $\mathbf{K}$ is as defined in Conjecture 1. If $\mathbf{v}$ is some eigenvector of $\mathbf{\Phi}$ corresponding to a unique eigenvalue ϱ, then $\mathbf{\Phi v} = \varrho \boldsymbol{v}$, and so

$$\begin{aligned} \mathbf{K\Phi v} &= \varrho \mathbf{K v} \\ \mathbf{\Phi^* K v} &= \varrho \mathbf{K v} \\ \mathbf{\Phi K v^*} &= \varrho \mathbf{K v^*} \end{aligned} \tag{5.24}$$

Clearly $\mathbf{Kv}^*$ is an eigenvector of $\mathbf{\Phi}$ corresponding to the real eigenvalue ϱ. Since this eigenvalue is distinct, and $\mathbf{v}$ has the same norm as $\mathbf{Kv}^*$, it follows that

$$\mathbf{v} = e^{j\theta} \mathbf{K v^*} \tag{5.25}$$

for some angle θ which is not necessarily the same for all of the eigenvectors. The symmetry property still holds when the eigenvalues are not distinct, although this is not so simple to prove [87].

Since the minimiser of the Rayleigh quotient is an eigenvector of $\mathbf{\Phi}$, it must have the same symmetry property. For minimisation over a discrete set (one of the difference alphabets of Fig. 5.2), there is the additional constraint that the elements of $\mathbf{e}$ belong to the discrete set, but the norm 1 constraint is removed. It is not obvious that the symmetry property necessarily holds, although it frequently does hold.

There is an additional complication: the minimisers appear to have the symmetry property only for their non-zero portion. Specifying a channel $\mathbf{f}$ specifies a family of square matrices $\mathbf{\Phi}$ of different sizes $R \times R$, the smaller ones of which are sub-matrices of the larger ones. Since they are Toeplitz, they all have the persymmetric property. In searching for a error event $\mathbf{e}$ which has the smallest distance $\delta(\mathbf{e}, \mathbf{f})$ for a particular channel $\mathbf{f}$, the length R of $\mathbf{e}$, and hence the size of $\mathbf{\Phi}$, is free to change.

The symmetry property is not only of academic interest. If all of the error events with minimum distance have this symmetry, then the range of value of R which can be searched is considerably extended. The total number of error events of length R is

$$N_n = (N_s - 1)^2 N s^{R-2} \tag{5.26}$$

where N_s is the number of symbols in the alphabet. The number of symmetric error events for even R is

$$N_m = (N_s - 1) N_s^{\frac{R-2}{2}} N_q \tag{5.27}$$

where N_q is the number of angles as rotational symmetry of the difference alphabet.

When R is odd, the centre error symbol of a symmetric sequence can only be one for which a rotation is the same as the conjugate. For most difference alphabets used, including QPSK, but excluding 16-QAM and higher orders of QAM, for each difference

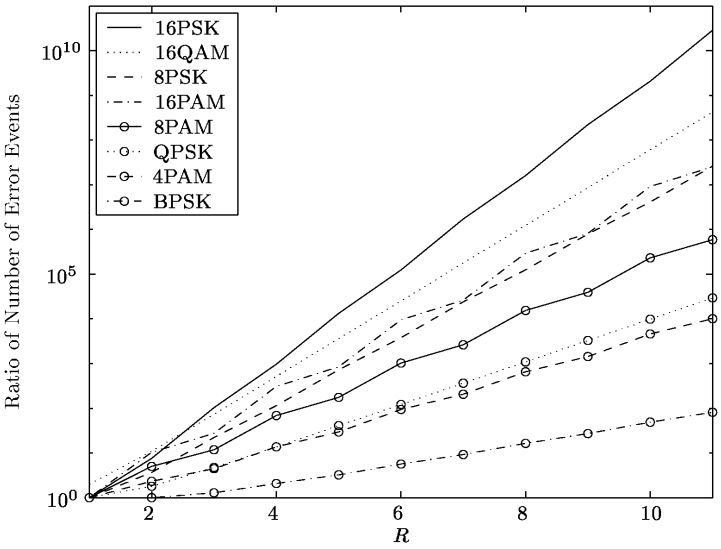


Figure 5.7: The Ratio of All Error Events of length $\leq R$ to those which are symmetric

symbol there is exactly one rotation (allowed by the symmetry of the alphabet) which will have the same effect as conjugation. For 16-QAM there are some difference symbols for which no rotations have the same effect as conjugation. The difference 0 will allow any rotation. If the number of non-zero difference symbols which have a conjugate rotation is N_r, then the number of symmetric error events of length R is

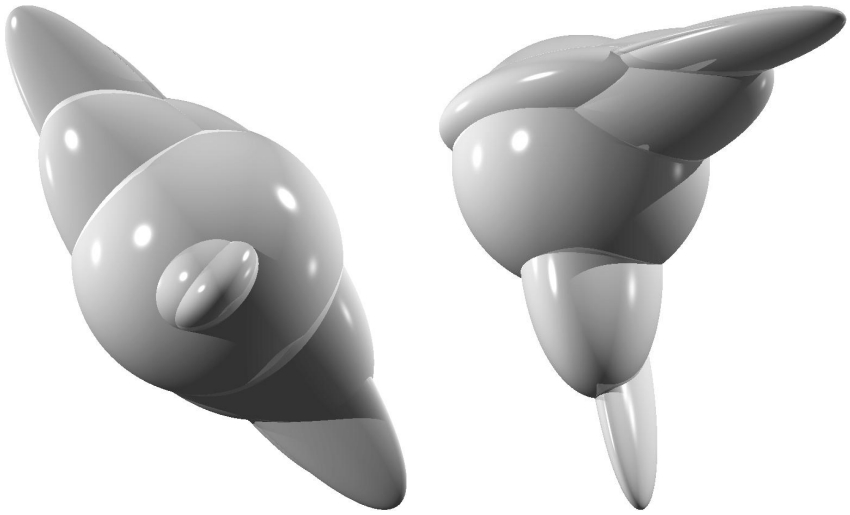
$$N_m = (N_s - 1)N_s^{\frac{R-1}{2}}(N_q + N_r - 1). \tag{5.28}$$

The ratio of the total number of error events less than or equal to R, to the number of error events having the conjugate symmetric property is shown in Fig. 5.7. Clearly R can be considerably extended by searching only the symmetric error events.

In a real system however, error events with large values of R will not really be significant. Such an error event will have the lowest distance for a set of channels which have a very small probability of occurrence. For the error event to contribute significantly the channel would have to remain within this small set of channels for the duration of at least R times the symbol period. For systems of even moderate rates of change, this will preclude long error events.

Considerable time and effort was spent in the search for either a proof or a counter example for the conjecture (over 6 months of 500MHz Intel processor time was spend in searching for counter examples). The following paragraphs are a disappointingly trivial report to show for the effort. Eventually several counter examples were found. These are listed in Appendix A in Section A.8.

The difficulty of stating results with confidence here is highlighted by the fact that it cannot yet be stated with certainty that these *are* in fact counter examples — only



(a) (b)

Figure 5.8: Union of Ellipsoids for 4-PAM, L=4. In (a) $\theta = 0.86\pi$ (equation 5.39). In (b) $f_4 = -0.59$ (equation 5.30).

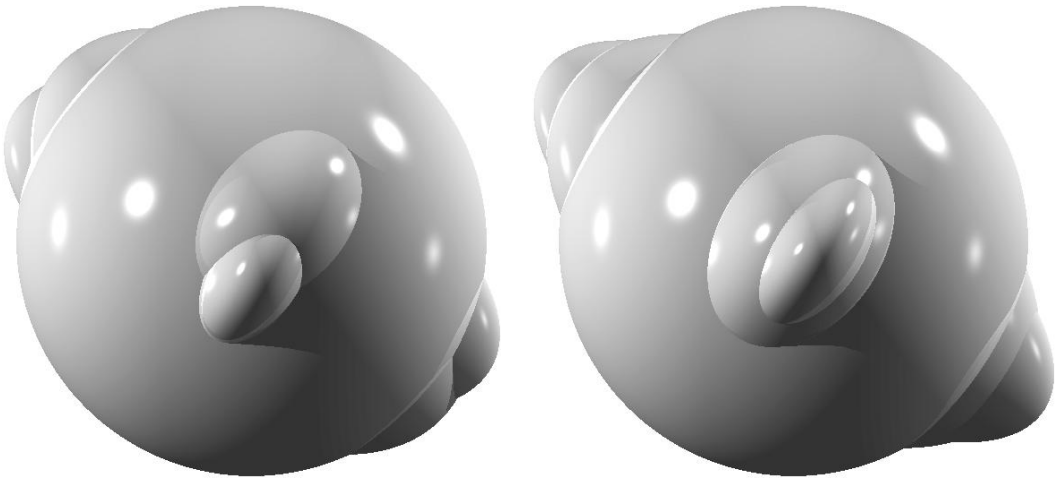
that for some channels these non-symmetric error events have a distance less than any symmetric error events of the length searched. There could be a longer sequence with lower distance which *does* have the symmetry property. The issue of channel variation of course makes sequences of large R only of academic interest, as mentioned above.

The counter examples (if indeed they are counter examples) were only found for the modulation schemes 8-PSK and 16-PSK. This leaves the tantalising possibility that the conjecture does hold under some conditions. It is somewhat surprising that these two modulation schemes should furnish the counter examples, since these have more "dense" difference alphabets, which one might have expected to more closely approximate the continuous case, where it is known that the minimising vector *does* have the symmetry property.

5.5 Visualisation

There are very few error events which contribute to the union for channels for which $L = 3$. For channels with more interesting behaviour, the number of dimensions (the number of discrete time channel coefficients) is greater than 3, and so visualisation of the union of ellipsoids is difficult. Some diagrams have been produced to enable some insight into the structure of the union in higher dimensions. These diagrams involve taking a two or three dimensional slice through a higher dimensional object.

If $\mathbf{A}$ is a real symmetric matrix with positive eigenvalues, then the equation $\mathbf{f}^H\mathbf{A}\mathbf{f} = 1$ describes an ellipsoid centred on the origin. For most of the matrices $\mathbf{A}$ described in this thesis $\mathbf{A}$ is also Toeplitz. If the first row of such a matrix is $[r_1, r_2, \ldots, r_L]$, then the equation $\mathbf{f}^H\mathbf{A}\mathbf{f} = 1$ can be expanded for some lower dimensional cases as:



(a) (b)

Figure 5.9: Union of Ellipsoids for QPSK, L=3, equation (5.41). In (a) $\theta_1 = -0.68\pi$, $\theta_2 = 0.05\pi$, $\theta_3 = 0$. In (b) $\theta_1 = 0.72\pi$, $\theta_2 = 0.86\pi$, $\theta_3 = 0$.

3-D:

$$r_1(f_1^2 + f_2^2 + f_3^2) + 2r_2(f_1f_2 + f_2f_3) + 2r_3f_1f_3 = 1 \tag{5.29}$$

4-D:

$$r_1(f_1^2 + f_2^2 + f_3^2 + f_4^2) + 2r_2(f_1f_2 + f_2f_3 + f_3f_4) + 2r_3(f_1f_3 + f_2f_4) + 2r_4f_1f_4 = 1 \tag{5.30}$$

5-D:

$$\begin{aligned} &r_1(f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5^2) + 2r_2(f_1f_2 + f_2f_3 + f_3f_4 + f_4f_5) \\ &+2r_3(f_1f_3 + f_2f_4 + f_3f_5) + 2r_4(f_1f_4 + f_2f_5) + 2r_5(f_1f_5) = 1. \end{aligned} \tag{5.31}$$

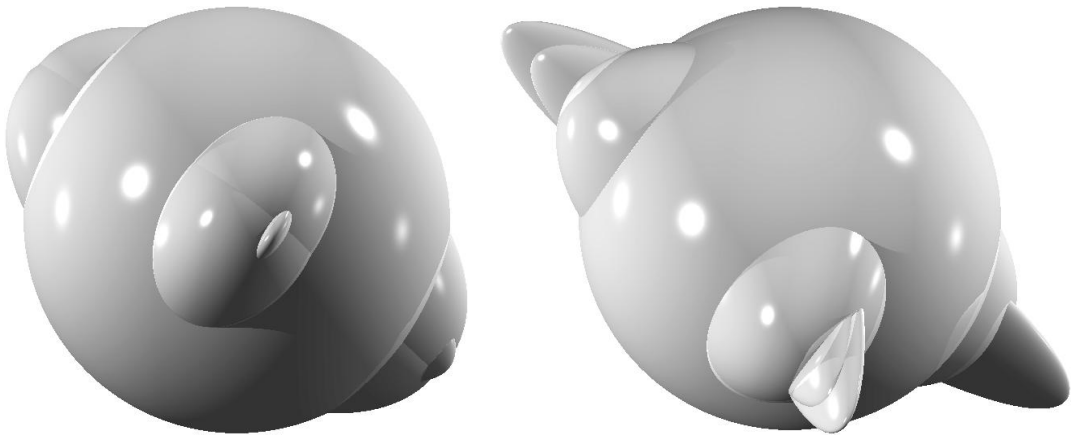
If some of the coefficients f_i are fixed, these expressions can be rearranged as an equation for an ellipse:

$$1 = p_1(x^2 + y^2) + 2p_2xy + 2p_3x + 2p_4y + p5 \tag{5.32}$$

or a three dimensional ellipsoid:

$$1 = p_1(x^2 + y^2 + z^2) + 2p_2xy + 2p_3xz + 2p_4yz + 2p_5x + 2p_6y + 2p_7z + p8. \tag{5.33}$$

where x, y, z are those coefficients which are not fixed.



(a) (b)

Figure 5.10: Union of Ellipsoids for 8-PSK, L=3, equation (5.41). In (a) $\theta_1 = -0.26\pi$, $\theta_2 = -0.58\pi$, $\theta_3 = 0$. In (b) $\theta_1 = 0.79\pi$, $\theta_2 = 0.89\pi$, $\theta_3 = 0$.

5.5.1 Three Dimensional Slices With One Co-ordinate Fixed

If an ellipse described by the quadratic form $\mathbf{x}^T\mathbf{A}_v\mathbf{x}$ where

$$\mathbf{A}_v = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_2 & a_1 & a_4 \\ a_3 & a_4 & a_1 \end{bmatrix}, \tag{5.34}$$

is translated to the point $\mathbf{x} = (x_0, y_0, z_0)^T$, the resulting equation is

$$\begin{aligned} 1 = {} & a_1\left(x^2 + y^2 + z^2\right) + 2a_2xy + 2a_3xz + 2a_4yz \\ & - 2x\left(a_1x_0 + a_2y_0 + a_3z_0\right) - 2y\left(a_2x_0 + a_1y_0 + a_4z_0\right) - 2z\left(a_3x_0 + a_4y_0 + a_1z_0\right) \\ & + a_1\left(x_0^2 + y_0^2 + z_0^2\right) + 2a_2x_0y_0 + 2a_3x_0z_0 + 2a_4y_0z_0. \end{aligned} \tag{5.35}$$

Equating the coefficients with those of (5.33):

$$\begin{aligned} k &= \frac{a_1}{p_1} = \frac{a_2}{p_2} = \frac{a_3}{p_3} = \frac{a_4}{p_4} \\ &= \frac{a_1x_0 + a_2y_0 + a_3z_0}{-p_5} = \frac{a_2x_0 + a_1y_0 + a_4z_0}{-p_6} = \frac{a_3x_0 + a_4y_0 + a_1z_0}{-p_7} \\ &= \frac{1 - a_1(x_0^2 + y_0^2 + z_0^2) - 2(a_2x_0y_0 + a_3x_0z_0 + a_4y_0z_0)}{1 - p_8}. \end{aligned} \tag{5.36}$$

This equation can be solved to obtain

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = -\begin{bmatrix} p_1 & p_2 & p_3 \\ p_2 & p_1 & p_4 \\ p_3 & p_4 & p_1 \end{bmatrix}^{-1} \begin{bmatrix} p_5 \\ p_6 \\ p_7 \end{bmatrix} \tag{5.37}$$

and

$$k = \frac{1}{1 - p_8 + p_1(x_0^2 + y_0^2 + z_0^2) + 2(p_2 x_0 y_0 + p_3 x_0 z_0 + p_4 y_0 z_0)} \tag{5.38}$$

The coefficients for a_1, a_2, a_3 and a_4 are then obtained by multiplying the corresponding p_i by k.

Many drawing packages allow a co-ordinate transformation matrix to be specified. Points on the ellipsoid defined by $\mathbf{A}_v$ satisfy $\mathbf{x}^T\mathbf{A}_v\mathbf{x} = 1$. If the matrix of orthonormal eigenvectors of $\mathbf{A}_v$ is $\mathbf{V}$, and the corresponding diagonal matrix of eigenvalues is $\mathbf{D}$, then it is easy to show that the matrix $\mathbf{V}\mathbf{D}^{-\frac{1}{2}}$ transforms the unit sphere to the required ellipsoid. A three dimensional slice produced in this way for 4-PAM is shown in Fig.5.8(b).

5.5.2 Rotating Hyper-plane Three Dimensional Slice

An alternative visualisation can be obtained by viewing the intersection of the union with a three dimensional hyper-plane which passes through the origin and rotates around one of the axes (note that once again the intersection will be an ellipsoid, but now centred on the origin). For example, if $f_1 = x$, $f_2 = y$, $f_3 = z\cos\theta$ and $f_4 = z\sin\theta$, then (5.30) leads to

$$\begin{aligned} &r_1(x^2 + y^2 + z^2) + 2r_2(xy + yz\cos\theta + z^2\sin\theta\cos\theta) \\ &\quad + 2r_3(xz\cos\theta + yz\sin\theta) + 2r_4 xz\sin\theta = 1. \end{aligned} \tag{5.39}$$

By equating coefficients of the terms in x, y and z, the matrix describing the resulting three dimensional ellipsoid is

$$\mathbf{E} = \begin{bmatrix} r_1 & r_2 & r3\cos\theta + r_4\sin\theta \\ r_2 & r_1 & r2\cos\theta + r_3\sin\theta \\ r3\cos\theta + r_4\sin\theta & r2\cos\theta + r_3\sin\theta & r_1 + r_2\sin(2\theta) \end{bmatrix}. \tag{5.40}$$

Such a slice for 4-PAM is shown in Fig.5.8(a).

This technique may also be used to view the complex three dimensional space. The angles θ_1, θ_2 and θ_3 define the angles made by the real part of the channel coefficient axes to each of the drawing axes x, y and z, so that $f_1 = xe^{j\theta_1}$, $f_2 = ye^{j\theta_2}$ and $f_3 = ze^{j\theta_3}$. If the first row of the Toeplitz Hermitian matrix $\mathbf{A}$ is $(a_1, a_2 + jb_2, a_3 + jb_3)$, then the equation $\mathbf{f}^H\mathbf{A}\mathbf{f} = 1$ can be shown to be equivalent to $\mathbf{x}^T\mathbf{E}\mathbf{x} = 1$ where $\mathbf{x} = (x, y, z)$

and the symmetric matrix

$$\mathbf{E} = \left[\begin{array}{c|c|c} a_1 & \begin{array}{c} a_2\cos(\theta_2-\theta_1) \\ -b_2\sin(\theta_2-\theta_1) \end{array} & \begin{array}{c} a_3\cos(\theta_3-\theta_1) \\ -b_3\sin(\theta_3-\theta_1) \end{array} \\ \hline \begin{array}{c} a_2\cos(\theta_2-\theta_1) \\ -b_2\sin(\theta_2-\theta_1) \end{array} & r_1 & \begin{array}{c} a_2\cos(\theta_3-\theta_2) \\ -b_2\sin(\theta_3-\theta_2) \end{array} \\ \hline \begin{array}{c} a_3\cos(\theta_3-\theta_1) \\ -b_3\sin(\theta_3-\theta_1) \end{array} & \begin{array}{c} a_2\cos(\theta_3-\theta_2) \\ -b_2\sin(\theta_3-\theta_2) \end{array} & r_1 \end{array}\right]. \tag{5.41}$$

The pairs of differences in (5.41) are not uniquely defined. This can be seen by observing that if the pairs of differences $(\alpha_1, \alpha_2, \alpha_3)^T = \mathbf{T}(\theta_1, \theta_2, \theta_3)^T$, the matrix $\mathbf{T}$ is singular. Thus in Fig. 5.9 (for QPSK) and Fig. 5.10 (8-PSK), θ_3 is chosen as zero.

5.6 Effect of Imperfect Channel Knowledge

In the previous sections it was assumed that, though the channel was noisy and dispersive, it was static, and perfectly known. In a mobile environment at least, this is never true.

Firstly, the channel is seldom static. Recall that the aim of the research presented in this thesis is to investigate the extent to which the behaviour of a rapidly varying channel can be predicted.

Secondly the channel is never perfectly known. Much research has been reported over the previous few decades on the performance of channel estimation techniques. This has included research on the optimal choice of training sequences, and the design and evaluation of many "blind" equalisation schemes; their convergence rates and excess error.

In this section the effect on system error probability of imperfect channel is investigated. The results obtained are similar to those of Gorokhov [63] but are more general, and in many cases more accurate. The derivations are explained in Appendices 5A and 5B at the end of this chapter. Another approach to the problem of imperfect channel information, based on considerations of mutual information is contained in [95]. Apart from these two papers, little has been published on this important problem.

5.6.1 Stochastic Channel Estimate and Noise

The symbols used here are similar to those of Section 5.1. $\mathbf{s}$ is a vector of transmitted symbols and $\mathbf{f}$ is a discrete time impulse response. Their respective convolution matrices $\mathbf{S}$ and $\mathbf{F}$ are as defined in (5.6) and (5.7). $\boldsymbol{\eta}$ is a vector of additive noise, $\mathbf{y}$ is a vector of received signal, with

$$\mathbf{y} = \mathbf{F}\mathbf{s} + \boldsymbol{\eta} = \mathbf{S}\mathbf{f} + \boldsymbol{\eta}. \tag{5.42}$$

The probability of a particular error event when the channel knowledge is perfect is the probability that the maximum likelihood metric of $\widehat{\mathbf{s}} \neq \mathbf{s}$ is smaller than that of $\mathbf{s}$, or

$$P(\mathcal{E}) = P(\|\mathbf{y} - \mathbf{F}\widehat{\mathbf{s}}\| \leq \|\mathbf{y} - \mathbf{F}\mathbf{s}\|). \tag{5.43}$$

If the estimate of the channel is $\widehat{\mathbf{f}} \neq \mathbf{f}$ (with convolution matrix $\widehat{\mathbf{F}}$), the probability of a particular error event is

$$\begin{aligned}
P(\mathcal{E}) &= P(\|\mathbf{y} - \widehat{\mathbf{S}}\widehat{\mathbf{f}}\| \leq \|\mathbf{y} - \mathbf{S}\widehat{\mathbf{f}}\|) \\
&= P(\|\mathbf{S}\mathbf{f} - \widehat{\mathbf{S}}\widehat{\mathbf{f}} + \boldsymbol{\eta}\| \leq \|\mathbf{S}\mathbf{f} - \mathbf{S}\widehat{\mathbf{f}} + \boldsymbol{\eta}\|) \\
&= P(\|\mathbf{S}\mathbf{f} - \mathbf{S}\widehat{\mathbf{f}} + \mathbf{S}\widehat{\mathbf{f}} - \widehat{\mathbf{S}}\widehat{\mathbf{f}} + \boldsymbol{\eta}\| \leq \|\mathbf{S}\mathbf{f} - \mathbf{S}\widehat{\mathbf{f}} + \boldsymbol{\eta}\|) \\
&= P(\|-\mathbf{S}\Delta_{\mathbf{f}} - \mathbf{E}\widehat{\mathbf{f}} + \boldsymbol{\eta}\|^2 \leq \|-\mathbf{S}\Delta_{\mathbf{f}} + \boldsymbol{\eta}\|^2) \\
&= P\left(\|\mathbf{E}\widehat{\mathbf{f}}\|^2 + 2\,\mathrm{Re}\left(\widehat{\mathbf{f}}^H\mathbf{E}^H\mathbf{S}\Delta_{\mathbf{f}} - \widehat{\mathbf{f}}^H\mathbf{E}^H\boldsymbol{\eta}\right) < 0\right) \\
&= P\left(\|\widehat{\boldsymbol{\epsilon}}\|^2 + 2\,\mathrm{Re}\left(\widehat{\boldsymbol{\epsilon}}^H\mathbf{S}\Delta_{\mathbf{f}} - \widehat{\boldsymbol{\epsilon}}^H\boldsymbol{\eta}\right) < 0\right) \\
&= P(\chi < 0),
\end{aligned} \tag{5.44}$$

where $\mathbf{E} = \widehat{\mathbf{S}} - \mathbf{S}$, $\Delta_{\mathbf{f}} = \widehat{\mathbf{f}} - \mathbf{f}$, $\widehat{\boldsymbol{\epsilon}} = \mathbf{E}\widehat{\mathbf{f}}$, and $\chi = \|\widehat{\boldsymbol{\epsilon}}\|^2 + 2\,\mathrm{Re}(\widehat{\boldsymbol{\epsilon}}^H\mathbf{S}\Delta_{\mathbf{f}} - \widehat{\boldsymbol{\epsilon}}^H\boldsymbol{\eta})$.

The exact distribution of χ involves products of two normal random variables, one of which has zero mean. The distribution of the product y of two normal random variables, $x_1 \sim \mathcal{N}(0, \sigma_1^2)$, $x_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$, calculated using the Mellin convolution is presented in [137, p136]. The true distribution can be shown to be very close to the normal distribution if the non-zero mean $\mu_2 > 4\sigma_2$. This will be true provided the channel estimation error is a few times smaller than the actual channel impulse response values. An estimate of the probability (5.44) is now derived based on the assumption that the distribution of the product is close to normal. Even when the distribution of just one product is not close to normal, the sum of many of these products leads to asymptotic normality and so an accurate result may be assured.

The following assumptions are used here and in Section 5.6.2.

1. The channel estimate $\widehat{\mathbf{f}}$ is an unbiased normal estimate of $\mathbf{f}$, so that $\Delta_{\mathbf{f}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\mathbf{f}})$.

2. The noise is also complex normal so that $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\boldsymbol{\eta}})$.

3. The channel length L is sufficient that the distribution of the quantity being compared with zero in (5.44) is approximately normal.

4. The channel estimation error is statistically independent of the noise. Whenever the received signal is used for estimating the channel this is not strictly true. However, the channel estimate is usually based on a much larger block of data

than the duration of one error event, and so such an estimate $\widehat{\mathbf{f}}$ will have little correlation with $\boldsymbol{\eta}$.

In [63] use is made of several assumptions in addition to these to show that $\mathbf{E}\widehat{\mathbf{f}} = \mathbf{E}\mathbf{f}(1+\xi_T)$ where ξ_T converges in probability to zero ($\xi_T \xrightarrow{p} 0$). The weakness in making these assumptions is that the convergence depends on the length L of the channel more than does Assumption 3 above. In effect, the dependence of $\widehat{\boldsymbol{\epsilon}}$ on the random variable $\Delta_{\mathbf{f}}$ is neglected.

The approach of this section is to take account of this dependence, and to use the χ in (5.44) to approximate the true distribution under the assumption that the resulting distribution is approximately normal. The result of the analysis is the following theorem which is proved in Appendix 5A.

Theorem 1 *Let $\Delta_{\mathbf{f}} \in \mathbb{C}^L$ with $\Delta_{\mathbf{f}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C_f})$, $\boldsymbol{\eta} \in \mathbb{C}^{R+L-1}$ with $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\boldsymbol{\eta}})$, and with $\Delta_{\mathbf{f}}$ and $\boldsymbol{\eta}$ independent. Let $\mathbf{S} \in \mathbb{C}^{(R+L-1)\times L}$ and $\mathbf{E} \in \mathbb{C}^{(R+L-1)\times L}$ be fixed matrices and let $\mathbf{f} \in \mathbb{C}^L$. Define $\widehat{\boldsymbol{\epsilon}} = \mathbf{E}(\mathbf{f}+\Delta_{\mathbf{f}})$, and $\chi = \|\widehat{\boldsymbol{\epsilon}}\|^2 + 2\,\mathrm{Re}(\widehat{\boldsymbol{\epsilon}}^H \mathbf{S}\Delta_{\mathbf{f}} - \widehat{\boldsymbol{\epsilon}}^H \boldsymbol{\eta})$. Then the mean of χ is*

$$E\{\chi\} = \|\boldsymbol{\epsilon}\|^2 + \mathrm{tr}(\mathbf{G}\mathbf{C_f}) + \mathrm{tr}(\mathbf{H}\mathbf{C_f}), \tag{5.45}$$

and the variance of χ is

$$V\{\chi\} = \mathrm{tr}((\mathbf{G}\mathbf{C_f})^2) + \mathrm{tr}((\mathbf{H}\mathbf{C_f})^2) + 2\boldsymbol{\epsilon}^H(\widehat{\mathbf{S}}\mathbf{C_f}\widehat{\mathbf{S}}^H + \mathbf{C}_{\boldsymbol{\eta}})\boldsymbol{\epsilon} + 2\mathbf{1}^T((\mathbf{E}\mathbf{C_f}\mathbf{E}^H) \odot \mathbf{C}_{\boldsymbol{\eta}})\mathbf{1}, \tag{5.46}$$

where $\widehat{\mathbf{S}} = \mathbf{S} + \mathbf{E}$, $\mathbf{H} = \mathbf{E}^H\mathbf{S} + \mathbf{S}^H\mathbf{E}$, $\boldsymbol{\epsilon} = \mathbf{E}\mathbf{f}$, $\mathbf{G} = \mathbf{E}^H\mathbf{E}$ and $\mathbf{1} = (1, 1, \ldots, 1)^T$.

If the distribution of χ was normal, the probability of the error event occurring, being the probability that χ is less than zero, would be

$$P(\mathcal{E}) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{(\|\boldsymbol{\epsilon}\|^2 + \mathrm{tr}(\mathbf{G}\mathbf{C_f}) + \mathrm{tr}(\mathbf{H}\mathbf{C_f}))}{\sqrt{\begin{array}{l}\mathrm{tr}((\mathbf{G}\mathbf{C_f})^2) + \mathrm{tr}((\mathbf{H}\mathbf{C_f})^2) \\ +2\mathbf{1}^T((\mathbf{E}\mathbf{C_f}\mathbf{E}^H) \odot \mathbf{C}_{\boldsymbol{\eta}})\mathbf{1} + 2\boldsymbol{\epsilon}^H(\widehat{\mathbf{S}}\mathbf{C_f}\widehat{\mathbf{S}}^H + \mathbf{C}_{\boldsymbol{\eta}})\boldsymbol{\epsilon}\end{array}}}\right) \tag{5.47}$$

The distribution of χ was investigated numerically, and compared with the normal distribution using the Kolmogorov Smirnov test for many different values of $\mathbf{E}$, $\mathbf{S}$, $\mathbf{f}$, $\mathbf{C_f}$ and $\mathbf{C}_{\boldsymbol{\eta}}$. The distribution was found to be a very good fit to normal, even with small values of L. Thus even though the distribution of χ is not strictly normal, (5.47) provides a reliable estimate of the probability of a particular error event occurring.

The expression is of limited usefulness, however, since it includes the transmitted symbol sequence $\mathbf{s}$ (actually a convolution matrix $\mathbf{S}$ of $\mathbf{s}$) as an explicit parameter. To

Modulation	Variance
BPSK	1/4
4-PAM	5/4
8-PAM	21/4
16-PAM	85/4
QPSK	1/2
8-PSK	2
16-PSK	2
16-QAM	5/2

Table 5.1: Variances of Modulation Schemes.

obtain the average probability of a particular error event requires averaging over all possible transmitted error sequences $\mathbf{s}$. The next section then derives a similar mean and variance where $\mathbf{s}$ is also a random quantity.

5.6.2 Stochastic transmitted symbols s

In this section the probability of an error event is derived under the conditions that the noise $\boldsymbol{\eta}$, the channel estimate error $\Delta_{\mathbf{f}}$ and the transmitted symbol sequence $\mathbf{s}$ are all stochastic unknown complex quantities, with zero mean and known variance. The variance of a white symbol sequence composed of symbols from the modulation schemes shown in Fig. 5.2 is shown in Table 5.1

The assumptions used in this section are the same as in Section 5.6.1, with the additional assumptions that

5. Successive transmitted symbols are independent (the data $\mathbf{s}$ is "white") so that $\{\mathbf{s}_i\mathbf{s}_j^*\} = \sigma_s^2\delta_{ij}$.

6. The data $\mathbf{s}$ and the channel estimate $\Delta_{\mathbf{f}}$ are uncorrelated.

The following theorem is proved in Appendix 5B.

Theorem 2 *Let $\Delta_{\mathbf{f}} \in \mathbb{C}^L$ with $\Delta_{\mathbf{f}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\mathbf{f}})$, $\boldsymbol{\eta} \in \mathbb{C}^{R+L-1}$ with $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\boldsymbol{\eta}})$, $\mathbf{s} \in \mathbb{C}^R$ with $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}, \sigma_s^2\mathbf{I})$ and with $\Delta_{\mathbf{f}}$, $\boldsymbol{\eta}$ and $\mathbf{s}$ all independent. Let $\mathbf{E} \in \mathbb{C}^{(R+L-1)\times L}$ be a fixed matrix, and let $\mathbf{f} \in \mathbb{C}^L$. Define the convolution matrix $\mathbf{S}$ of $\mathbf{s}$ using (5.7), $\widehat{\boldsymbol{\epsilon}} = \mathbf{E}(\mathbf{f} + \Delta_{\mathbf{f}})$, and $\chi = \|\widehat{\boldsymbol{\epsilon}}\|^2 + 2\,\mathrm{Re}(\widehat{\boldsymbol{\epsilon}}^H\mathbf{S}\Delta_{\mathbf{f}} - \widehat{\boldsymbol{\epsilon}}^H\boldsymbol{\eta})$. Then the mean of χ is*

$$E\{\chi\} = \|\boldsymbol{\epsilon}\|^2 + \mathrm{tr}(\mathbf{G}\mathbf{C}_{\mathbf{f}}) \tag{5.48}$$

and the variance of χ is

$$\begin{aligned} V\{\chi\} = {} & \mathrm{tr}((\mathbf{G}\mathbf{C}_{\mathbf{f}})^2) + 2\boldsymbol{\epsilon}^H(\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H + \mathbf{C}_{\boldsymbol{\eta}} + \sigma_s^2\mathbf{C}_{\mathbf{s}})\boldsymbol{\epsilon} \\ & + 2\mathbf{1}^T((\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H) \odot (\mathbf{C}_{\boldsymbol{\eta}} + \sigma_s^2\mathbf{C}_{\mathbf{s}}{}^*))\mathbf{1} + 2\sigma_s^2\mathbf{1}^T(\mathbf{E}\mathbf{C}_{\mathbf{f}} \odot (\mathbf{E}\mathbf{C})_{\mathbf{s}}^*)\mathbf{1}, \end{aligned} \tag{5.49}$$

where $\boldsymbol{\epsilon} = \mathbf{E}\mathbf{f}$, $\mathbf{G} = \mathbf{E}^H\mathbf{E}$, *the* ij*-th element of the matrix* $\mathbf{C_s}$ *is*

$$[\mathbf{C_s}]_{ij} = \sum_{\substack{p=1 \\ p-q=i-j \\ 1\le i-p+1\le R}}^{L} \sum_{q=1}^{L} [\mathbf{C_f}]_{pq}, \tag{5.50}$$

the ij*-th element of the matrix* $(\mathbf{EC})_\mathbf{s}$ *is*

$$[(\mathbf{EC})_\mathbf{s}]_{ij} = \sum_{\substack{p=1 \\ p-q=i-j \\ 1\le i-p+1\le R}}^{R+L-1} \sum_{q=1}^{L} [\mathbf{EC_f}]_{pq}, \tag{5.51}$$

and $\mathbf{1} = (1, 1, \ldots, 1)^T$, *of length* L *or* $R + L - 1$.

If the distribution of χ was normal, the probability of the error event corresponding to $\mathbf{E}$ occurring would be

$$P(\mathcal{E}) = \frac{1}{2}\operatorname{erfc}\left(\frac{\|\boldsymbol{\epsilon}\|^2 + \operatorname{tr}(\mathbf{GC_f})}{\sqrt{\begin{array}{c}\operatorname{tr}((\mathbf{GC_f})^2) + 2\boldsymbol{\epsilon}^H(\mathbf{EC_f}\mathbf{E}^H + \mathbf{C}_{\boldsymbol{\eta}} + \sigma_s^2\mathbf{C_s})\boldsymbol{\epsilon} \\ +2\mathbf{1}^T((\mathbf{EC_f}\mathbf{E}^H) \odot (\mathbf{C}_{\boldsymbol{\eta}} + \sigma_s^2\mathbf{C_s}^*))\mathbf{1} + 2\sigma_s^2\mathbf{1}^T(\mathbf{EC_f} \odot (\mathbf{EC})_\mathbf{s}^*)\mathbf{1}\end{array}}}\right). \tag{5.52}$$

The true distribution was investigated numerically, as in Section 5.6.1, but this time the distribution was found not to converge so rapidly to normality. Though in some respects resembling a normal distribution, the left tail tends to be steeper than the right tail. Despite this, the expression (5.52) still provides a useful approximation to the actual error probability.

This is demonstrated in a simple example. The modulation scheme is 8-PSK, and the error event $\mathbf{e}$ being investigated is the single ($R = 1$) error a (refer to Fig. A.1 in Appendix A). The actual channel is given by $f_l = \exp(-3l/L)$, normalised to have unit power, and $L = 4$. The channel error covariance matrix is $\mathbf{C_f} = 0.01\mathbf{I}$, and the noise covariance $\mathbf{C}_{\boldsymbol{\eta}} = 0.1\mathbf{I}$. The density function of χ estimated from numerical experiment of 4×10^6 transmissions, and the normal distribution with the same mean and variance (calculated using (5.48) and (5.49)) are compared in Fig. 5.11. The probability of the error event occurring is $p(\mathcal{E}) = p(\chi < 0)$. The two cumulative distribution functions are compared in Fig. 5.12. The probability of error for both the simulated and approximated cases is 1.48×10^{-4}. The approximation here appears to be quite accurate. One would expect that for larger values of L the approximation would be more accurate, since the distribution of χ converges to the normal. The approximated probability of longer error events may not be so accurate however.

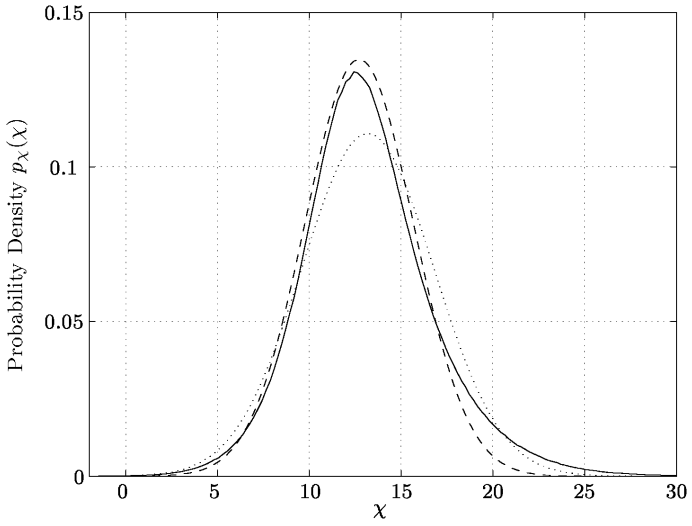


Figure 5.11: Comparison of the density functions estimated from numerical experiment (solid line), and a normal distribution (dotted line) with parameters calculated using (5.48) and (5.49). The dashed line is the distribution used in (5.53).

Note that in [63] it is assumed that $\sqrt{K}(\widehat{\mathbf{f}} - \mathbf{f}) \xrightarrow{d} \mathcal{CN}(0, \mathbf{C_f})$, where K is the number of data samples from which the estimate $\widehat{\mathbf{f}}$ of the channel $\mathbf{f}$ is obtained.[1] If $\mathbf{C_f}$ is assumed to tend to zero in this manner, except in the calculation of $\mathbf{C_s}$, and $\mathbf{C_\eta} = \sigma_\eta \mathbf{I}$ the probability of error becomes

$$\begin{aligned} P(\mathcal{E}) &= \frac{1}{2}\operatorname{erfc}\left(\frac{\|\boldsymbol{\epsilon}\|^2}{\sqrt{2\left(\sigma_s^2 2\boldsymbol{\epsilon}^H \mathbf{C_s}\boldsymbol{\epsilon} + 2\boldsymbol{\epsilon}^H \mathbf{C_\eta}\boldsymbol{\epsilon}\right)}}\right) \\ &= \frac{1}{2}\operatorname{erfc}\left(\frac{\|\boldsymbol{\epsilon}\|}{2\sigma_\eta}\left(1 + \frac{\boldsymbol{\epsilon}^H \mathbf{C_s}\boldsymbol{\epsilon}}{\|\boldsymbol{\epsilon}\|^2}\frac{\sigma_s^2}{\sigma_\eta^2}\right)^{-\frac{1}{2}}\right) \end{aligned} \tag{5.53}$$

which is very similar to the expression obtained in [63]. It can be seen from Fig. 5.11 and Fig. 5.12 that while the distribution corresponding to this simplification is close to the true distribution, the *cumulative* distribution used to calculate $p(\chi < 0)$ is quite inaccurate.

The problem of finding a set of error events $\mathcal{U}$ smaller than $\mathcal{D}^R$ for which to evaluate the sum (5.14) when the channel knowledge is imperfect remains to be investigated. It is quite possible that a similar approach to that used in Section 5.2 would prove effective.

From this section it is concluded that even in a situation where the channel impulse response is not perfectly known, the symbol error rate can be determined from the inaccurate channel estimate and some estimate of the channel estimation error.

[1] A useful method of choosing a training sequence to minimise the error in a channel estimate is described in [29].

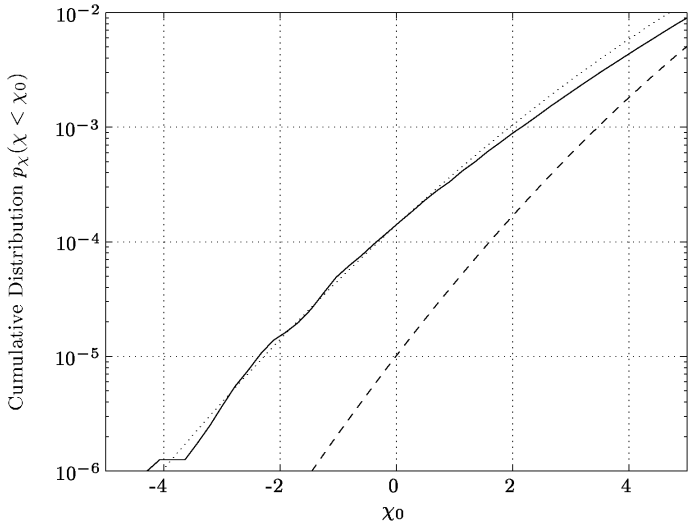


Figure 5.12: Comparison of the cumulative distribution functions estimated from numerical experiment (solid line), and a normal distribution (dotted line) with parameters calculated using (5.48) and (5.49). The dashed line is the distribution used in (5.53). Although the approximate expression of (5.53) results in a good approximation to the true density (Fig 5.11), the *cumulative* distribution near zero is much better represented by (5.48) and (5.49).

5.7 Summary and Contributions

In this chapter the application of predicted channel information to predicting the error rate of a broad-band system has been considered.

We itemise some specific contributions made in this chapter:

i. A new insight into the error behaviour of a broad-band system has been presented — that the error performance can be accurately estimated from knowledge of the channel and of the relatively small set of error events which contribute most of the errors.

ii. A method has been proposed which allows the set of contributing error events to be found in a way which is *channel independent*, so the information, once calculated may be incorporated into a real system, so that it can predict its own performance. The set of such error events for some uncoded modulation schemes has been found.

iii. The approach has been extended to the case where the channel information is inaccurate, and an asymptotic expression derived which allows the probability of occurrence of an error event in such a situation to be estimated.

5A Proof of Theorem 1

Proof

The aim of this appendix is to find the mean and variance of χ, where $\chi = \|\widehat{\boldsymbol{\epsilon}}\|^2 + 2\,\mathrm{Re}(\widehat{\boldsymbol{\epsilon}}^H \mathbf{S}\Delta_{\mathbf{f}} - \widehat{\boldsymbol{\epsilon}}^H \boldsymbol{\eta})$.

Expanding the expression for χ

$$\begin{aligned}\chi = \mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{f} + {\Delta_{\mathbf{f}}}^H\mathbf{E}^H\mathbf{E}\Delta_{\mathbf{f}} + 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_{\mathbf{f}}) + 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_{\mathbf{f}}) \\ + 2\,\mathrm{Re}({\Delta_{\mathbf{f}}}^H\mathbf{E}^H\mathbf{S}\Delta_{\mathbf{f}}) - 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\boldsymbol{\eta}) - 2\,\mathrm{Re}({\Delta_{\mathbf{f}}}^H\mathbf{E}^H\boldsymbol{\eta}). \qquad (5.54)\end{aligned}$$

For convenience of reference, the following terms are defined:

$$\begin{aligned} t_1 &= \mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{f} \\ t_2 &= {\Delta_{\mathbf{f}}}^H\mathbf{E}^H\mathbf{E}\Delta_{\mathbf{f}} \\ t_3 &= 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_{\mathbf{f}}) \\ t_4 &= 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_{\mathbf{f}}) \\ t_5 &= 2\,\mathrm{Re}({\Delta_{\mathbf{f}}}^H\mathbf{E}^H\mathbf{S}\Delta_{\mathbf{f}}) \\ t_6 &= -2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\boldsymbol{\eta}) \\ t_7 &= -2\,\mathrm{Re}({\Delta_{\mathbf{f}}}^H\mathbf{E}^H\boldsymbol{\eta}) \\ \chi &= t_1 + t_2 + t_3 + t_4 + t_5 + t_6 + t_7. \qquad (5.55)\end{aligned}$$

The terms $t_1, \ldots, t_7$ are examined to find their mean and variance, and the covariances of pairs of terms.

5A.1 The term $t_1 = \mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{f}$

This term is constant, and so has mean t_1, and zero variance.

5A.2 The term $t_2 = {\Delta_{\mathbf{f}}}^H\mathbf{E}^H\mathbf{E}\Delta_{\mathbf{f}}$

Since $E\{\Delta_{\mathbf{f}}{\Delta_{\mathbf{f}}}^H\} = \mathbf{C}_{\mathbf{f}}$, it follows that

$$\begin{aligned} E\{{\Delta_{\mathbf{f}}}^H\mathbf{E}^H\mathbf{E}\Delta_{\mathbf{f}}\} &= E\{\mathrm{tr}(\mathbf{E}\Delta_{\mathbf{f}}{\Delta_{\mathbf{f}}}^H\mathbf{E}^H)\} \\ &= \mathrm{tr}(\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H) \\ &= \mathrm{tr}(\mathbf{E}^H\mathbf{E}\mathbf{C}_{\mathbf{f}}). \qquad (5.56)\end{aligned}$$

The variance of this term can be found by considering the eigendecomposition of $\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H$ which can be expressed as $\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H\mathbf{V} = \mathbf{V}\mathbf{D}$. Then if $\mathbf{g} = \mathbf{V}^H\mathbf{E}\Delta_{\mathbf{f}}$, and $E\{\mathbf{g}\mathbf{g}^H\} = \mathbf{D}$, the elements of $\mathbf{g}$ are uncorrelated, with variances given by $E\{g_i g_i^*\} =$

σ_i^2. Then

$$\begin{aligned} E\{{\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f}{\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f}\} &= E\{\mathbf{g}^H\mathbf{V}^H\mathbf{V}\mathbf{g}\mathbf{g}^H\mathbf{V}^H\mathbf{V}\mathbf{g}\} \\ &= E\{\mathbf{g}^H\mathbf{g}\mathbf{g}^H\mathbf{g}\} \\ &= E\left\{\left(\sum_{i=1}^{L}|g_i|^2\right)^2\right\} \\ &= \sum_{i=1}^{L}E\left\{|g_i|^4\right\} + 2\sum_{i=1}^{L}\sum_{j\neq i}E\{|g_i|^2|g_j|^2\}. \end{aligned} \tag{5.57}$$

At this point it is necessary to know the expectation of x^4 for $x \sim \mathcal{N}(0,\sigma_x^2), n \in \mathbb{N}$. Using the integral (3.381-4) of [64]

$$E(x^n) = \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi\sigma_x^2}}\, e^{-\frac{x^2}{2\sigma_x^2}} \quad dx = \begin{cases} \sigma_x^n(n-1)!! & n \text{ even} \\ 0 & n \text{ odd}, \end{cases} \tag{5.58}$$

where $n!! = 1.3.5.\cdots.n$. For complex $z = x + jy \sim \mathcal{CN}(0,\sigma_z^2)$,

$$\begin{aligned} E(|z|^4) &= E\{(x^4 + y^4 + 2x^2y^2\} \\ &= E\left\{3(\frac{1}{2}\sigma_z^2)^2 + 3(\frac{1}{2}\sigma_z^2)^2 + 2(\frac{1}{2}\sigma_z^2)^2\right\} \\ &= 2\sigma_z^4. \end{aligned} \tag{5.59}$$

Thus

$$\begin{aligned} \sum_{i=1}^{L}E\left\{|g_i|^4\right\} + 2\sum_{i=1}^{L}\sum_{j\neq i}E\{|g_i|^2|g_j|^2\} &= \sum_{i=1}^{L}2\sigma_i^4 + 2\sum_{i=1}^{L}\sum_{j\neq i}\sigma_i^2\sigma_j^2 \\ &= \mathrm{tr}(\mathbf{D}^2) + (\mathrm{tr}(\mathbf{D}))^2 \\ &= \mathrm{tr}((\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f})^2) + (\mathrm{tr}(\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f}))^2. \end{aligned} \tag{5.60}$$

Therefore the variance of t_2 is given by

$$V(t_2) = \mathrm{tr}((\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f})^2). \tag{5.61}$$

5A.3 The term $t_3 = 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f})$

The only random part of this term is $\Delta_\mathbf{f}$, which has zero mean, and hence

$$E\{2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f})\} = 0. \tag{5.62}$$

If $\boldsymbol{\alpha} = \mathbf{f}^H\mathbf{E}^H\mathbf{E}$, then

$$\begin{aligned} & V\{2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f})\} \\ &= V\{2\,\mathrm{Re}(\boldsymbol{\alpha}\Delta_\mathbf{f})\} \\ &= E\{(\boldsymbol{\alpha}^H\Delta_\mathbf{f} + \Delta_\mathbf{f}\boldsymbol{\alpha})^2\} \\ &= E\{\boldsymbol{\alpha}^H\Delta_\mathbf{f}\Delta_\mathbf{f}{}^H\boldsymbol{\alpha} + \mathrm{tr}(\boldsymbol{\alpha}^H\Delta_\mathbf{f}\Delta_\mathbf{f}{}^H\boldsymbol{\alpha}) + \boldsymbol{\alpha}^H\Delta_\mathbf{f}\Delta_\mathbf{f}{}^T\boldsymbol{\alpha}^* + \boldsymbol{\alpha}^T\Delta_\mathbf{f}{}^*\Delta_\mathbf{f}{}^H\boldsymbol{\alpha}\} \\ &= 2\boldsymbol{\alpha}^H\mathbf{C}_\mathbf{f}\boldsymbol{\alpha} \\ &= 2\mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H\mathbf{E}\mathbf{f}. \end{aligned} \tag{5.63}$$

5A.4 The term $t_4 = 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})$

Identical reasoning to that for t_3 shows that

$$E\{2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\} = 0, \qquad \text{and} \tag{5.64}$$

$$V\{2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\} = 2\mathbf{f}^H\mathbf{E}^H\mathbf{S}\mathbf{C}_\mathbf{f}\mathbf{S}^H\mathbf{E}\mathbf{f}. \tag{5.65}$$

5A.5 The term $t_5 = 2\,\mathrm{Re}(\Delta_\mathbf{f}{}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})$

If $\mathbf{H} = \mathbf{E}^H\mathbf{S} + \mathbf{S}^H\mathbf{E}$, then

$$\begin{aligned} E\{2\,\mathrm{Re}(\Delta_\mathbf{f}{}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\} &= E\{\Delta_\mathbf{f}{}^H\mathbf{H}\Delta_\mathbf{f}\} \\ &= E\{\mathrm{tr}(\mathbf{H}\Delta_\mathbf{f}\Delta_\mathbf{f}{}^H)\} \\ &= \mathrm{tr}(\mathbf{H}\mathbf{C}_\mathbf{f}). \end{aligned} \tag{5.66}$$

If $\mathbf{H}$, being Hermitian, is written in the form $\mathbf{H} = \mathbf{X}^H\mathbf{X}$, then using the same reasoning as for Section 5A.2 the variance is

$$\begin{aligned} V\{2\,\mathrm{Re}(\Delta_\mathbf{f}{}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\} &= \mathrm{tr}((\mathbf{X}^H\mathbf{X}\mathbf{C}_\mathbf{f})^2) \\ &= \mathrm{tr}((\mathbf{H}\mathbf{C}_\mathbf{f})^2) \end{aligned} \tag{5.67}$$

In fact $\mathbf{H}$ can only be written as $\mathbf{H} = \mathbf{X}^H\mathbf{X}$ if the eigenvalues are positive. Equation 5.67 is valid even if some eigenvalues are non-positive, and a proof is simple to construct using an almost identical approach to that of Section 5A.9.

5A.6 The term $t_6 = -2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\boldsymbol{\eta})$

The only stochastic part of this term is the zero mean noise $\boldsymbol{\eta}$. Following Section 5A.3,

$$E\{-2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\boldsymbol{\eta})\} = 0 \tag{5.68}$$

and

$$V\{-2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\boldsymbol{\eta})\} = 2\mathbf{f}^H\mathbf{E}^H\mathbf{C}_{\boldsymbol{\eta}}\mathbf{E}\mathbf{f} \tag{5.69}$$

5A.7 The term $t_7 = -2\,\mathrm{Re}(\Delta_{\mathbf{f}}{}^H\mathbf{E}^H\boldsymbol{\eta})$

Using the assumption that the channel estimation error is statistically independent of the noise

$$E\{-2\,\mathrm{Re}(\Delta_{\mathbf{f}}{}^H\mathbf{E}^H\boldsymbol{\eta})\} = 0, \tag{5.70}$$

and if $\mathbf{h} = \mathbf{E}\Delta_{\mathbf{f}}$

$$\begin{aligned} V\{(\Delta_{\mathbf{f}}{}^H\mathbf{E}^H\boldsymbol{\eta})\} &= E\{\mathbf{h}^H\boldsymbol{\eta}\boldsymbol{\eta}^H\mathbf{h}\} \\ &= E\left\{\left(\sum_{i=1}^{L} h_i^*\eta_i\right)\left(\sum_{j=1}^{L}\eta_j^* h_j\right)\right\} \\ &= \sum_{i=1}^{L}\sum_{j=1}^{L} E\{h_i^* h_j \eta_j^*\eta_i\} \\ &= \sum_{i=1}^{L}\sum_{j=1}^{L} E\{h_i^* h_j\}E\{\eta_j^*\eta_i\} \\ &= \mathbf{1}^T((\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H)\odot\mathbf{C}_{\boldsymbol{\eta}})\mathbf{1} \end{aligned} \tag{5.71}$$

where $\mathbf{1} = (1, 1, \ldots, 1)^T$. Use has again been made of the assumption that h_i and η_i are independent, so that the variance of their product is equal to the product of their variances [81, p245]. If $\mathbf{C}_{\boldsymbol{\eta}}$ is diagonal, $E\{\eta_i^*\eta_j\}$ will be zero for $i \neq j$, and so

$$\begin{aligned} \sum_{i=1}^{L}\sum_{j=1}^{L} E\{h_i^* h_j\}E\{\eta_j^*\eta_i\} &= \sum_{i=1}^{L} E\{h_i^* h_i\}E\{\eta_i^*\eta_i\} \\ &= \mathrm{tr}(\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H\mathbf{C}_{\boldsymbol{\eta}}). \end{aligned} \tag{5.72}$$

Taking the real part halves this variance, and doubling the value quadruples the variance, so

$$V\{-2\,\mathrm{Re}(\Delta_{\mathbf{f}}{}^H\mathbf{E}^H\boldsymbol{\eta})\} = 2\mathbf{1}^T((\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H)\odot\mathbf{C}_{\boldsymbol{\eta}})\mathbf{1} \tag{5.73}$$

(or $2\,\mathrm{tr}(\mathbf{E}\mathbf{C}_{\mathbf{f}}\mathbf{E}^H\mathbf{C}_{\boldsymbol{\eta}})$ when $\mathbf{C}_{\boldsymbol{\eta}}$ is diagonal).

5A.8 The cross-correlation of t_3 and t_4

The expression to be evaluated here is

$$\begin{aligned}
&E\{2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f})2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})^*\} \\
&= E\{(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f} + {\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{E}\mathbf{f})(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f} + {\Delta_\mathbf{f}}^H\mathbf{S}^H\mathbf{E}\mathbf{f})\} \\
&= E\{\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f}{\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{E}\mathbf{f} + \mathbf{f}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f} + {\Delta_\mathbf{f}}^H\mathbf{S}^H\mathbf{E}\mathbf{f}\} \\
&= 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{S}^H\mathbf{E}\mathbf{f}).
\end{aligned} \tag{5.74}$$

The third line follows from the fact that $E\{{\Delta_\mathbf{f}}^T\Delta_\mathbf{f}\} = E\{{\Delta_\mathbf{f}}^H{\Delta_\mathbf{f}}^*\} = 0$.

5A.9 The cross correlation of t_2 and t_5

Recall from Section 5A.5 $\mathbf{H} = \mathbf{E}^H\mathbf{S} + \mathbf{S}^H\mathbf{E}$. The eigendecomposition (different from that of Section 5A.2) of $\mathbf{C}_\mathbf{f}$ is written as $\mathbf{C}_\mathbf{f}\mathbf{V} = \mathbf{V}\mathbf{D}$, and $\mathbf{g} = \mathbf{V}^H\Delta_\mathbf{f}$, so that $E\{\mathbf{g}\mathbf{g}^H\} = \mathbf{D}$, so the elements of $\mathbf{g}$ are uncorrelated, with variances given by $E\{g_i g_i^*\} = \sigma_i^2$.

If $\mathbf{G}' = \mathbf{V}^H\mathbf{E}^H\mathbf{E}\mathbf{V}$, and $\mathbf{H}' = \mathbf{V}^H(\mathbf{E}^H\mathbf{S} + \mathbf{S}^H\mathbf{E})\mathbf{V}$, then

$$\begin{aligned}
&E\{{\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f}2\,\mathrm{Re}({\Delta_\mathbf{f}}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\} \\
&= E\{(\mathbf{g}^H\mathbf{V}^H\mathbf{E}^H\mathbf{E}\mathbf{V}\mathbf{g})(\mathbf{g}^H\mathbf{V}^H(\mathbf{E}^H\mathbf{S} + \mathbf{S}^H\mathbf{E})\mathbf{V}\mathbf{g})\} \\
&= E\{(\mathbf{g}^H\mathbf{G}'\mathbf{g})(\mathbf{g}^H\mathbf{H}'\mathbf{g})\} \\
&= E\left\{\left(\sum_{i=1}^{L}\sum_{j=1}^{L} g_i^* G'_{ij} g_j\right)\left(\sum_{k=1}^{L}\sum_{l=1}^{L} g_k^* H'_{kl} g_l\right)\right\} \\
&= \sum_{i=1}^{L}\sum_{j=1}^{L}\sum_{k=1}^{L}\sum_{l=1}^{L} E\left\{g_i^* g_k^* g_j g_l G'_{ij} H'_{kl}\right\} \\
&= \sum_{i=1}^{L}\sum_{k=1}^{L} \sigma_i^2\sigma_k^2 G'_{jj} H'_{kk} \\
&= \left(\sum_{i=1}^{L}\sigma_i^2 G'_{jj}\right)\left(\sum_{k=1}^{L}\sigma_k^2 H'_{kk}\right) \\
&= \mathrm{tr}(\mathbf{G}'\mathbf{D})\,\mathrm{tr}(\mathbf{H}'\mathbf{D}) \\
&= \mathrm{tr}(\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f})\,\mathrm{tr}(\mathbf{H}\mathbf{C}_\mathbf{f}).
\end{aligned} \tag{5.75}$$

Note that only those terms of the quadruple summation are non-zero for which both $i = j$ *and* $k = l$, and if the elements of $\mathbf{g}$ are independent (they are jointly normal and uncorrelated) and have zero mean, the variance of the product is the product of the variances.

5A.10 The covariance of t_2 and t_3

Using the eigendecomposition of Section 5A.2, allowing $\mathbf{c} = \mathbf{f}^H\mathbf{E}^H\mathbf{E}$, $g_i = x_i + jy_i$, and $c_i = a_i + jb_i$,

$$\begin{aligned} E\{\Delta_\mathbf{f}\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f}\mathbf{f}\mathbf{E}^H\mathbf{E}\Delta_\mathbf{f}\} &= E\{\mathbf{g}^H\mathbf{g}\mathbf{b}^H\mathbf{g}\} \\ &= E\left\{\sum_{k=1}^{L}\sum_{l=1}^{L}(x_k^2+y_k^2)(x_l+jy_l)(a_l-jb_l)\right\}. \end{aligned} \tag{5.76}$$

It can be seen that each of the terms involved in this expression involves an odd power of a zero mean normal random variable, and from (5.58) the expectation of each of these is zero. Hence the covariance of terms t_2 and t_3 is zero.

If is fairly straightforward to show that all of the remaining cross correlations are also zero.

5A.11 The Combined Expression

Given that the only terms with non-zero means are t_1, t_2 and t_5, and the only pairs of terms with non-zero cross correlation are (t_1, t_2), (t_1, t_5), (t_2, t_5) and (t_3, t_4), the mean and variance of the combined expression χ can be readily calculated. Recall that $\mathbf{H} = \mathbf{E}^H\mathbf{S} + \mathbf{S}^H\mathbf{E}$. Setting $\boldsymbol{\epsilon} = \mathbf{E}\mathbf{f}$ and $\mathbf{G} = \mathbf{E}^H\mathbf{E}$,

$$\begin{aligned} E\{\chi\} &= E\{t_1 + t_2 + t_5\} \\ &= \mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{f} + \mathrm{tr}(\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f}) + \mathrm{tr}(\mathbf{H}\mathbf{C}_\mathbf{f}) \\ &= \|\boldsymbol{\epsilon}\|^2 + \mathrm{tr}(\mathbf{G}\mathbf{C}_\mathbf{f}) + \mathrm{tr}(\mathbf{H}\mathbf{C}_\mathbf{f}), \end{aligned} \tag{5.77}$$

and

$$\begin{aligned} V\{\chi\} &= E\{\chi^2\} - (E\{\chi\})^2 \\ &= \mathrm{tr}((\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f})^2) + 2\mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H\mathbf{E}\mathbf{f} + 2\mathbf{f}^H\mathbf{E}^H\mathbf{S}\mathbf{C}_\mathbf{f}\mathbf{S}^H\mathbf{E}\mathbf{f} + \mathrm{tr}((\mathbf{H}\mathbf{C}_\mathbf{f})^2) \\ &\quad + 2\mathbf{f}^H\mathbf{E}^H\mathbf{C}_{\boldsymbol{\eta}}\mathbf{E}\mathbf{f} + 2\mathbf{1}^T((\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H)\odot\mathbf{C}_{\boldsymbol{\eta}})\mathbf{1} + 2\,\mathrm{tr}(\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f})\,\mathrm{tr}(\mathbf{H}\mathbf{C}_\mathbf{f}) \\ &\quad + 4\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{S}^H\mathbf{E}\mathbf{f}) - 2\,\mathrm{tr}(\mathbf{E}^H\mathbf{E}\mathbf{C}_\mathbf{f})\,\mathrm{tr}(\mathbf{H}\mathbf{C}_\mathbf{f}) \\ &= \mathrm{tr}((\mathbf{G}\mathbf{C}_\mathbf{f})^2) + \mathrm{tr}((\mathbf{H}\mathbf{C}_\mathbf{f})^2) \\ &\quad + 2\boldsymbol{\epsilon}^H(\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H + \mathbf{S}\mathbf{C}_\mathbf{f}\mathbf{S}^H + \mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{S}^H + \mathbf{S}\mathbf{C}_\mathbf{f}\mathbf{E}^H + \mathbf{C}_{\boldsymbol{\eta}})\boldsymbol{\epsilon} \\ &\quad + 2\mathbf{1}^T((\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H)\odot\mathbf{C}_{\boldsymbol{\eta}})\mathbf{1} \\ &= \mathrm{tr}((\mathbf{G}\mathbf{C}_\mathbf{f})^2) + \mathrm{tr}((\mathbf{H}\mathbf{C}_\mathbf{f})^2) + 2\boldsymbol{\epsilon}^H(\widehat{\mathbf{S}}\mathbf{C}_\mathbf{f}\widehat{\mathbf{S}}^H + \mathbf{C}_{\boldsymbol{\eta}})\boldsymbol{\epsilon} \\ &\quad + 2\mathbf{1}^T((\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H)\odot\mathbf{C}_{\boldsymbol{\eta}})\mathbf{1}. \end{aligned} \tag{5.78}$$

∎

5B Proof of Theorem 2

Proof

The terms are the same as in Appendix 5A, but the variance of the terms involving $\mathbf{s}$ are different.

5B.1 The term $t_4 = 2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})$

As before, this term has zero mean. To calculate the variance, the combination of $\mathbf{S}$ and $\Delta_\mathbf{f}$ is first investigated:

$$\begin{aligned}[\mathbf{S}\Delta_\mathbf{f}{\Delta_\mathbf{f}}^H\mathbf{S}^H]_{ij} &= \sum_{p=1}^{L}\sum_{q=1}^{L}\mathbf{S}_{ip}\Delta_{\mathbf{f}p}({\Delta_\mathbf{f}}^H)_q{\mathbf{S}^H}_{qj} \\ &= \sum_{p=1}^{L}\sum_{q=1}^{L}\mathbf{s}_{i-p+1}\mathbf{s}^*_{j-q+1}\Delta_{\mathbf{f}p}\Delta^*_{\mathbf{f}q} \\ &= \sigma_s^2\sum_{\substack{p-q\\=i-j}}[\mathbf{C}_\mathbf{f}]_{pq}, \end{aligned} \tag{5.79}$$

where the last line follows from

- The definition of $\mathbf{S}$ in (5.7)
- The assumption of whiteness of $\mathbf{s}$, so that $\{\mathbf{s}_i\mathbf{s}_j^*\} = \sigma_s^2\delta_{ij}$
- The assumption that $\Delta_\mathbf{f}$ and $\mathbf{s}$ are uncorrelated.

If the data block is of finite size, the summation over p and q is further limited by the fact that $s_i = 0$ for $i < 1$ and for $i > R$. Thus the summation is more strictly over $p - q = i - j$ where $i + 1 - R \le p \le i$. The following definition of the ij-th element of the matrix $\mathbf{C}_\mathbf{s}$ is then convenient:

$$[\mathbf{C}_\mathbf{s}]_{ij} = \sum_{\substack{p=1\\p-q=i-j\\1\le i-p+1\le R}}^{L}\sum_{q=1}^{L}[\mathbf{C}_\mathbf{f}]_{pq}. \tag{5.80}$$

It follows that

$$E\{(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\Delta\mathbf{f}^H\mathbf{S}^H\mathbf{E}\mathbf{f}\} = \sigma_s^2\boldsymbol{\epsilon}^H\mathbf{C}_\mathbf{s}\boldsymbol{\epsilon} \tag{5.81}$$

The variance of t_4 is then given by

$$\begin{aligned}V\{2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\} &= E\{2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})2\,\mathrm{Re}(\mathbf{f}^H\mathbf{E}^H\mathbf{S}\Delta_\mathbf{f})\} \\ &= 2\sigma_s^2\boldsymbol{\epsilon}^H\mathbf{C}_\mathbf{s}\boldsymbol{\epsilon}\end{aligned} \tag{5.82}$$

5B.2 The term $t_5 = 2\,\mathrm{Re}({\Delta_\mathbf{f}}^H \mathbf{E}^H \mathbf{S} \Delta_\mathbf{f})$

Once again, the mean of this term is zero. The variance is derived using

$$
\begin{aligned}
& V\{{\Delta_\mathbf{f}}^H \mathbf{E}^H \mathbf{S} \Delta_\mathbf{f}\} \\
&= E\{{\Delta_\mathbf{f}}^H \mathbf{E}^H \mathbf{S} \Delta_\mathbf{f} {\Delta_\mathbf{f}}^H \mathbf{S}^H \mathbf{E} \Delta_\mathbf{f}\} \\
&= E\left\{ \sum_{i=1}^{L}\sum_{j=1}^{L}\sum_{k=1}^{R+L-1}\sum_{l=1}^{R+L-1}\sum_{m=1}^{L}\sum_{n=1}^{L} \Delta_{f_i}\Delta_{f_j}\Delta^*_{f_n}\Delta^*_{f_m}\mathbf{S}_{km}\mathbf{S}^H_{nl}\mathbf{E}^H_{ik}\mathbf{E}_{lj} \right\} \\
&= E\left\{ \sum_{i=1}^{L}\sum_{j=1}^{L}\sum_{k=1}^{R+L-1}\sum_{l=1}^{R+L-1}\sum_{m=1}^{L}\sum_{n=1}^{L} \Delta_{f_i}\Delta_{f_j}\Delta^*_{f_n}\Delta^*_{f_m}\mathbf{s}_{k-m+1}\mathbf{s}^*_{l-n+1}\mathbf{E}^H_{ik}\mathbf{E}_{lj} \right\} \\
&= \sum_{i=1}^{L}\sum_{j=1}^{L}\sum_{k=1}^{R+L-1}\sum_{l=1}^{R+L-1}\sum_{m=1}^{L}\sum_{n=1}^{L} (\mathbf{C}_{\mathbf{f}mn}\mathbf{C}_{\mathbf{f}ji} + \mathbf{C}_{\mathbf{f}jn}\mathbf{C}_{\mathbf{f}mi})\sigma_s^2\delta_{m-n,k-l}\mathbf{E}^H_{ik}\mathbf{E}_{lj} \\
&= \sigma_s^2 \sum_{k=1}^{R+L-1}\sum_{l=1}^{R+L-1}\sum_{m=1}^{L}\sum_{n=1}^{L} \delta_{m-n,k-l}\left(\mathbf{C}_{\mathbf{f}mn}(\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H)_{lk} + (\mathbf{E}\mathbf{C}_\mathbf{f})_{ln}(\mathbf{C}_\mathbf{f}\mathbf{E}^H)_{mk}\right) \\
&= \sigma_s^2 \left(\sum_{k=1}^{R+L-1}\sum_{l=1}^{R+L-1} \mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H \sum_{\substack{m-n \\ =k-l}} \right) \\
&\quad + \sigma_s^2 \left(\sum_{l=1}^{R+L-1}\sum_{n=1}^{L} \mathbf{E}\mathbf{C}_\mathbf{f} \sum_{k=1}^{R+L-1}\sum_{m=1}^{L} [\mathbf{E}\mathbf{C}_\mathbf{f}{}^*]_{km}\delta_{k-m,l-n} \right) \\
&= \sigma_s^2(\mathbf{1}^H(\mathbf{E}\mathbf{C}\mathbf{E}^H \odot \mathbf{C_s}{}^*)\mathbf{1} + \mathbf{1}^H(\mathbf{E}\mathbf{C}_\mathbf{f} \odot (\mathbf{E}\mathbf{C})^*_\mathbf{s})\mathbf{1})
\end{aligned}
\tag{5.83}
$$

where the elements of the matrix $(\mathbf{E}\mathbf{C})_\mathbf{s}$ are defined by

$$
[(\mathbf{E}\mathbf{C})_\mathbf{s}]_{ij} = \sum_{\substack{p=1 \\ p-q=i-j \\ 1\le i-p+1\le R}}^{R+L-1}\sum_{q=1}^{L} [\mathbf{E}\mathbf{C}_\mathbf{f}]_{pq} \tag{5.84}
$$

and the vector $\mathbf{1} = (1, 1, \ldots, 1)^T$ is of length appropriate to the context, which may be L or $R + L - 1$. The variance of t_5 is thus given by

$$
V\{2\,\mathrm{Re}({\Delta_\mathbf{f}}^H \mathbf{E}^H \mathbf{S} \Delta_\mathbf{f})\} = 2\sigma_s^2(\mathbf{1}^T(\mathbf{E}\mathbf{C}_\mathbf{f}\mathbf{E}^H \odot \mathbf{C_s}{}^*)\mathbf{1} + \mathbf{1}^T(\mathbf{E}\mathbf{C}_\mathbf{f} \odot (\mathbf{E}\mathbf{C})^*_\mathbf{s})\mathbf{1}) \tag{5.85}
$$

5B.3 The combined expression for random s

It is fairly straightforward to show that all of the terms of the expression $\chi = t_1 + t_2 + t_3 + t_4 + t_5 + t_6 + t_7$ have zero cross correlation. The mean χ is given by

$$
E\{\chi\} = \|\boldsymbol{\epsilon}\|^2 + \mathrm{tr}(\mathbf{G}\mathbf{C}_\mathbf{f}) \tag{5.86}
$$

and the variance by

$$V\{\chi\} = \mathrm{tr}((\mathbf{GC_f})^2) + 2\boldsymbol{\epsilon}^H(\mathbf{EC_f E}^H + \mathbf{C}_{\boldsymbol{\eta}} + \sigma_s^2\mathbf{C_s})\boldsymbol{\epsilon} + 2\mathbf{1}^T((\mathbf{EC_f E}^H) \odot (\mathbf{C}_{\boldsymbol{\eta}} + \sigma_s^2\mathbf{C_s}^*))\mathbf{1} + 2\sigma_s^2\mathbf{1}^T(\mathbf{EC_f} \odot (\mathbf{EC})^*_{\mathbf{s}})\mathbf{1} \quad (5.87)$$

∎

5C Note on 'A New Method for Determining "Unknown" Worst-Case Channels for Maximum Likelihood Sequence Estimation'

In [132] a method is presented for finding the minimum distance error events and channels. Although the algorithm appears to be have been successful in the example of BPSK for channels of length up to $L = 10$, there is an error in the proof of the theorem which forms the basis of the method. Theorem 2 of the paper states that if $m \geq 2$, and p_1 and p_2 are integers such that $p_1 > p_2 \geq m - 1$, then $B_{p_1}^{(m,l)} \geq B_{p_2}^{(m,l)}$.

$B_p^{(m,l)}$ is defined as

$$B_p^{(m,l)} = \begin{cases} \min\limits_{\mathbf{g}_m \in \Gamma_m, \hat{\mathbf{e}}_{m+p-1} \in \Omega_{m+p-1}} \left\| \begin{pmatrix} \mathbf{M}_m^{(1)} \\ \mathbf{M}_{m,p}^{(2)} \\ \mathbf{M}_{m,p}^{(3)} \end{pmatrix} \mathbf{g}_m \right\|^2, & \text{if } p \geq m \\ \min\limits_{\mathbf{g}_m \in \Gamma_m, \hat{\mathbf{e}}_{m+p-1} \in \Omega_{m+p-1}} \left\| \begin{pmatrix} \mathbf{M}_m^{(1)} \\ \mathbf{M}_{m,p}^{(3)} \end{pmatrix} \mathbf{g}_m \right\|^2, & \text{if } p = m - 1 \end{cases} \tag{5.88}$$

where

$$\hat{\mathbf{e}}_{m+p-1} = (\hat{e}_0, \hat{e}_1, \ldots, \hat{e}_{m+p-2})^T \in \Omega_{m+p-1}, \tag{5.89}$$

Ω_n is the set of possible error events of length n, which for PAM modulations, as considered in this paper, can be expressed as

$$\begin{aligned} \Omega_n \quad = \quad & \Big\{ (y_0, y_1, \ldots, y_{n-1})^T \in \mathbb{R}^{n \times 1} : \\ & y_0, y_1, \cdots, y_{n-1} \in \{-l, -l+1, \cdots, l\}, y_0, y_{n-1} \neq 0 \Big\}, \end{aligned} \tag{5.90}$$

$\mathbf{g}_m$ is a channel description of length m, the domain of which is given by

$$\mathbf{g}_m = (g_0, g_1, \ldots, g_{m-1})^T \in \Gamma_m, \tag{5.91}$$

where

$$\Gamma_m = \left\{ (y_0, y_1, \ldots, y_{m-1})^T \in \mathbb{C}^{m \times 1} : \sum_{k=0}^{m-1} |y_k|^2 = 1 \right\}, \tag{5.92}$$

$\mathbf{M}_m^{(1)}$ is an $m-1$ by m matrix whose (i,j) entry is equal to $\hat{e}_{i-j}$ if $i \geq j$ and equal to 0 if $i < j$, $\mathbf{M}_{m,p}^{(2)}$ is a $p-m+1$ by m matrix whose (i,j) entry is equal to $\hat{e}_{m+i-j-1}$, and $\mathbf{M}_{m,p}^{(3)}$ is an $m-1$ by m matrix whose (i,j) entry is equal to $\hat{e}_{m+p+i-j-1}$ if $i \leq j-1$ and equal to 0 if $i > j-1$.

Since $p_1 > p_2 \geq m-1$, $\mathbf{M}^{(2)}_{m,p_1}$ may be expressed as

$$\mathbf{M}^{(2)}_{m,p_1} = \begin{bmatrix} \hat{\mathbf{M}}^{(2)}_{m,p_1} \\ \tilde{\mathbf{M}}^{(2)}_{m,p_1} \end{bmatrix} \tag{5.93}$$

where $\hat{\mathbf{M}}^{(2)}_{m,p_1}$ is a sub-matrix of $\mathbf{M}^{(2)}_{m,p_1}$ containing the first (p_2-m+1) rows, and $\tilde{\mathbf{M}}^{(2)}_{m,p_1}$ containing the last (p_1-p_2) rows. The proof (which is presented in the paper) of the theorem then requires that

$$\min_{\mathbf{g}_m \in \Gamma_m, \hat{\mathbf{e}}_{m+p-1} \in \Omega_{m+p-1}} \left\| \begin{pmatrix} \mathbf{M}^{(1)}_m \\ \hat{\mathbf{M}}^{(2)}_{m,p} \\ \mathbf{M}^{(3)}_{m,p} \end{pmatrix} \mathbf{g}_m \right\|^2 = B^{(m,l)}_{p_2} \tag{5.94}$$

However, this is not always true. The left and right hand sides of (5.94) (when $p \geq m-1$) are

$$\left[\begin{array}{cccc} e_0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ e_{m-2} & \dots & e_0 & 0 \\ \hline e_{m-1} & e_{m-2} & \dots & e_0 \\ \vdots & \ddots & \ddots & \vdots \\ e_{p_2-1} & \dots & & e_{p_2-m} \\ \hline 0 & e_{m+p_1-2} & \dots & e_{p_1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & e_{m+p_1-2} \end{array}\right] \neq \left[\begin{array}{cccc} e_0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ e_{m-2} & \dots & e_0 & 0 \\ \hline e_{m-1} & e_{m-2} & \dots & e_0 \\ \vdots & \ddots & \ddots & \vdots \\ e_{p_2-1} & \dots & & e_{p_2-m} \\ \hline 0 & e_{m+p_2-2} & \dots & e_{p_2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & e_{m+p_2-2} \end{array}\right]. \tag{5.95}$$

The last block (corresponding to $\mathbf{M}^{(3)}_{m,p_1}$) of each matrix is not the same, and the operation

$$\min_{\mathbf{g}_m \in \Gamma_m, \hat{\mathbf{e}}_{m+p-1} \in \Omega_{m+p-1}} \|\mathbf{M}\mathbf{g}_m\|^2 \tag{5.96}$$

will not usually produce equality of these two expressions.

Chapter 6

Conclusions and Further Research

IN this chapter the conclusions of this research are summarised, and further research directions are outlined.

6.1 Conclusions

The problem of real time characterisation of the mobile fading channel has been considered in detail. The primary objective has been to determine the extent to which the mobile channel can be predicted for short time periods into the future. The method has been to develop models which firstly allow estimation of the model parameters, and secondly have parameters which do not change rapidly, and so allow channel prediction.

One might expect that the extent of spatial correlation would provide a bound for the range of accurate prediction. It was concluded in Section 2.3.5 that this is not true in general. The range of accurate prediction may be very long or very short in scenarios where the extent of spatial correlation is either large or small.

In the case of a moving receiver and far field point sources, and for moderate SNR, the subspace algorithms (Section 3.2) provide an excellent method of estimating the parameters of a simple channel model which allows good channel prediction. These may be enhanced by gradient methods (Section 3.1.2).

It has been shown in Sections 4.2.3 and 4.4 that the range of channel prediction depends *critically* on the number of significant scatterers. This conclusion is supported by both simulations and considerations of the Cramer Rao bound. The idea is widespread in the literature that in many circumstances in mobile communications there are only a few significant scatterers. In Section 2.3.5 it has been shown that the reason usually given for this assertion is not well founded.

The parameters of a far field model which correspond to near field sources may change too rapidly to allow long range prediction. While array processing techniques

may be used to estimate the parameters of a near field model, it has been shown in Section 3.2.2 that they are problematical in the case of the single array snapshot provided by the trajectory of a single moving receiver.

The effect of even moderately rough surfaces has also been shown in Section 4.5 to be a limiting factor in long range channel prediction.

It is known (see Section 4.7) that the field in a relatively small region can be recreated to within a given tolerance with only a small number of sources. This result has been used in Section 4.7.6 to show that the error encountered in extrapolation of the field beyond the region grows rapidly with extrapolation distance.

In the light of the above discoveries, and the experience of prediction with measured channel data (Section 4.3) it is concluded that while prediction of the mobile multipath channel may indeed be feasible in some situations, for only a small proportion of the time can a system predict more than one fade into the future. Factors which can limit the range of useful prediction include the presence of a large number of significant scatterers, macroscopic changes in the scattering environment, and the presence of rough surfaces.

An alternative method of overcoming multipath fading is space diversity, discussed in Chapter 2. Given the advanced state this technique, and the feasibility of incorporating multiple antennas on a compact handset (shown by the experiments in Section 2.4), and the limitations of prediction discovered in the research so far, it is likely that this technique will enter the market place long before channel prediction. However, channel prediction may yet have an alternative or complementary role; alternative, where multiple antennas may not be feasible; and complementary, where the multiple antennas may be treated as an array to provide superior information on the scattering environment than could be obtained by one moving antenna alone, or where prediction can optimise the channel choice made in a selection diversity system.

It has been shown in Section 5.1 that predicted channel information can be used to predict system performance, although the relationship between channel information and system performance can be a non-trivial one. A method has been presented which allows the performance to be calculated in a relatively efficient manner provided that the "contributing" error events are known. An algorithm for finding these has been presented in Section 5.3, along with results which may be useful for some simple modulation schemes (Appendix A). The problem of estimating the system performance in the case of channel estimation error has also been examined and it has been shown in Section 5.6 how reliable indications of system performance can be obtained.

6.2 Future Research

Based on the material in this thesis three specific research directions are here proposed which could lead to a deeper understanding of real time channel characterisation.

The nature of the mobile channel

The actual probability density of the power of sources in a typical mobile scenario remains largely unknown. It has been pointed out that many researchers have assumed with perhaps inadequate justification that there are only a few significant discrete sources. This represents one extreme of the density function. The other extreme is a completely diffuse (if not omni-directional) field. This problem requires considerably more investigation that it has received to date. It would require a considerable investment in both measurement equipment and either arrays or directional antennas.

With better knowledge of the channel, models based on distributed sources may also be developed and usefully applied.

The frequency of macroscopic changes in the scattering environment (e.g., changes in shadowing) has a direct impact on the frequency of prediction failures. This requires measurement equipment capable of storing very long data records. In a private correspondence with Torbjörn Ekman and Mikael Sternad (authors of [43]) it would appear that they are investigating this problem.

Adaptive algorithms have not been considered in this thesis. In this context, adaptive algorithms are ones which can efficiently use previous estimates of model parameters, and adjust these as new data becomes available. If a model is found for which the parameters vary relatively slowly, then adaptive techniques ought to work well. As mentioned in Section 4.2.7, the linear model is readily adapted to an adaptive algorithm. Future work should include some assessment of the effectiveness of such a technique. A situation in which adaptive algorithms will not work well is of course the situation where there are abrupt and radical changes in the scattering environment, and thus the model parameters. Some results of the application of adaptive techniques have already been presented by the authors mentioned above.

The use of support vector machines (see Section 3.7) for either regression or classification of channel data has not been investigated at great length. This is another area in which significant progress may be made. Progress is of course subject to certain bounds, which are the subject of the next suggested research area.

Bounds

Some bounds have already been derived based on several assumptions about the nature of the mobile channel. Other approaches may also be helpful. For instance, bounds on the prediction range may be found based on theoretical considerations of the mutual information for near field situations, rather than trying to estimate the mutual information from data records.

Another area of useful research is into the *confidence* with which predictions can be made. It would be very useful for a channel predictor to be able to estimate this confidence, so that predicted channel information may be used judiciously.

Applications

The area of applications has been touched in Chapter 5, but there is still a lot more which could be explored. Some obvious steps include simulation of systems which use predicted channel information in some way — for power control or data rate control, or request for change of frequency for example. Another system which should be investigated is a Multiple Input Multiple Output system. Such systems are of great interest to researchers at present, and real time channel characterisation if not channel prediction may have an important part to play in optimising their performance.

The problem of *efficiently* calculating the error rate from a channel estimate when the channel estimate is in error requires further investigation.

The prediction accuracy requirements for various applications need to be quantified in greater detail than has been done here.

Equalisation based on a channel model with only slowly varying parameters (proposed in Section 1.3.3) has not been explored. Since only a small range of prediction is required for such a technique to perform well, this may be profitable research area.

Appendix A

Contributing Error Events

IN this appendix is presented the set of error events which will contain the worst case error event for *any* channel. It should be noted that in every case each listed error event actually represents all possible error events that can be generated by rotation of the listed sequence around the centre of the difference alphabet. For instance, for QPSK, the error event $(1, j)$ has the same distance for a given channel as has $(j, -1)$, $(-1, -j)$ and $(-j, 1)$, but only the error event $(1, j)$ is listed.

Note also that the complete error event actually consists of the listed sequence followed by $L - 1$ zeros. It is only after this number of correctly interpreted symbols that the paths of a Viterbi equaliser can be considered to have converged, and the error event is complete.

The number in each entry is the percentage of unit power channels for which the corresponding error event has the smallest distance. These percentages do not correspond to a uniform distribution of all channels since the grid used has more points close to the poles of the unit hyper-sphere, and fewer points near its equator. The percentages do however give some indication of the likely contribution of each error event to the overall error rate.

A dash indicates the sequence does not have the smallest distance for *any* channels, whereas a zero indicates that it has the smallest distance for less than 0.05% of the

Modulation Type	Symmetric Sequences	Non-symmetric sequences
BPSK	18	16
4PAM	14	8
8PAM	12	7
QPSK	14	9
8PSK	10	5
16QAM	10	5
16PSK	8	4

Table A.1: Length of sequences searched for contributing error events

channels tried. Some sequences which appear added at the end of each table are those which produce a minimum worst case distance, and therefore must contribute to the union, but were not detected in the search.

The maximum length of error sequences for which the search was conducted are listed in table A.1.

A.1 BPSK

L	2	3	4	5	6	7	8	9	10	11
1	100	83	83.6	81.9	82.6	82.4	82.4	79.9	80.7	84.3
1, 1	-	8.5	5.2	4.7	4.6	4.7	4.9	4.8	3.6	3.7
1, −1	-	8.5	5.2	4.7	4.6	4.7	4.9	4.8	3.6	3.7
1, 1, 1	-	-	3	2.2	1.8	1.6	1.5	2.1	2.1	1.6
1, −1, 1	-	-	3	2.2	1.8	1.6	1.5	2.1	2.1	1.6
1, 0, 1	-	-	-	1.7	1.3	1.1	1	1.3	1.2	1
1, 0, −1	-	-	-	1.1	0.9	0.7	0.6	1.1	0.8	0.4
1, 1, 1, 1	-	-	-	0.7	0.7	0.6	0.6	0.7	1.2	0.7
1, −1, 1, −1	-	-	-	0.7	0.7	0.6	0.6	0.7	1.2	0.7
1, 1, 1, 1, 1	-	-	-	-	0.3	0.4	0.3	0.3	0.5	0.4
1, −1, 1, −1, 1	-	-	-	-	0.3	0.4	0.3	0.3	0.5	0.4
1, 1, 0, −1, −1	-	-	-	-	0.2	0.2	0.1	0.2	0.3	0.1
1, −1, 0, 1, −1	-	-	-	-	0.2	0.2	0.1	0.2	0.3	0.1
1, 0, 1, 0, 1	-	-	-	-	-	0.2	0.2	0.2	0.2	0.1
1, 0, 0, 1	-	-	-	-	-	0.2	0.1	0.2	0.1	0.1
1, 0, 0, −1	-	-	-	-	-	0.2	0.1	0.2	0.1	0.1
1, 0, −1, 0, 1	-	-	-	-	-	0.2	0.1	0.2	0.2	0.1
1, 1, 1, 1, 1, 1	-	-	-	-	-	0.2	0.2	0.1	0.2	0.2
1, −1, 1, −1, 1, −1	-	-	-	-	-	0.2	0.2	0.1	0.2	0.2
1, 0, 1, 1, 0, 1	-	-	-	-	-	0	-	0	0	0
1, 0, 1, −1, 0, −1	-	-	-	-	-	0	-	0	0	0
1, 1, 1, 0, −1, −1, −1	-	-	-	-	-	0	0	0	0	0
1, −1, 1, 0, −1, 1, −1	-	-	-	-	-	0	0	0	0	0
1, 1, −1, −1, 1, 1	-	-	-	-	-	0	0	0	-	-
1, −1, −1, 1, 1, −1	-	-	-	-	-	0	0	0	-	-
1, 1, 1, 1, 1, 1, 1	-	-	-	-	-	-	0.1	0.1	0.1	0.1
1, −1, 1, −1, 1, −1, 1	-	-	-	-	-	-	0.1	0.1	0.1	0.1
1, 1, 0, 0, 1, 1	-	-	-	-	-	-	0	0	0.1	0
1, −1, 0, 0, 1, −1	-	-	-	-	-	-	0	0	0.1	0
1, 1, 0, −1, 0, 1, 1	-	-	-	-	-	-	0	0	0	0
1, −1, 0, 1, 0, −1, 1	-	-	-	-	-	-	0	0	0	0
1, 1, −1, −1, 1, 1, −1, −1	-	-	-	-	-	-	0	0	-	-
1, −1, −1, 1, 1, −1, −1, 1	-	-	-	-	-	-	0	0	-	-
1, 0, 1, 0, 1, 0, 1	-	-	-	-	-	-	-	0.1	0.1	0
1, 0, −1, 0, 1, 0, −1	-	-	-	-	-	-	-	0.1	0.1	0
1, 1, 1, 1, 1, 1, 1, 1	-	-	-	-	-	-	-	0	0	0
1, −1, 1, −1, 1, −1, 1, −1	-	-	-	-	-	-	-	0	0	0
1, 0, 0, 0, 1	-	-	-	-	-	-	-	0	0	0
1, 0, 0, 0, −1	-	-	-	-	-	-	-	0	0	0
1, 1, 0, −1, −1, 0, 1, 1	-	-	-	-	-	-	-	0	0.1	0
1, −1, 0, 1, −1, 0, 1, −1	-	-	-	-	-	-	-	0	0.1	0
1, 1, 0, −1, −1, −1, 0, 1, 1	-	-	-	-	-	-	-	0	0	0
1, −1, 0, 1, −1, 1, 0, −1, 1	-	-	-	-	-	-	-	0	0	0
1, 1, −1, −1	-	-	-	-	-	-	-	0	-	0
1, −1, −1, 1	-	-	-	-	-	-	-	0	-	0
1, 1, 0, −1, 0, 1, 0, −1, −1	-	-	-	-	-	-	-	0	0	-
1, −1, 0, 1, 0, −1, 0, 1, −1	-	-	-	-	-	-	-	0	0	-
1, 0, 0, 1, 0, 0, 1	-	-	-	-	-	-	-	-	0	0
1, 0, 0, −1, 0, 0, 1	-	-	-	-	-	-	-	-	0	0
1, 0, −1, −1, 1, 1, 0, −1	-	-	-	-	-	-	-	-	0	-
1, 0, −1, 1, 1, −1, 0, 1	-	-	-	-	-	-	-	-	0	-
1, 1, 0, 0, 0, −1, −1	-	-	-	-	-	-	-	-	0	0
1, −1, 0, 0, 0, 1, −1	-	-	-	-	-	-	-	-	0	0
1, 1, 1, 0, 0, −1, −1, −1	-	-	-	-	-	-	-	-	0	-
1, −1, 1, 0, 0, 1, −1, 1	-	-	-	-	-	-	-	-	0	-
1, 1, 1, 1, 1, 1, 1, 1, 1	-	-	-	-	-	-	-	-	0	0
1, −1, 1, −1, 1, −1, 1, −1, 1	-	-	-	-	-	-	-	-	0	0
1, 1, 1, 1, 0, −1, −1, −1, −1	-	-	-	-	-	-	-	-	0	0
1, −1, 1, −1, 0, 1, −1, 1, −1	-	-	-	-	-	-	-	-	0	0
1, 1, 0, −1, −1, −1, −1, 0, 1, 1	-	-	-	-	-	-	-	-	0	-
1, −1, 0, 1, −1, 1, −1, 0, 1, −1	-	-	-	-	-	-	-	-	0	-

L	2	3	4	5	6	7	8	9	10	11
$1,1,1,0,-1,-1,0,1,1,1$	-	-	-	-	-	-	-	-	0	-
$1,-1,1,0,-1,1,0,-1,1,-1$	-	-	-	-	-	-	-	-	0	-
$1,0,-1,0,1,0,-1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,0,1,0,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,0,0,0,-1,0,-1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,0,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,0,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,1,1,1,1,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,-1,-1,-1,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,-1,1,-1,1,-1,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,-1,1,-1,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,-1,-1,0,1,1,0,-1$	-	-	-	-	-	-	-	-	-	0
$1,0,-1,1,0,-1,1,0,-1$	-	-	-	-	-	-	-	-	-	0

A.2 4PAM

L	2	3	4	5	6	7	8	9	10	11
1	100	97.1	83.6	81.9	82.6	82.4	82.4	79.9	80.7	84.3
$1,1$	-	1.4	5	4.6	4.5	4.5	4.8	4.6	3.6	3.7
$1,-1$	-	1.4	5	4.6	4.5	4.5	4.8	4.6	3.6	3.7
$1,2,2,1$	-	0	0.2	0.1	0.1	0.1	0	0.2	-	-
$1,-2,2,-1$	-	0	0.2	0.1	0.1	0.1	0	0.2	-	-
$1,1,1$	-	-	3	2.1	1.7	1.5	1.5	2	2.1	1.6
$1,-1,1$	-	-	3	2.1	1.7	1.5	1.5	2	2.1	1.6
$1,2,3,3,2,1$	-	-	0	0	0	0	0	0.1	-	-
$1,-2,3,-3,2,-1$	-	-	0	0	0	0	0	0.1	-	-
$1,2,3,3,3,2,1$	-	-	0	0	0	0	0	0	-	-
$1,-2,3,-3,3,-2,1$	-	-	0	0	0	0	0	0	-	-
$1,0,1$	-	-	-	1.7	1.3	1	1	1.3	1.2	1
$1,0,-1$	-	-	-	1.1	0.9	0.7	0.6	1.1	0.8	0.4
$1,1,1,1$	-	-	-	0.7	0.7	0.6	0.6	0.7	1.2	0.7
$1,-1,1,-1$	-	-	-	0.7	0.7	0.6	0.6	0.7	1.2	0.7
$1,1,-1,-2,0,2,1,-1,-1$	-	-	-	0	-	0	0	0	-	0.1
$1,-1,-1,2,0,-2,1,1,-1$	-	-	-	0	-	0	0	0	-	0.1
$1,2,1,-1,-2,-1$	-	-	-	0	0	0	0	-	-	-
$1,-2,1,1,-2,1$	-	-	-	0	0	0	0	-	-	-
$1,2,2,2,2,1$	-	-	-	0	0	0	0	0	-	-
$1,-2,2,-2,2,-1$	-	-	-	0	0	0	0	0	-	-
$1,2,1$	-	-	-	0	0	0	-	-	-	-
$1,-2,1$	-	-	-	0	0	0	-	-	-	-
$1,2,2,2,1$	-	-	-	0	0	0	0	0	-	-
$1,-2,2,-2,1$	-	-	-	0	0	0	0	0	-	-
$1,1,1,1,1$	-	-	-	-	0.3	0.4	0.3	0.3	0.5	0.4
$1,-1,1,-1,1$	-	-	-	-	0.3	0.4	0.3	0.3	0.5	0.4
$1,1,0,-1,-1$	-	-	-	-	0.2	0.1	0.1	0.2	0.3	0.1
$1,-1,0,1,-1$	-	-	-	-	0.2	0.1	0.1	0.2	0.3	0.1
$1,0,2,0,2,0,1$	-	-	-	-	0	0	0	0	-	-
$1,0,-2,0,2,0,-1$	-	-	-	-	0	0	0	0	-	-
$1,2,3,3,2,0,-2,-3,-3,-2,-1$	-	-	-	-	0	0	0	0	-	0
$1,-2,3,-3,2,0,-2,3,-3,2,-1$	-	-	-	-	0	0	0	0	-	0
$1,2,1,-1,-1,1,2,1$	-	-	-	-	0	0	-	-	-	-
$1,-2,1,1,-1,-1,2,-1$	-	-	-	-	0	0	-	-	-	-
$1,0,1,0,1$	-	-	-	-	-	0.2	0.2	0.2	0.2	0.1
$1,0,0,1$	-	-	-	-	-	0.2	0.1	0.2	0.1	0.1
$1,0,0,-1$	-	-	-	-	-	0.2	0.1	0.2	0.1	0.1
$1,0,-1,0,1$	-	-	-	-	-	0.2	0.1	0.2	0.2	0.1
$1,1,1,1,1,1$	-	-	-	-	-	0.2	0.2	0.1	0.2	0.2
$1,-1,1,-1,1,-1$	-	-	-	-	-	0.2	0.2	0.1	0.2	0.2
$1,0,1,1,0,1$	-	-	-	-	-	0	-	0	0	0
$1,0,1,-1,0,-1$	-	-	-	-	-	0	-	0	0	0
$1,1,2,1,1$	-	-	-	-	-	0	0	0	0.1	0
$1,-1,2,-1,1$	-	-	-	-	-	0	0	0	0.1	0
$1,1,2,2,2,2,1,1$	-	-	-	-	-	0	0	0	0	-
$1,-1,2,-2,2,-2,1,-1$	-	-	-	-	-	0	0	0	0	-
$1,1,1,0,-1,-1,-1$	-	-	-	-	-	0	0	0	0	0
$1,-1,1,0,-1,1,-1$	-	-	-	-	-	0	0	0	0	0
$1,1,-1,-1,1,1$	-	-	-	-	-	0	0	0	-	-
$1,-1,-1,1,1,-1$	-	-	-	-	-	0	0	0	-	-
$1,2,2,1,-1,-2,-2,-1$	-	-	-	-	-	0	-	-	-	-
$1,-2,2,-1,-1,2,-2,1$	-	-	-	-	-	0	-	-	-	-

L	2	3	4	5	6	7	8	9	10	11
$1, 2, 2, 2, 2, 2, 1$	-	-	-	-	-	0	-	0	-	-
$1, -2, 2, -2, 2, -2, 1$	-	-	-	-	-	0	-	0	-	-
$1, 2, 3, 3, 3, 3, 2, 1$	-	-	-	-	-	0	0	0	-	-
$1, -2, 3, -3, 3, -3, 3, -2, 1$	-	-	-	-	-	0	0	0	-	-
$1, 2, 2, 2, 3, 3, 2, 2, 2, 1$	-	-	-	-	-	0	-	-	-	-
$1, -2, 2, -2, 3, -3, 2, -2, 2, -1$	-	-	-	-	-	0	-	-	-	-
$1, 2, 1, -1, -2, -2, -1, 1, 2, 1$	-	-	-	-	-	0	-	-	-	-
$1, -2, 1, 1, -2, 2, -1, -1, 2, -1$	-	-	-	-	-	0	-	-	-	-
$1, 1, 1, 1, 1, 1, 1$	-	-	-	-	-	-	0.1	0.1	0.1	0.1
$1, -1, 1, -1, 1, -1, 1$	-	-	-	-	-	-	0.1	0.1	0.1	0.1
$1, 1, 0, 0, 1, 1$	-	-	-	-	-	-	0	0	0.1	0
$1, -1, 0, 0, 1, -1$	-	-	-	-	-	-	0	0	0.1	0
$1, 1, 0, -1, 0, 1, 1$	-	-	-	-	-	-	0	0	0	0
$1, -1, 0, 1, 0, -1, 1$	-	-	-	-	-	-	0	0	0	0
$1, 1, 2, 2, 2, 1, 1$	-	-	-	-	-	-	0	0	0	0
$1, -1, 2, -2, 2, -1, 1$	-	-	-	-	-	-	0	0	0	0
$1, 1, 1, 2, 2, 2, 2, 2, 1, 1, 1$	-	-	-	-	-	-	0	0	0	0
$1, -1, 1, -2, 2, -2, 2, -2, 2, -1, 1, -1$	-	-	-	-	-	-	0	0	0	0
$1, 2, 2, 1, 0, -1, -2, -2, -1$	-	-	-	-	-	-	0	0	-	-
$1, -2, 2, -1, 0, 1, -2, 2, -1$	-	-	-	-	-	-	0	0	-	-
$1, 1, -1, -1, 1, 1, -1, -1$	-	-	-	-	-	-	0	-	-	-
$1, -1, -1, 1, 1, -1, -1, 1$	-	-	-	-	-	-	0	-	-	-
$1, 0, 1, 0, 1, 0, 1$	-	-	-	-	-	-	-	0.1	0.1	0
$1, 0, -1, 0, 1, 0, -1$	-	-	-	-	-	-	-	0.1	0.1	0
$1, 1, 1, 1, 1, 1, 1, 1, 1$	-	-	-	-	-	-	-	0	0	0
$1, -1, 1, -1, 1, -1, 1, -1$	-	-	-	-	-	-	-	0	0	0
$1, 0, 0, 0, 1$	-	-	-	-	-	-	-	0	0	0
$1, 0, 0, 0, -1$	-	-	-	-	-	-	-	0	0	0
$1, 1, 0, -1, -1, 0, 1, 1$	-	-	-	-	-	-	-	0	0.1	0
$1, -1, 0, 1, -1, 0, 1, -1$	-	-	-	-	-	-	-	0	0.1	0
$1, 1, 0, -1, -1, -1, 0, 1, 1$	-	-	-	-	-	-	-	0	0	0
$1, -1, 0, 1, -1, 1, 0, -1, 1$	-	-	-	-	-	-	-	0	0	0
$1, 1, -1, 0, 2, 0, -1, 1, 1$	-	-	-	-	-	-	-	0	-	-
$1, -1, -1, 0, 2, 0, -1, -1, 1$	-	-	-	-	-	-	-	0	-	-
$1, 2, 3, 3, 3, 3, 3, 3, 2, 1$	-	-	-	-	-	-	-	0	-	-
$1, -2, 3, -3, 3, -3, 3, -3, 2, -1$	-	-	-	-	-	-	-	0	-	-
$1, 1, 1, 0, -1, -2, -2, -1, 0, 1, 1, 1$	-	-	-	-	-	-	-	0	-	0
$1, -1, 1, 0, -1, 2, -2, 1, 0, -1, 1, -1$	-	-	-	-	-	-	-	0	-	0
$1, 1, 2, 2, 3, 2, 2, 1, 1$	-	-	-	-	-	-	-	0	-	-
$1, -1, 2, -2, 3, -2, 2, -1, 1$	-	-	-	-	-	-	-	0	-	-
$1, 1, 2, 2, 2, 2, 2, 2, 1, 1$	-	-	-	-	-	-	-	0	0	0
$1, -1, 2, -2, 2, -2, 2, -1, 1$	-	-	-	-	-	-	-	0	0	0
$1, 1, 1, 2, 2, 2, 2, 1, 1, 1$	-	-	-	-	-	-	-	0	-	-
$1, -1, 1, -2, 2, -2, 2, -1, 1, -1$	-	-	-	-	-	-	-	0	-	-
$1, 0, -2, -1, 2, 2, -1, -2, 0, 1$	-	-	-	-	-	-	-	0	-	-
$1, 0, -2, 1, 2, -2, -1, 2, 0, -1$	-	-	-	-	-	-	-	0	-	-
$1, 1, 0, -1, 0, 1, 0, -1, -1$	-	-	-	-	-	-	-	0	0	-
$1, -1, 0, 1, 0, -1, 0, 1, -1$	-	-	-	-	-	-	-	0	0	-
$1, 0, 0, 1, 0, 0, 1$	-	-	-	-	-	-	-	-	0	0
$1, 0, 0, -1, 0, 0, 1$	-	-	-	-	-	-	-	-	0	0
$1, 0, -1, -1, 1, 1, 0, -1$	-	-	-	-	-	-	-	-	0	-
$1, 0, -1, 1, 1, -1, 0, 1$	-	-	-	-	-	-	-	-	0	-
$1, 1, 1, 2, 1, 1, 1$	-	-	-	-	-	-	-	-	0	0
$1, -1, 1, -2, 1, -1, 1$	-	-	-	-	-	-	-	-	0	0
$1, 1, 0, 0, 0, -1, -1$	-	-	-	-	-	-	-	-	0	0
$1, -1, 0, 0, 0, 1, -1$	-	-	-	-	-	-	-	-	0	0
$1, 1, 1, 1, 2, 1, 1, 1, 1$	-	-	-	-	-	-	-	-	0	0
$1, -1, 1, -1, 2, -1, 1, -1, 1$	-	-	-	-	-	-	-	-	0	0
$1, 1, 1, 0, 0, -1, -1, -1$	-	-	-	-	-	-	-	-	0	-
$1, -1, 1, 0, 0, 1, -1, 1$	-	-	-	-	-	-	-	-	0	-
$1, 1, 1, 1, 1, 1, 1, 1, 1$	-	-	-	-	-	-	-	-	0	0
$1, -1, 1, -1, 1, -1, 1, -1, 1$	-	-	-	-	-	-	-	-	0	0
$1, 1, 1, 1, 0, -1, -1, -1, -1$	-	-	-	-	-	-	-	-	0	0
$1, -1, 1, -1, 0, 1, -1, 1, -1$	-	-	-	-	-	-	-	-	0	0
$1, 1, 0, -1, -1, -1, -1, 0, 1, 1$	-	-	-	-	-	-	-	-	0	-
$1, -1, 0, 1, -1, 1, -1, 0, 1, -1$	-	-	-	-	-	-	-	-	0	-
$1, 1, 1, 0, -1, -1, 0, 1, 1, 1$	-	-	-	-	-	-	-	-	0	-
$1, -1, 1, 0, -1, 1, 0, -1, 1, -1$	-	-	-	-	-	-	-	-	0	-
$1, 1, -1, -1$	-	-	-	-	-	-	-	-	-	0
$1, -1, -1, 1$	-	-	-	-	-	-	-	-	-	0
$1, 0, -1, 0, 1, 0, -1, 0, 1$	-	-	-	-	-	-	-	-	-	0
$1, 1, 0, 1, 2, 1, 0, 1, 1$	-	-	-	-	-	-	-	-	-	0
$1, -1, 0, -1, 2, -1, 0, -1, 1$	-	-	-	-	-	-	-	-	-	0
$1, 0, 1, 0, 1, 0, 1, 0, 1$	-	-	-	-	-	-	-	-	-	0
$1, 0, 1, 0, 0, 0, -1, 0, -1$	-	-	-	-	-	-	-	-	-	0

L	2	3	4	5	6	7	8	9	10	11
$1,1,1,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,0,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,0,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,1,1,1,1,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,-1,-1,-1,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,-1,1,-1,1,-1,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,-1,1,-1,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,-1,-1,0,1,1,0,-1$	-	-	-	-	-	-	-	-	-	0
$1,0,-1,1,0,-1,1,0,-1$	-	-	-	-	-	-	-	-	-	0
$1,3,3,0,-2,0,3,3,1$	-	-	-	-	-	0	0	0	0	0
$1,-3,3,0,-2,0,3,-3,1$	-	-	-	-	-	0	0	0	0	0

A.3 8PAM

L	2	3	4	5	6	7	8	9	10	11
1	100	83	83.6	81.9	82.6	82.4	82.4	79.9	80.7	84.3
$1,1$	-	8.2	5	4.6	4.5	4.5	4.8	4.6	3.6	3.7
$1,-1$	-	8.2	5	4.6	4.5	4.5	4.8	4.6	3.6	3.7
$1,2,2,1$	-	0.4	0.2	0.1	0.1	0.1	0	0.2	-	-
$1,-2,2,-1$	-	0.4	0.2	0.1	0.1	0.1	0	0.2	-	-
$1,1,1$	-	-	3	2.1	1.7	1.5	1.5	2	2.1	1.6
$1,-1,1$	-	-	3	2.1	1.7	1.5	1.5	2	2.1	1.6
$1,2,3,3,2,1$	-	-	0	0	0	0	0	0.1	-	-
$1,-2,3,-3,2,-1$	-	-	0	0	0	0	0	0.1	-	-
$1,2,3,4,4,4,3,2,1$	-	-	0	0	0	0	0	0	-	-
$1,-2,3,-4,4,-4,3,-2,1$	-	-	0	0	0	0	0	0	-	-
$1,3,5,6,6,5,3,1$	-	-	0	0	0	0	-	-	-	-
$1,-3,5,-6,6,-5,3,-1$	-	-	0	0	0	0	-	-	-	-
$1,2,3,3,3,2,1$	-	-	0	0	0	0	0	0	-	-
$1,-2,3,-3,3,-2,1$	-	-	0	0	0	0	0	0	-	-
$1,0,1$	-	-	-	1.7	1.3	1	1	1.3	1.2	1
$1,0,-1$	-	-	-	1.1	0.9	0.7	0.6	1.1	0.8	0.4
$1,1,1,1$	-	-	-	0.7	0.7	0.6	0.6	0.7	1.2	0.7
$1,-1,1,-1$	-	-	-	0.7	0.7	0.6	0.6	0.7	1.2	0.7
$1,1,-1,-2,0,2,1,-1,-1$	-	-	-	0	-	0	0	0	-	0.1
$1,-1,-1,2,0,-2,1,1,-1$	-	-	-	0	-	0	0	0	-	0.1
$1,2,1,-1,-2,-1$	-	-	-	0	0	0	0	-	-	-
$1,-2,1,1,-2,1$	-	-	-	0	0	0	0	-	-	-
$1,2,2,2,2,1$	-	-	-	0	0	0	0	0	-	-
$1,-2,2,-2,2,-1$	-	-	-	0	0	0	0	0	-	-
$1,2,3,4,5,5,5,4,3,2,1$	-	-	-	0	0	0	0	-	-	-
$1,-2,3,-4,5,-5,5,-4,3,-2,1$	-	-	-	0	0	0	0	-	-	-
$1,2,3,4,4,3,2,1$	-	-	-	0	0	0	-	0	-	-
$1,-2,3,-4,4,-3,2,-1$	-	-	-	0	0	0	-	0	-	-
$1,2,3,4,5,6,6,5,4,3,2,1$	-	-	-	0	0	-	-	-	-	-
$1,-2,3,-4,5,-6,6,-5,4,-3,2,-1$	-	-	-	0	0	-	-	-	-	-
$1,2,3,4,5,5,4,3,2,1$	-	-	-	0	0	-	-	0	-	-
$1,-2,3,-4,5,-5,4,-3,2,-1$	-	-	-	0	0	-	-	0	-	-
$1,2,1$	-	-	-	0	0	0	-	-	-	-
$1,-2,1$	-	-	-	0	0	0	-	-	-	-
$1,3,5,6,5,3,1$	-	-	-	0	0	0	0	-	-	-
$1,-3,5,-6,5,-3,1$	-	-	-	0	0	0	0	-	-	-
$1,2,2,2,1$	-	-	-	0	0	0	0	0	-	-
$1,-2,2,-2,1$	-	-	-	0	0	0	0	0	-	-
$1,1,1,1,1$	-	-	-	-	0.3	0.4	0.3	0.3	0.5	0.4
$1,-1,1,-1,1$	-	-	-	-	0.3	0.4	0.3	0.3	0.5	0.4
$1,1,0,-1,-1$	-	-	-	-	0.2	0.1	0.1	0.2	0.3	0.1
$1,-1,0,1,-1$	-	-	-	-	0.2	0.1	0.1	0.2	0.3	0.1
$1,0,2,0,2,0,1$	-	-	-	-	0	0	0	0	-	-
$1,0,-2,0,2,0,-1$	-	-	-	-	0	0	0	0	-	-
$1,2,3,3,2,0,-2,-3,-3,-2,-1$	-	-	-	-	0	0	0	0	-	0
$1,-2,3,-3,2,0,-2,3,-3,2,-1$	-	-	-	-	0	0	0	0	-	0
$1,2,1,-1,-1,1,2,1$	-	-	-	-	0	0	-	-	-	-
$1,-2,1,1,-1,-1,2,-1$	-	-	-	-	0	0	-	-	-	-
$1,2,3,4,4,4,4,3,2,1$	-	-	-	-	0	0	0	0	-	-
$1,-2,3,-4,4,-4,4,-3,2,-1$	-	-	-	-	0	0	0	0	-	-
$1,3,5,6,5,2,-2,-5,-6,-5,-3,-1$	-	-	-	0	0	0	-	-	-	-
$1,-3,5,-6,5,-2,-2,5,-6,5,-3,1$	-	-	-	0	0	0	-	-	-	-
$1,2,3,4,5,5,5,5,4,3,2,1$	-	-	-	-	0	0	-	0	-	-
$1,-2,3,-4,5,-5,5,-5,4,-3,2,-1$	-	-	-	-	0	0	-	0	-	-
$1,0,1,0,1$	-	-	-	-	-	0.2	0.2	0.2	0.2	0.1

L	2	3	4	5	6	7	8	9	10	11
$1,0,0,1$	-	-	-	-	-	0.2	0.1	0.2	0.1	0.1
$1,0,0,-1$	-	-	-	-	-	0.2	0.1	0.2	0.1	0.1
$1,0,-1,0,1$	-	-	-	-	-	0.2	0.1	0.2	0.2	0.1
$1,1,1,1,1,1$	-	-	-	-	-	0.2	0.2	0.1	0.2	0.2
$1,-1,1,-1,1,-1$	-	-	-	-	-	0.2	0.2	0.1	0.2	0.2
$1,0,1,1,0,1$	-	-	-	-	-	0	-	0	0	0
$1,0,1,-1,0,-1$	-	-	-	-	-	0	-	0	0	0
$1,1,2,1,1$	-	-	-	-	-	0	0	0	0.1	0
$1,-1,2,-1,1$	-	-	-	-	-	0	0	0	0.1	0
$1,1,2,2,2,2,1,1$	-	-	-	-	-	0	0	0	0	-
$1,-1,2,-2,2,-2,1,-1$	-	-	-	-	-	0	0	0	0	-
$1,1,1,0,-1,-1,-1$	-	-	-	-	-	0	0	0	0	0
$1,-1,1,0,-1,1,-1$	-	-	-	-	-	0	0	0	0	0
$1,1,-1,-1,1,1$	-	-	-	-	-	0	0	0	-	-
$1,-1,-1,1,1,-1$	-	-	-	-	-	0	0	0	-	-
$1,2,2,1,-1,-2,-2,-1$	-	-	-	-	-	0	-	-	-	-
$1,-2,2,-1,-1,2,-2,1$	-	-	-	-	-	0	-	-	-	-
$1,2,2,2,2,2,1$	-	-	-	-	-	0	-	0	-	-
$1,-2,2,-2,2,-2,1$	-	-	-	-	-	0	-	0	-	-
$1,2,3,3,3,3,3,2,1$	-	-	-	-	-	0	0	0	-	-
$1,-2,3,-3,3,-3,3,-2,1$	-	-	-	-	-	0	0	0	-	-
$1,2,2,2,3,3,2,2,2,1$	-	-	-	-	-	0	-	-	-	-
$1,-2,2,-2,3,-3,2,-2,2,-1$	-	-	-	-	-	0	-	-	-	-
$1,3,5,6,6,6,6,5,3,1$	-	-	-	-	-	0	-	-	-	-
$1,-3,5,-6,6,-6,6,-6,5,-3,1$	-	-	-	-	-	0	-	-	-	-
$1,2,4,5,6,5,4,2,1$	-	-	-	-	-	0	-	-	-	-
$1,-2,4,-5,6,-5,4,-2,1$	-	-	-	-	-	0	-	-	-	-
$1,2,1,-1,-2,-2,-1,1,2,1$	-	-	-	-	-	0	-	-	-	-
$1,-2,1,1,-2,2,-1,-1,2,-1$	-	-	-	-	-	0	-	-	-	-
$1,1,1,1,1,1,1$	-	-	-	-	-	-	0.1	0.1	0.1	0.1
$1,-1,1,-1,1,-1,1$	-	-	-	-	-	-	0.1	0.1	0.1	0.1
$1,1,0,0,1,1$	-	-	-	-	-	-	0	0	0.1	0
$1,-1,0,0,1,-1$	-	-	-	-	-	-	0	0	0.1	0
$1,1,0,-1,0,1,1$	-	-	-	-	-	-	0	0	0	0
$1,-1,0,1,0,-1,1$	-	-	-	-	-	-	0	0	0	0
$1,1,2,2,2,1,1$	-	-	-	-	-	-	0	0	0	0
$1,-1,2,-2,2,-1,1$	-	-	-	-	-	-	0	0	0	0
$1,1,1,2,2,2,2,2,1,1,1$	-	-	-	-	-	-	0	0	0	0
$1,-1,1,-2,2,-2,2,-2,2,-1,1,-1$	-	-	-	-	-	-	0	0	0	0
$1,2,2,1,0,-1,-2,-2,-1$	-	-	-	-	-	-	0	0	-	-
$1,-2,2,-1,0,1,-2,2,-1$	-	-	-	-	-	-	0	0	-	-
$1,1,-1,-1,1,1,-1,-1$	-	-	-	-	-	-	0	-	-	-
$1,-1,-1,1,1,-1,-1,1$	-	-	-	-	-	-	0	-	-	-
$1,3,5,5,3,1$	-	-	-	-	-	-	0	-	-	-
$1,-3,5,-5,3,-1$	-	-	-	-	-	-	0	-	-	-
$1,0,1,0,1,0,1$	-	-	-	-	-	-	-	0.1	0.1	0
$1,0,-1,0,1,0,-1$	-	-	-	-	-	-	-	0.1	0.1	0
$1,1,1,1,1,1,1,1$	-	-	-	-	-	-	-	0	0	0
$1,-1,1,-1,1,-1,1,-1$	-	-	-	-	-	-	-	0	0	0
$1,0,0,0,1$	-	-	-	-	-	-	-	0	0	0
$1,0,0,0,-1$	-	-	-	-	-	-	-	0	0	0
$1,1,0,-1,-1,0,1,1$	-	-	-	-	-	-	-	0	0.1	0
$1,-1,0,1,-1,0,1,-1$	-	-	-	-	-	-	-	0	0.1	0
$1,1,0,-1,-1,-1,0,1,1$	-	-	-	-	-	-	-	0	0	0
$1,-1,0,1,-1,1,0,-1,1$	-	-	-	-	-	-	-	0	0	0
$1,1,-1,0,2,0,-1,1,1$	-	-	-	-	-	-	-	0	-	-
$1,-1,-1,0,2,0,-1,-1,1$	-	-	-	-	-	-	-	0	-	-
$1,2,3,3,3,3,3,3,2,1$	-	-	-	-	-	-	-	0	-	-
$1,-2,3,-3,3,-3,3,-3,2,-1$	-	-	-	-	-	-	-	0	-	-
$1,1,1,0,-1,-2,-2,-1,0,1,1,1$	-	-	-	-	-	-	-	0	-	0
$1,-1,1,0,-1,2,-2,1,0,-1,1,-1$	-	-	-	-	-	-	-	0	-	0
$1,1,2,2,3,2,2,1,1$	-	-	-	-	-	-	-	0	-	-
$1,-1,2,-2,3,-2,2,-1,1$	-	-	-	-	-	-	-	0	-	-
$1,1,2,2,2,2,2,1,1$	-	-	-	-	-	-	-	0	0	0
$1,-1,2,-2,2,-2,2,-1,1$	-	-	-	-	-	-	-	0	0	0
$1,1,1,2,2,2,2,1,1,1$	-	-	-	-	-	-	-	0	-	-
$1,-1,1,-2,2,-2,2,-1,1,-1$	-	-	-	-	-	-	-	0	-	-
$1,0,-2,-1,2,2,-1,-2,0,1$	-	-	-	-	-	-	-	0	-	-
$1,0,-2,1,2,-2,-1,2,0,-1$	-	-	-	-	-	-	-	0	-	-
$1,1,0,-1,0,1,0,-1,-1$	-	-	-	-	-	-	-	0	0	-
$1,-1,0,1,0,-1,0,1,-1$	-	-	-	-	-	-	-	0	0	-
$1,0,0,1,0,0,1$	-	-	-	-	-	-	-	-	0	0
$1,0,0,-1,0,0,1$	-	-	-	-	-	-	-	-	0	0
$1,0,-1,-1,1,1,0,-1$	-	-	-	-	-	-	-	-	0	-
$1,0,-1,1,1,-1,0,1$	-	-	-	-	-	-	-	-	0	-

L	2	3	4	5	6	7	8	9	10	11
$1,1,1,2,1,1,1$	-	-	-	-	-	-	-	-	0	0
$1,-1,1,-2,1,-1,1$	-	-	-	-	-	-	-	-	0	0
$1,1,0,0,0,-1,-1$	-	-	-	-	-	-	-	-	0	0
$1,-1,0,0,0,1,-1$	-	-	-	-	-	-	-	-	0	0
$1,1,1,1,2,1,1,1,1$	-	-	-	-	-	-	-	-	0	0
$1,-1,1,-1,2,-1,1,-1,1$	-	-	-	-	-	-	-	-	0	0
$1,1,1,0,0,-1,-1,-1$	-	-	-	-	-	-	-	-	0	-
$1,-1,1,0,0,1,-1,1$	-	-	-	-	-	-	-	-	0	-
$1,1,1,1,1,1,1,1,1$	-	-	-	-	-	-	-	-	0	0
$1,-1,1,-1,1,-1,1,-1,1$	-	-	-	-	-	-	-	-	0	0
$1,1,1,1,0,-1,-1,-1,-1$	-	-	-	-	-	-	-	-	0	0
$1,-1,1,-1,0,1,-1,1,-1$	-	-	-	-	-	-	-	-	0	0
$1,1,0,-1,-1,-1,-1,0,1,1$	-	-	-	-	-	-	-	-	0	-
$1,-1,0,1,-1,1,-1,0,1,-1$	-	-	-	-	-	-	-	-	0	-
$1,1,1,0,-1,-1,0,1,1,1$	-	-	-	-	-	-	-	-	0	-
$1,-1,1,0,-1,1,0,-1,1,-1$	-	-	-	-	-	-	-	-	0	-
$1,1,-1,-1$	-	-	-	-	-	-	-	-	-	0
$1,-1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,-1,0,1,0,-1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,1,0,1,2,1,0,1,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,0,-1,2,-1,0,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,0,1,0,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,0,0,0,-1,0,-1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,0,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,0,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,1,1,1,1,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,1,1,0,-1,-1,-1,0,1,1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,1,-1,1,-1,1,-1,1,0,1$	-	-	-	-	-	-	-	-	-	0
$1,-1,1,0,-1,1,-1,0,1,-1,1$	-	-	-	-	-	-	-	-	-	0
$1,0,-1,-1,0,1,1,0,-1$	-	-	-	-	-	-	-	-	-	0
$1,0,-1,1,0,-1,1,0,-1$	-	-	-	-	-	-	-	-	-	0
$1,3,5,6,6,6,6,5,3,1$	-	-	-	-	0	-	-	-	-	-
$1,-3,5,-6,6,-6,6,-5,3,-1$	-	-	-	-	0	-	-	-	-	-
$1,4,7,6,0,-6,-6,0,6,7,4,1$	-	-	-	-	-	0	0	-	-	-
$1,-4,7,-6,0,6,-6,0,6,-7,4,-1$	-	-	-	-	-	0	0	-	-	-
$1,4,7,6,1,-2,1,6,7,4,1$	-	-	-	-	-	-	-	0	0	0
$1,-4,7,-6,1,2,1,-6,7,-4,1$	-	-	-	-	-	-	-	0	0	0

A.4 QPSK

L	2	3	4	5	6
1	100	93.9	94.8	92.8	100
$1,-1$	-	1.3	1	1.3	0
$1,1$	-	1.3	1	1.3	0
$1,-j$	-	1.2	0.9	1	-
$1,j$	-	1.2	0.9	1	-
$1,-1-j,j$	-	0.3	0.2	0.3	-
$1,1-j,-j$	-	0.3	0.2	0.3	-
$1,1+j,j$	-	0.3	0.2	0.3	-
$1,-1+j,-j$	-	0.3	0.2	0.3	-
$1,-1,1$	-	-	0.2	0.2	0
$1,1,1$	-	-	0.2	0.2	0
$1,-j,-1$	-	-	0.2	0.2	-
$1,j,-1$	-	-	0.2	0.2	-
$1,0,1$	-	-	-	0	0
$1,0,-1$	-	-	-	0	0
$1,-1,1,-1$	-	-	-	0.1	0
$1,1,1,1$	-	-	-	0.1	0
$1,-j,-1,j$	-	-	-	0.1	-
$1,j,-1,-j$	-	-	-	0.1	-
$1,0,-j$	-	-	-	0	-
$1,0,j$	-	-	-	0	-
$1,-1-j,j,1-j,-1$	-	-	-	0	-
$1,1-j,-j,-1-j,-1$	-	-	-	0	-
$1,1+j,j,-1+j,-1$	-	-	-	0	-
$1,-1+j,-j,1+j,-1$	-	-	-	0	-
$1,-1,1,-1,1$	-	-	-	-	0
$1,1,1,1,1$	-	-	-	-	0
$1,-1,0,1,-1$	-	-	-	-	0

L	2	3	4	5	6
$1, 1, 0, -1, -1$	-	-	-	-	0
$1, -1-j, 0, -1+j, 1$	-	-	-	0	0
$1, 1-j, 0, 1+j, 1$	-	-	-	0	0
$1, 1+j, 0, 1-j, 1$	-	-	-	0	0
$1, -1+j, 0, -1-j, 1$	-	-	-	0	0

A.5 8-PSK

The nomenclature of [87, p30] has been used for the error events in the case of 8-PSK. This is represented in Fig. A.1. q is the factor required to rotate by 1/8 of a revolution which is $(1+j)/\sqrt{2}$. The question marks represent sequences which for some channels have distance virtually indistinguishable from that of some other sequence for the same channel, and so cannot with confidence be excluded.

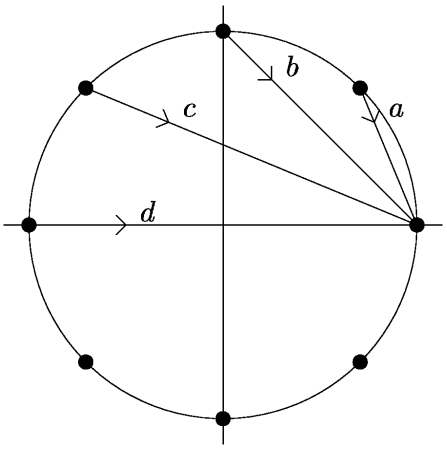


Figure A.1: Labelling of difference symbols for 8-PSK

L	2	3	4	5	6
a	100	91.5	92.5	89.7	82.8
a, q^4a	-	1.1	0.9	1.1	4.8
a, a	-	1.1	0.9	1.1	4.7
a, q^5a	-	1.1	0.8	1	-
a, qa	-	1.1	0.8	1	-
a, q^7a	-	1.1	0.8	1	-
a, q^3a	-	1.1	0.8	1	-
a, q^6a	-	1	0.7	0.9	-
a, q^2a	-	1	0.7	0.9	-
$a, q^7b, q^6c, q^6b, q^6a$	-	0	0	-	-
$a, q^6b, q^4c, q^2d, qc, b, q^7a$	-	0	-	-	-
$a, q^2b, q^4c, q^6d, qc, q^4b, q^7a$	-	0	-	-	-
$a, q^4b, c, q^4d, qc, q^6b, q^3a$	-	0	0	?	0
$a, b, c, d, qc, q^2b, q^3a$	-	0	0	?	0
$a, q^3b, q^6c, qd, q^5c, qb, q^5a$	-	0	0	0	0
a, q^4b, c, q^5b, q^2a	-	0	0	-	-
a, b, c, qb, q^2a	-	0	0	-	-
$a, q^3b, q^6c, q^2b, q^6a$	-	0	0	-	-
$a, q^6b, q^4c, q^3b, q^2a$	-	0	0	-	-
$a, q^2b, q^4c, q^7b, q^2a$	-	0	0	-	-
a, a, a	-	-	0.2	0.2	1.8
a, q^4a, a	-	-	0.2	0.2	2
a, q^7c, q^7c, a	-	-	0	0	0
$a, q^7b, q^6c, q^5d, q^5c, q^5b, q^5a$	-	-	0	0	0
a, q^7a, q^6a	-	-	0.2	0.2	-
a, q^3a, q^6a	-	-	0.2	0.2	-

L	2	3	4	5	6
a, q^5a, q^2a	-	-	0.2	0.2	-
a, qa, q^2a	-	-	0.2	0.2	-
a, q^6a, q^4a	-	-	0.1	0.2	-
a, q^2a, q^4a	-	-	0.1	0.2	-
a, q^7a, q^5a	-	-	0	0	-
a, q^2a, q^5a	-	-	0	0	-
a, q^6a, q^3a	-	-	0	0	-
a, qa, q^3a	-	-	0	0	-
a, q^5a, qa	-	-	0	0	-
a, q^7a, q^7a	-	-	0	0	-
a, qa, qa	-	-	0	0	-
a, q^3a, q^7a	-	-	0	0	-
$a, q^5b, q^2c, q^7d, q^5c, q^3b, qa$	-	-	0	-	-
$a, qb, q^2c, q^3d, q^5c, q^7b, qa$	-	-	0	-	-
a, q^4b, qa	-	-	0	-	-
a, b, qa	-	-	0	-	-
a, q^7b, q^7a	-	-	0	-	-
a, q^3b, q^7a	-	-	0	-	-
a, q^6b, q^5a	-	-	0	-	-
a, q^2b, q^5a	-	-	0	-	-
a, q^5b, q^2c, b, q^6a	-	-	0	-	-
a, qb, q^2c, q^4b, q^6a	-	-	0	-	-
a, q^5b, q^3b, qb, q^7a	-	-	0	-	-
a, qb, q^3b, q^5b, q^7a	-	-	0	-	-
$a, q^5b, q^2b, q^7b, q^5a$	-	-	0	-	-
a, qb, q^2b, q^3b, q^5a	-	-	0	-	-
$a, 0, a$	-	-	-	0	1.3
$a, 0, q^4a$	-	-	-	0	0.9
a, q^4a, a, q^4a	-	-	-	0.1	-
a, a, a, a	-	-	-	0.1	0.7
a, q^3c, q^7c, q^4a	-	-	-	0	0
$a, q^7a, q^5a, q^3a, q^2a$	-	-	-	0	-
$a, q^3a, q^5a, q^7a, q^2a$	-	-	-	0	-
a, q^2a, q^4a, q^6a	-	-	-	0.1	-
a, q^6a, q^4a, q^2a	-	-	-	0.1	-
a, q^5a, q^2a, q^7a	-	-	-	0	-
a, q^7a, q^6a, q^5a	-	-	-	0	-
a, qa, q^2a, q^3a	-	-	-	0	-
a, q^3a, q^6a, qa	-	-	-	0	-
$a, 0, q^6a$	-	-	-	0	-
$a, 0, q^2a$	-	-	-	0	-
$a, 0, q^5a$	-	-	-	0	-
$a, 0, q^7a$	-	-	-	0	-
$a, 0, qa$	-	-	-	0	-
$a, 0, q^3a$	-	-	-	0	-
a, q^4a, qa, q^5a	-	-	-	0	-
a, a, q^7a, q^7a	-	-	-	0	-
a, a, qa, qa	-	-	-	0	-
a, q^4a, q^7a, q^3a	-	-	-	0	-
a, q^7a, q^7a, q^6a	-	-	-	0	-
a, qa, qa, q^2a	-	-	-	0	-
a, q^3a, q^7a, q^2a	-	-	-	0	-
a, q^6a, q^3a, qa	-	-	-	0	-
a, q^6a, q^5a, q^3a	-	-	-	0	-
a, q^2a, q^3a, q^5a	-	-	-	0	-
a, q^2a, q^5a, q^7a	-	-	-	0	-
a, q^7a, q^5a, q^4a	-	-	-	0	-
a, qa, q^3a, q^4a	-	-	-	0	-
a, q^3a, q^5a, a	-	-	-	0	-
a, q^5a, q^3a, a	-	-	-	0	-
a, q^5a, qa, q^5a, q^2a	-	-	-	0	-
$a, q^7a, q^7a, q^7a, q^6a$	-	-	-	0	-
$a, q^3a, q^7a, q^3a, q^6a$	-	-	-	0	-
a, qa, qa, qa, q^2a	-	-	-	0	-
a, q^5a, q^3a, qa, q^6a	-	-	-	0	-
a, qa, q^3a, q^5a, q^6a	-	-	-	0	-
a, q^6a, q^3a, a, q^6a	-	-	-	0	-
$a, q^2a, q^3a, q^4a, q^6a$	-	-	-	0	-
$a, q^6a, q^5a, q^4a, q^2a$	-	-	-	0	-
a, q^2a, q^5a, a, q^2a	-	-	-	0	-
a, a, q^7a, q^6a, q^6a	-	-	-	0	-
$a, q^4a, q^7a, q^2a, q^6a$	-	-	-	0	-
a, q^4a, qa, q^6a, q^2a	-	-	-	0	-
a, a, qa, q^2a, q^2a	-	-	-	0	-
$a, qa, q^3a, q^4b, q^5b, q^7a, qa, q^2a$	-	-	-	0	-

L	2	3	4	5	6
$a, q^7a, q^5a, q^3b, q^2b, qa, q^7a, q^6a$	-	-	-	0	-
$a, q^3a, q^5a, q^7b, q^2b, q^5a, q^7a, q^2a$	-	-	-	0	-
a, q^4a, a, q^4a, a	-	-	-	-	0.5
a, a, a, a, a	-	-	-	-	0.3
$a, q^4a, q^7a, q^3a, q^6a, q^2a$	-	-	-	-	0
$a, q^7c, q^7c, a, q^4a, q^3c, q^3c, q^4a$	-	-	-	-	0
$a, q^3c, q^7c, q^4a, q^4a, q^7c, q^3c, a$	-	-	-	-	0
a, q^4b, b, q^4b, qa	-	-	-	-	0
$a, q^7a, q^5a, q^4a, q^2a, qa$	-	-	-	-	0
$a, q^3a, q^5a, a, q^2a, q^5a$	-	-	-	-	0
a, q^4c, qc, q^7a	-	-	-	0	0
a, c, qc, q^3a	-	-	-	0	0
a, qc, q^3c, q^6a	-	-	-	0	0
a, q^5c, q^3c, q^2a	-	-	-	0	0
$a, q^3b, q^7b, q^3b, q^7a$	-	-	-	-	?
$a, q^4a, qa, q^5a, q^2a, q^6a$	-	-	-	-	?
$a, q^5a, q^3a, a, q^6a, q^3a$	-	-	-	-	?
$a, qa, q^3a, q^4a, q^6a, q^7a$	-	-	-	-	?
a, q^6c, q^5c, q^5a	-	-	-	-	0
a, q^2c, q^5c, qa	-	-	-	-	0

A.6 16-QAM

L	2	3	4	5	6
1	100	93.9	94.8	92.8	100
$1, -1$	-	1.3	1	1.3	0
$1, 1$	-	1.3	1	1.3	0
$1, -j$	-	1.2	0.9	1	-
$1, j$	-	1.2	0.9	1	-
$1, -1-j, j$	-	0.3	0.2	0.3	-
$1, 1-j, -j$	-	0.3	0.2	0.3	-
$1, 1+j, j$	-	0.3	0.2	0.3	-
$1, -1+j, -j$	-	0.3	0.2	0.3	-
$1, -2, 2, -2, 2, -2, 2, -2, 1$	-	0	-	-	-
$1, -2j, -2, 2j, 2, -2j, -2, 2j, 1$	-	0	-	-	-
$1, 2, 2, 2, 2, 2, 2, 2, 1$	-	0	-	-	-
$1, 2j, -2, -2j, 2, 2j, -2, -2j, 1$	-	0	-	-	-
$1, -1, 1$	-	-	0.2	0.2	0
$1, 1, 1$	-	-	0.2	0.2	0
$1, -j, -1$	-	-	0.2	0.2	-
$1, j, -1$	-	-	0.2	0.2	-
$1, -1-j, 2j, 1-j, -1$	-	-	0	0	-
$1, 1-j, -2j, -1-j, -1$	-	-	0	0	-
$1, 1+j, 2j, -1+j, -1$	-	-	0	0	-
$1, -1+j, -2j, 1+j, -1$	-	-	0	0	-
$1, -2, 2, -1$	-	-	0	0	0
$1, 2, 2, 1$	-	-	0	0	0
$1, -2j, -2, j$	-	-	0	-	-
$1, 2j, -2, -j$	-	-	0	-	-
$1, -1-j, -1+j, 2+j, -2j, -2, 1+2j, 1-j, -1-j, j$	-	-	0	-	-
$1, 1-j, -1-j, -2+j, 2j, 2, 1-2j, -1-j, -1+j, j$	-	-	0	-	-
$1, 1+j, -1+j, -2-j, -2j, 2, 1+2j, -1+j, -1-j, -j$	-	-	0	-	-
$1, -1+j, -1-j, 2-j, 2j, -2, 1-2j, 1+j, -1+j, -j$	-	-	0	-	-
$1, -1-j, -1+j, 1+j, 1-j, -1-j, -1+j, 1+j, -j$	-	-	0	-	-
$1, 1-j, -1-j, -1+j, 1+j, 1-j, -1-j, -1+j, j$	-	-	0	-	-
$1, 1+j, -1+j, -1-j, 1-j, 1+j, -1+j, -1-j, -j$	-	-	0	-	-
$1, -1+j, -1-j, 1-j, 1+j, -1+j, -1-j, 1-j, j$	-	-	0	-	-
$1, -2, 3, -3, 2, -1$	-	-	0	0	0
$1, 2, 3, 3, 2, 1$	-	-	0	0	0
$1, -2, 3, -3, 3, -3, 3, -2, 1$	-	-	0	-	-
$1, 2, 3, 3, 3, 3, 3, 2, 1$	-	-	0	-	-
$1, -j, -1, j$	-	-	-	0.1	-
$1, j, -1, -j$	-	-	-	0.1	-
$1, -1, 1, -1$	-	-	-	0.1	0
$1, 1, 1, 1$	-	-	-	0.1	0
$1, 0, -j$	-	-	-	0	-
$1, 0, j$	-	-	-	0	-
$1, 0, 1$	-	-	-	0	0
$1, 0, -1$	-	-	-	0	0
$1, -1-j, j, 1-j, -1$	-	-	-	0	-
$1, 1-j, -j, -1-j, -1$	-	-	-	0	-

L	2	3	4	5	6
$1, 1+j, j, -1+j, -1$	-	-	-	0	-
$1, -1+j, -j, 1+j, -1$	-	-	-	0	-
$1, -1-j, 2j, 1-2j, -2+j, 2, -1-j, j$	-	-	-	0	-
$1, -1-j, 2j, 2-j, -2-j, 2j, 1-j, -1$	-	-	-	0	-
$1, 1-j, -2j, -2-j, -2+j, 2j, 1+j, 1$	-	-	-	0	-
$1, 1-j, -2j, -1-2j, -2-j, -2, -1+j, j$	-	-	-	0	-
$1, 1+j, 2j, -1+2j, -2+j, -2, -1-j, -j$	-	-	-	0	-
$1, 1+j, 2j, -2+j, -2-j, -2j, 1-j, 1$	-	-	-	0	-
$1, -1+j, -2j, 2+j, -2+j, -2j, 1+j, -1$	-	-	-	0	-
$1, -1+j, -2j, 1+2j, -2-j, 2, -1+j, -j$	-	-	-	0	-
$1, -1-j, 1+j, -1-2j, 2j, -2j, -1+2j, 1-j, -1+j, 1$	-	-	-	0	-
$1, 1-j, 1-j, 1-2j, -2j, -2j, -1-2j, -1-j, -1-j, -1$	-	-	-	0	-
$1, 1+j, 1+j, 1+2j, 2j, 2j, -1+2j, -1+j, -1+j, -1$	-	-	-	0	-
$1, -1+j, 1-j, -1+2j, -2j, 2j, -1-2j, 1+j, -1-j, 1$	-	-	-	0	-
$1, -1-j, 2j, 2-2j, -3, 2+2j, -2j, -1+j, 1$	-	-	-	0	-
$1, 1-j, -2j, -2-2j, -3, -2+2j, 2j, 1+j, 1$	-	-	-	0	-
$1, 1+j, 2j, -2+2j, -3, -2-2j, -2j, 1-j, 1$	-	-	-	0	-
$1, -1+j, -2j, 2+2j, -3, 2-2j, 2j, -1-j, 1$	-	-	-	0	-
$1, -1-j, j, 1, -1-j, j, 1, -1-j, j$	-	-	-	0	-
$1, 1-j, -j, -1, -1+j, j, 1, 1-j, -j$	-	-	-	0	-
$1, 1+j, j, -1, -1-j, -j, 1, 1+j, j$	-	-	-	0	-
$1, -1+j, -j, 1, -1+j, -j, 1, -1+j, -j$	-	-	-	0	-
$1, -j, -1, 1+j, 1-j, -1-j, j, 1, -j$	-	-	-	0	-
$1, -j, -1, -1+j, 1+j, 1-j, -j, -1, j$	-	-	-	0	-
$1, j, -1, -1-j, 1-j, 1+j, j, -1, -j$	-	-	-	0	-
$1, j, -1, 1-j, 1+j, -1+j, -j, 1, j$	-	-	-	0	-
$1, -1, -1, 2, 0, -2, 1, 1, -1$	-	-	-	0	0
$1, 1, -1, -2, 0, 2, 1, -1, -1$	-	-	-	0	0
$1, -2, 3, -3, 3, -2, 1$	-	-	-	0	0
$1, 2, 3, 3, 3, 2, 1$	-	-	-	0	0
$1, -2, 1, 1, -2, 1$	-	-	-	0	0
$1, 2, 1, -1, -2, -1$	-	-	-	0	0
$1, -2, 2, -2, 2, -1$	-	-	-	0	0
$1, 2, 2, 2, 2, 1$	-	-	-	0	0
$1, -2, 1$	-	-	-	0	0
$1, 2, 1$	-	-	-	0	0
$1, -1, 1, -1, 1$	-	-	-	-	0
$1, 1, 1, 1, 1$	-	-	-	-	0
$1, -1, 0, 1, -1$	-	-	-	-	0
$1, 1, 0, -1, -1$	-	-	-	-	0
$1, -1, 2, -2, 3, -2, 2, -1, 1$	-	-	-	-	0
$1, 1, 2, 2, 3, 2, 2, 1, 1$	-	-	-	-	0
$1, 0, 1, -1, 0, -1, 1, 0, 1$	-	-	-	-	0
$1, 0, 1, 1, 0, 1, 1, 0, 1$	-	-	-	-	0
$1, -2, 2, -2, 1$	-	-	-	-	0
$1, 2, 2, 2, 1$	-	-	-	-	0
$1, -1, 2, -1, 1, 1, -1, 2, -1, 1$	-	-	-	-	0
$1, 1, 2, 1, 1, -1, -1, -2, -1, -1$	-	-	-	-	0
$1, 0, 2, 0, 2, 0, 2, 0, 1$	-	-	-	-	0
$1, -1, -1, 1, 1, -1, -1, 1, 1, -1$	-	-	-	-	0
$1, 1, -1, -1, 1, 1, -1, -1, 1, 1$	-	-	-	-	0
$1, 0, -2, 0, 2, 0, -2, 0, 1$	-	-	-	-	0
$1, 0, 1, -1, 1, -1, 1, 0, 1$	-	-	-	-	0
$1, 0, 1, 1, 1, 1, 1, 0, 1$	-	-	-	-	0
$1, -2, 1, 1, -1, -1, 2, -1$	-	-	-	-	0
$1, 2, 1, -1, -1, 1, 2, 1$	-	-	-	-	0
$1, -1, 1, -2, 2, -2, 2, -1, 1, -1$	-	-	-	-	0
$1, 1, 1, 2, 2, 2, 2, 1, 1, 1$	-	-	-	-	0
$1, -1, -1, 1, 2, -2, -1, 1, 1, -1$	-	-	-	-	0
$1, 1, -1, -1, 2, 2, -1, -1, 1, 1$	-	-	-	-	0
$1, -1-2j, -1+2j, 1-j, -1+j, 1-j, -2+j, 2+j, -j$	-	-	-	0	-
$1, -2-j, 1+2j, 1-j, -1-j, -1+j, 2+j, -1-2j, j$	-	-	-	0	-
$1, 2-j, 1-2j, -1-j, -1+j, 1+j, 2-j, 1-2j, -j$	-	-	-	0	-
$1, 1-2j, -1-2j, -1-j, -1-j, -1-j, -2-j, -2+j, j$	-	-	-	0	-
$1, 1+2j, -1+2j, -1+j, -1+j, -1+j, -2+j, -2-j, -j$	-	-	-	0	-
$1, 2+j, 1+2j, -1+j, -1-j, 1-j, 2+j, 1+2j, j$	-	-	-	0	-
$1, -2+j, 1-2j, 1+j, -1+j, -1-j, 2-j, -1+2j, -j$	-	-	-	0	-
$1, -1+2j, -1-2j, 1+j, -1-j, 1+j, -2-j, 2-j, j$	-	-	-	0	-
$1, -1-3j, -3+3j, 3, -2, 2, -3, 3+3j, 1-3j, -1$	-	-	-	-	0
$1, -3-j, 3+3j, -3j, -2, 2j, 3, -3-3j, 1+3j, -j$	-	-	-	-	0
$1, 3-j, 3-3j, -3j, -2, 2j, 3, 3-3j, 1-3j, -j$	-	-	-	-	0
$1, 1-3j, -3-3j, -3, -2, -2, -3, -3+3j, 1+3j, 1$	-	-	-	-	0
$1, 1+3j, -3+3j, -3, -2, -2, -3, -3-3j, 1-3j, 1$	-	-	-	-	0
$1, 3+j, 3+3j, 3j, -2, -2j, 3, 3+3j, 1+3j, j$	-	-	-	-	0
$1, -3+j, 3-3j, 3j, -2, -2j, 3, -3+3j, 1-3j, j$	-	-	-	-	0

L	2	3	4	5	6
$1, -1+3j, -3-3j, 3, -2, 2, -3, 3-3j, 1+3j, -1$	-	-	-	-	0

A.7 16-PSK

Again the nomenclature of [87, p30] has been used for the error events in the case of 16-PSK. This is represented in Fig. A.2. q is the factor required to rotate by 1/16 of a revolution which is $(\sqrt{2+\sqrt{2}}+j\sqrt{2-\sqrt{2}})/2$.

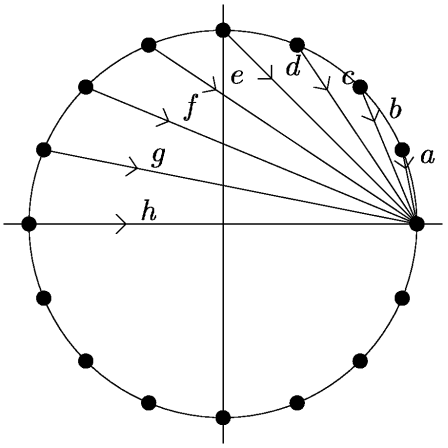


Figure A.2: Labelling of difference symbols for 16-PSK

L	2	3	4	5	6
a	100	90.3	91.4	87.5	82.6
a, a	-	0.6	0.5	0.6	4.6
a, q^8a	-	0.6	0.5	0.6	4.6
a, qa	-	0.6	0.5	0.7	-
$a, q^{15}a$	-	0.6	0.5	0.7	-
a, q^9a	-	0.6	0.5	0.7	-
a, q^7a	-	0.6	0.5	0.7	-
a, q^2a	-	0.6	0.4	0.6	-
$a, q^{10}a$	-	0.6	0.4	0.6	-
$a, q^{14}a$	-	0.6	0.4	0.6	-
a, q^6a	-	0.6	0.4	0.6	-
$a, q^{13}a$	-	0.6	0.4	0.5	-
a, q^5a	-	0.6	0.4	0.5	-
a, q^3a	-	0.6	0.4	0.5	-
$a, q^{11}a$	-	0.6	0.4	0.5	-
$a, q^{12}a$	-	0.6	0.4	0.6	-
a, q^4a	-	0.6	0.4	0.6	-
$a, q^3b, q^7b, q^{11}a$	-	0	0	-	-
$a, q^{11}b, q^7b, q^3a$	-	0	0	-	-
$a, q^9b, q^2b, q^{12}a$	-	0	0	-	-
$a, q^{13}b, q^{11}b, q^9a$	-	0	0	-	-
$a, q^5b, q^{11}b, qa$	-	0	0	-	-
$a, q^7b, q^{14}b, q^6a$	-	0	0	0	0
$a, q^{14}b, q^{12}b, q^{11}a$	-	0	0	-	-
$a, q^{10}b, q^4b, q^{15}a$	-	0	0	-	-
a, q^2b, q^4b, q^7a	-	0	0	-	-
a, q^2b, q^5b, q^8a	-	0	0	-	-
$a, q^{11}b, q^6b, q^2a$	-	0	0	-	-
$a, q^6b, q^{12}b, q^3a$	-	0	0	-	-
$a, q^{10}b, q^5b, a$	-	0	0	-	-
$a, q^6b, q^{13}b, q^4a$	-	0	0	-	-
a, q^8a, a	-	-	0.1	0.1	1.8
a, a, a	-	-	0.1	0.1	1.8
$a, q^{15}b, q^{15}b, q^{15}a$	-	-	0	0	0.1

L	2	3	4	5	6
$a, q^8b, c, q^8c, qb, q^{10}a$	-	-	0	0	-
$a, q^7c, q^{14}e, q^6e, q^{15}c, q^8a$	-	-	0	0	0
$a, q^{15}c, q^{14}e, q^{14}e, q^{15}c, a$	-	-	0	-	-
$a, q^7b, q^{15}b, q^8a, q^4a, q^{12}b, q^4b, q^{12}a$	-	-	0	-	-
$a, q^{14}a, q^{12}a$	-	-	0.1	0.1	-
$a, q^6a, q^{12}a$	-	-	0.1	0.1	-
$a, q^{10}a, q^4a$	-	-	0.1	0.1	-
a, q^2a, q^4a	-	-	0.1	0.1	-
a, q^9a, q^2a	-	-	0.1	0.1	-
a, qa, q^2a	-	-	0.1	0.1	-
$a, q^{15}a, q^{14}a$	-	-	0.1	0.1	-
$a, q^7a, q^{14}a$	-	-	0.1	0.1	-
$a, q^{13}a, q^{10}a$	-	-	0.1	0	-
$a, q^5a, q^{10}a$	-	-	0.1	0	-
$a, q^{11}a, q^6a$	-	-	0.1	0	-
a, q^3a, q^6a	-	-	0.1	0	-
$a, q^{12}a, q^8a$	-	-	0.1	0.2	-
a, q^4a, q^8a	-	-	0.1	0.2	-
$a, q^{14}a, q^{13}a, q^{11}a$	-	-	0	0	-
$a, q^6a, q^{13}a, q^3a$	-	-	0	0	-
$a, q^{10}a, q^3a, q^{13}a$	-	-	0	0	-
a, q^2a, q^3a, q^5a	-	-	0	0	-
$a, q^{12}a, q^7a, q^3a$	-	-	0	0	-
$a, q^{12}a, q^9a, q^5a$	-	-	0	0	-
$a, q^4a, q^9a, q^{13}a$	-	-	0	0	-
$a, q^4a, q^7a, q^{11}a$	-	-	0	0	-
$a, q^{13}a, q^{11}a, q^8a$	-	-	0	0	-
$a, q^5a, q^{11}a, a$	-	-	0	0	-
a, q^3a, q^5a, q^8a	-	-	0	0	-
$a, q^{11}a, q^5a, a$	-	-	0	0	-
$a, q^9a, qa, q^{10}a$	-	-	0	0	-
a, qa, qa, q^2a	-	-	0	0	-
$a, q^{15}a, q^{15}a, q^{14}a$	-	-	0	0	-
$a, q^7a, q^{15}a, q^6a$	-	-	0	0	-
$a, a, q^{15}a, q^{15}a$	-	-	0	0	-
$a, q^8a, q^{15}a, q^7a$	-	-	0	0	-
a, q^8a, qa, q^9a	-	-	0	0	-
a, a, qa, qa	-	-	0	0	-
$a, q^{13}a, q^9a, q^6a$	-	-	0	0	-
$a, q^5a, q^9a, q^{14}a$	-	-	0	0	-
$a, q^{11}a, q^7a, q^2a$	-	-	0	0	-
$a, q^3a, q^7a, q^{10}a$	-	-	0	0	-
$a, q^{14}a, q^{11}a, q^9a$	-	-	0	0	-
$a, q^6a, q^{11}a, qa$	-	-	0	0	-
$a, q^{10}a, q^5a, q^{15}a$	-	-	0	0	-
a, q^2a, q^5a, q^7a	-	-	0	0	-
$a, q^9a, q^3a, q^{12}a$	-	-	0	0	-
a, qa, q^3a, q^4a	-	-	0	0	-
$a, q^{15}a, q^{13}a, q^{12}a$	-	-	0	0	-
$a, q^7a, q^{13}a, q^4a$	-	-	0	0	-
$a, q^{10}a, q^3a$	-	-	0	0	-
a, q^2a, q^3a	-	-	0	0	-
$a, q^{15}a, q^{13}a$	-	-	0	0	-
$a, q^7a, q^{13}a$	-	-	0	0	-
a, q^9a, qa	-	-	0	0	-
a, qa, qa	-	-	0	0	-
$a, q^{15}a, q^{15}a$	-	-	0	0	-
$a, q^7a, q^{15}a$	-	-	0	0	-
$a, q^{13}a, q^{11}a$	-	-	0	0	-
$a, q^5a, q^{11}a$	-	-	0	0	-
$a, q^{11}a, q^5a$	-	-	0	0	-
a, q^3a, q^5a	-	-	0	0	-
$a, q^{13}a, q^9a$	-	-	0	0	-
a, q^5a, q^9a	-	-	0	0	-
$a, q^{11}a, q^7a$	-	-	0	0	-
a, q^3a, q^7a	-	-	0	0	-
$a, q^8b, qb, q^{10}a$	-	-	0	-	-
a, b, qb, q^2a	-	-	0	-	-
a, qb, q^2b, q^4a	-	-	0	-	-
$a, q^{12}b, q^8b, q^5a$	-	-	0	-	-
$a, q^4b, q^8b, q^{13}a$	-	-	0	-	-
$a, q^{12}b, q^9b, q^6a$	-	-	0	-	-
$a, q^{13}b, q^{10}b, q^8a$	-	-	0	-	-
$a, q^5b, q^{10}b, a$	-	-	0	-	-
$a, q^{14}b, q^{13}b, q^{12}a$	-	-	0	-	-

L	2	3	4	5	6
$a, q^{14}b, q^{13}b, q^{12}b, q^{11}a$	-	-	0	-	-
$a, q^{6}b, q^{13}b, q^{4}b, q^{11}a$	-	-	0	-	-
$a, q^{10}b, q^{5}b, b, q^{11}a$	-	-	0	-	-
$a, q^{2}b, q^{5}b, q^{8}b, q^{11}a$	-	-	0	-	-
$a, q^{15}b, q^{14}b, q^{14}a$	-	-	0	0	0
$a, 0, a$	-	-	-	0	1.3
$a, 0, q^{8}a$	-	-	-	0	0.9
a, a, a, a	-	-	-	0	0.7
$a, q^{8}a, a, q^{8}a$	-	-	-	0	0.7
$a, b, c, c, qb, q^{2}a$	-	-	-	0	0
$a, q^{7}b, q^{15}b, q^{8}a, q^{5}a, q^{13}b, q^{5}b, q^{13}a$	-	-	-	0	-
$a, q^{15}b, q^{15}a, q^{7}a, q^{6}b, q^{6}a$	-	-	-	0	0
$a, q^{7}b, q^{15}a, q^{15}a, q^{6}b, q^{14}a$	-	-	-	0	-
$a, q^{14}a, q^{10}a, q^{6}a, q^{4}a$	-	-	-	0	-
$a, q^{6}a, q^{10}a, q^{14}a, q^{4}a$	-	-	-	0	-
$a, q^{7}b, q^{15}b, q^{7}b, q^{15}b, q^{7}a$	-	-	-	0	0
$a, q^{9}a, q^{6}a, q^{15}b, q^{9}a, q^{6}a, q^{15}a$	-	-	-	0	-
$a, qa, q^{6}a, q^{7}b, q^{9}a, q^{14}a, q^{15}a$	-	-	-	0	-
$a, q^{8}b, qb, q^{11}a, q^{9}a, q^{2}b, q^{11}b, q^{4}a$	-	-	-	0	0
$a, q^{7}b, q^{14}b, q^{5}a, q^{8}a, q^{14}b, q^{5}b, q^{13}a$	-	-	-	0	-
$a, q^{15}b, q^{14}b, q^{13}a, q^{7}a, q^{5}b, q^{4}b, q^{4}a$	-	-	-	0	-
$a, q^{15}b, q^{14}b, q^{13}a, q^{8}a, q^{6}b, q^{5}b, q^{5}a$	-	-	-	0	-
$a, q^{15}b, q^{15}a$	-	-	-	0	0
$a, q^{7}b, q^{15}a$	-	-	-	0	0
$a, q^{12}a, q^{8}a, q^{4}a$	-	-	-	0	-
$a, q^{4}a, q^{8}a, q^{12}a$	-	-	-	0	-
$a, q^{14}a, q^{12}a, q^{10}a$	-	-	-	0	-
$a, q^{6}a, q^{12}a, q^{2}a$	-	-	-	0	-
$a, q^{10}a, q^{4}a, q^{14}a$	-	-	-	0	-
$a, q^{2}a, q^{4}a, q^{6}a$	-	-	-	0	-
$a, 0, q^{12}a$	-	-	-	0	-
$a, 0, q^{4}a$	-	-	-	0	-
$a, q^{9}a, q^{2}a, q^{11}a$	-	-	-	0	-
$a, qa, q^{2}a, q^{3}a$	-	-	-	0	-
$a, q^{15}a, q^{14}a, q^{13}a$	-	-	-	0	-
$a, q^{7}a, q^{14}a, q^{5}a$	-	-	-	0	-
$a, 0, q^{15}a$	-	-	-	0	-
$a, 0, q^{7}a$	-	-	-	0	-
$a, 0, q^{9}a$	-	-	-	0	-
$a, 0, qa$	-	-	-	0	-
$a, 0, q^{14}a$	-	-	-	0	-
$a, 0, q^{2}a$	-	-	-	0	-
$a, 0, q^{10}a$	-	-	-	0	-
$a, 0, q^{6}a$	-	-	-	0	-
$a, 0, q^{3}a$	-	-	-	0	-
$a, 0, q^{13}a$	-	-	-	0	-
$a, 0, q^{11}a$	-	-	-	0	-
$a, 0, q^{5}a$	-	-	-	0	-
$a, q^{9}a, q^{4}a, q^{13}a$	-	-	-	0	-
$a, qa, q^{4}a, q^{5}a$	-	-	-	0	-
$a, q^{15}a, q^{12}a, q^{11}a$	-	-	-	0	-
$a, q^{7}a, q^{12}a, q^{3}a$	-	-	-	0	-
$a, q^{13}a, q^{12}a, q^{9}a$	-	-	-	0	-
$a, q^{5}a, q^{12}a, qa$	-	-	-	0	-
$a, q^{11}a, q^{4}a, q^{15}a$	-	-	-	0	-
$a, q^{3}a, q^{4}a, q^{7}a$	-	-	-	0	-
$a, q^{8}a, q^{2}a, q^{10}a$	-	-	-	0	-
$a, a, q^{14}a, q^{14}a$	-	-	-	0	-
$a, a, q^{2}a, q^{2}a$	-	-	-	0	-
$a, q^{8}a, q^{14}a, q^{6}a$	-	-	-	0	-
$a, q^{13}a, q^{10}a, q^{7}a$	-	-	-	0	-
$a, q^{5}a, q^{10}a, q^{15}a$	-	-	-	0	-
$a, q^{11}a, q^{6}a, qa$	-	-	-	0	-
$a, q^{3}a, q^{6}a, q^{9}a$	-	-	-	0	-
$a, q^{12}a, q^{6}a, q^{2}a$	-	-	-	0	-
$a, q^{12}a, q^{10}a, q^{6}a$	-	-	-	0	-
$a, q^{4}a, q^{6}a, q^{10}a$	-	-	-	0	-
$a, q^{4}a, q^{10}a, q^{14}a$	-	-	-	0	-
$a, q^{14}a, q^{14}a, q^{12}a$	-	-	-	0	-
$a, q^{6}a, q^{14}a, q^{4}a$	-	-	-	0	-
$a, q^{2}a, q^{2}a, q^{4}a$	-	-	-	0	-
$a, q^{10}a, q^{2}a, q^{12}a$	-	-	-	0	-
$a, q^{14}a, q^{10}a, q^{8}a$	-	-	-	0	-
$a, q^{6}a, q^{10}a, a$	-	-	-	0	-
$a, q^{10}a, q^{6}a, a$	-	-	-	0	-

L	2	3	4	5	6
a, q^2a, q^6a, q^8a	-	-	-	0	-
$a, q^8a, qa, q^{10}a, q^2a$	-	-	-	0	-
a, a, qa, q^2a, q^2a	-	-	-	0	-
$a, a, q^{15}a, q^{14}a, q^{14}a$	-	-	-	0	-
$a, q^8a, q^{15}a, q^6a, q^{14}a$	-	-	-	0	-
$a, q^{13}a, q^8a, q^5a$	-	-	-	0	-
$a, q^5a, q^8a, q^{13}a$	-	-	-	0	-
$a, q^3a, q^8a, q^{11}a$	-	-	-	0	-
$a, q^{11}a, q^8a, q^3a$	-	-	-	0	-
$a, q^{12}a, q^9a, q^6a, q^2a$	-	-	-	0	-
$a, q^{12}a, q^7a, q^2a, q^{14}a$	-	-	-	0	-
$a, q^4a, q^9a, q^{14}a, q^2a$	-	-	-	0	-
$a, q^4a, q^7a, q^{10}a, q^{14}a$	-	-	-	0	-
a, q^9a, a, q^9a	-	-	-	0	-
a, qa, a, qa	-	-	-	0	-
$a, q^{15}a, a, q^{15}a$	-	-	-	0	-
a, q^7a, a, q^7a	-	-	-	0	-
$a, q^{14}a, q^{13}a, q^{12}a, q^{10}a$	-	-	-	0	-
$a, q^6a, q^{13}a, q^4a, q^{10}a$	-	-	-	0	-
$a, q^{10}a, q^3a, q^{12}a, q^6a$	-	-	-	0	-
$a, q^2a, q^3a, q^4a, q^6a$	-	-	-	0	-
$a, q^{14}a, q^{11}a, q^8a, q^6a$	-	-	-	0	-
$a, q^6a, q^{11}a, a, q^6a$	-	-	-	0	-
$a, q^{10}a, q^5a, a, q^{10}a$	-	-	-	0	-
$a, q^2a, q^5a, q^8a, q^{10}a$	-	-	-	0	-
$a, q^8a, a, q^9a, qa, q^9a$	-	-	-	0	-
$a, a, a, q^{15}a, q^{15}a, q^{15}a$	-	-	-	0	-
a, a, a, qa, qa, qa	-	-	-	0	-
$a, q^8a, a, q^7a, q^{15}a, q^7a$	-	-	-	0	-
$a, q^{12}a, q^8a, q^3a, q^{15}a, q^{11}a$	-	-	-	0	-
$a, q^{12}a, q^8a, q^5a, qa, q^{13}a$	-	-	-	0	-
$a, q^4a, q^8a, q^{11}a, q^{15}a, q^3a$	-	-	-	0	-
$a, q^4a, q^8a, q^{13}a, qa, q^5a$	-	-	-	0	-
a, q^9a, qa, q^9a, q^2a	-	-	-	0	-
a, qa, qa, qa, q^2a	-	-	-	0	-
$a, q^{15}a, q^{15}a, q^{15}a, q^{14}a$	-	-	-	0	-
$a, q^7a, q^{15}a, q^7a, q^{14}a$	-	-	-	0	-
$a, q^{13}a, q^9a, q^5a, q^2a$	-	-	-	0	-
$a, q^5a, q^9a, q^{13}a, q^2a$	-	-	-	0	-
$a, q^{11}a, q^7a, q^3a, q^{14}a$	-	-	-	0	-
$a, q^3a, q^7a, q^{11}a, q^{14}a$	-	-	-	0	-
$a, q^{12}a, q^6a, a, q^{12}a$	-	-	-	0	-
$a, a, q^{14}a, q^{12}a, q^{12}a$	-	-	-	0	-
$a, q^8a, q^{14}a, q^4a, q^{12}a$	-	-	-	0	-
$a, q^{12}a, q^{10}a, q^8a, q^4a$	-	-	-	0	-
$a, q^8a, q^2a, q^{12}a, q^4a$	-	-	-	0	-
a, a, q^2a, q^4a, q^4a	-	-	-	0	-
$a, q^4a, q^6a, q^8a, q^{12}a$	-	-	-	0	-
$a, q^4a, q^{10}a, a, q^4a$	-	-	-	0	-
$a, q^{14}a, q^{12}a, q^9a, q^7a, q^5a$	-	-	-	0	-
$a, q^6a, q^{12}a, qa, q^7a, q^{13}a$	-	-	-	0	-
$a, q^2a, q^4a, q^7a, q^9a, q^{11}a$	-	-	-	0	-
$a, q^{10}a, q^4a, q^{15}a, q^9a, q^3a$	-	-	-	0	-
$a, q^{14}a, q^{14}a, q^{14}a, q^{12}a$	-	-	-	0	-
$a, q^6a, q^{14}a, q^6a, q^{12}a$	-	-	-	0	-
$a, q^{10}a, q^2a, q^{10}a, q^4a$	-	-	-	0	-
$a, q^2a, q^2a, q^2a, q^4a$	-	-	-	0	-
$a, q^{14}a, q^{12}a, q^{11}a, q^9a, q^7a$	-	-	-	0	-
$a, q^6a, q^{12}a, q^3a, q^9a, q^{15}a$	-	-	-	0	-
$a, q^{10}a, q^4a, q^{13}a, q^7a, qa$	-	-	-	0	-
$a, q^2a, q^4a, q^5a, q^7a, q^9a$	-	-	-	0	-
$a, q^{10}a, q^6a, q^2a, q^{12}a$	-	-	-	0	-
$a, q^2a, q^6a, q^{10}a, q^{12}a$	-	-	-	0	-
$a, q^9a, q^4a, q^{15}a, q^8a$	-	-	-	0	-
a, qa, q^4a, q^7a, q^8a	-	-	-	0	-
$a, q^{13}a, q^{12}a, q^{11}a, q^8a$	-	-	-	0	-
$a, q^5a, q^{12}a, q^3a, q^8a$	-	-	-	0	-
$a, q^9a, q^3a, q^{13}a, q^6a$	-	-	-	0	-
a, qa, q^3a, q^5a, q^6a	-	-	-	0	-
$a, q^{15}a, q^{12}a, q^9a, q^8a$	-	-	-	0	-
$a, q^7a, q^{12}a, qa, q^8a$	-	-	-	0	-
$a, q^{11}a, q^4a, q^{13}a, q^8a$	-	-	-	0	-
$a, q^3a, q^4a, q^5a, q^8a$	-	-	-	0	-
$a, q^{15}a, q^{13}a, q^{11}a, q^{10}a$	-	-	-	0	-
$a, q^7a, q^{13}a, q^3a, q^{10}a$	-	-	-	0	-

L	2	3	4	5	6
$a, q^{13}a, q^{11}a, q^9a, q^6a$	-	-	-	0	-
$a, q^5a, q^{11}a, qa, q^6a$	-	-	-	0	-
$a, q^{11}a, q^5a, q^{15}a, q^{10}a$	-	-	-	0	-
$a, q^3a, q^5a, q^7a, q^{10}a$	-	-	-	0	-
$a, q^9a, qa, q^{11}a, q^3a, q^{12}a$	-	-	-	0	-
$a, qa, qa, q^3a, q^3a, q^4a$	-	-	-	0	-
$a, q^{13}a, q^{10}a, q^6a, q^3a, a$	-	-	-	0	-
$a, q^5a, q^{10}a, q^{14}a, q^3a, q^8a$	-	-	-	0	-
$a, q^2a, q^5a, q^6a, q^9a, q^{11}a$	-	-	-	0	-
$a, q^{10}a, q^5a, q^{14}a, q^9a, q^3a$	-	-	-	0	-
$a, q^{14}a, q^{11}a, q^{10}a, q^7a, q^5a$	-	-	-	0	-
$a, q^6a, q^{11}a, q^2a, q^7a, q^{13}a$	-	-	-	0	-
$a, q^3a, q^6a, q^{10}a, q^{13}a, a$	-	-	-	0	-
$a, q^{11}a, q^6a, q^2a, q^{13}a, q^8a$	-	-	-	0	-
$a, q^{15}a, q^{15}a, q^{13}a, q^{13}a, q^{12}a$	-	-	-	0	-
$a, q^7a, q^{15}a, q^5a, q^{13}a, q^4a$	-	-	-	0	-
a, a, a, a, a	-	-	-	-	0.3
a, q^8a, a, q^8a, a	-	-	-	-	0.3
$a, q^8a, 0, a, q^8a$	-	-	-	-	0.2
$a, a, 0, q^8a, q^8a$	-	-	-	-	0.2
a, q^8b, b, q^9a	-	-	-	-	0.1
$a, q^{15}c, q^{14}e, q^{13}g, q^{13}g, q^{14}e, q^{15}c, a$	-	-	-	-	0
$a, q^7c, q^{14}e, q^5g, q^{13}g, q^6e, q^{15}c, q^8a$	-	-	-	-	0
$a, q^{15}a, q^9a, q^7b, q^6a, a, q^{15}a$	-	-	-	-	0
$a, q^7a, q^9a, q^{15}b, q^6a, q^8a, q^{15}a$	-	-	-	-	0
$a, q^7b, q^{14}c, q^6c, q^{14}b, q^6a$	-	-	-	-	0
$a, q^{15}b, q^{15}b, q^{15}b, q^{15}a$	-	-	-	-	0
$a, q^7b, q^{15}b, q^7b, q^{15}a$	-	-	-	-	0
$a, q^{15}a, q^{10}a, q^8a, q^3a, q^2a$	-	-	-	-	0
$a, q^7a, q^{10}a, a, q^3a, q^{10}a$	-	-	-	-	0
$a, q^7b, q^{15}a, a, q^7b, q^{15}a$	-	-	-	-	0
$a, q^{15}b, q^{15}a, q^8a, q^7b, q^7a$	-	-	-	-	0
$a, q^{15}b, q^{15}b, q^{15}b, q^{15}a$	-	-	-	-	0
$a, q^{15}c, q^{14}e, q^{13}g, q^{14}e, q^{15}c, a$	-	-	-	-	0
$a, q^7c, q^{14}e, q^5g, q^{14}e, q^7c, a$	-	-	-	-	0
$a, b, c, qc, q^2c, q^3c, q^4b, q^5a$	-	-	-	-	0
$a, q^8b, c, q^9c, q^2c, q^{11}c, q^4b, q^{13}a$	-	-	-	-	0
$a, q^9a, q^6a, a, q^{13}a, q^6a$	-	-	-	-	0
$a, b, qb, q^3a, q^9a, q^{10}b, q^{11}b, q^{12}a$	-	-	-	-	0
$a, q^{15}b, q^{15}b, b, b, a$	-	-	-	-	0
$a, q^{15}b, q^{14}b, q^{12}b, q^{11}b, q^{11}a$	-	-	-	-	0
$a, q^8b, b, q^7b, q^{15}b, q^8a$	-	-	-	-	0
$a, q^7b, q^{15}b, q^8b, b, q^8a$	-	-	-	-	0
$a, q^7b, q^{14}b, q^4b, q^{11}b, q^3a$	-	-	-	-	0
$a, q^{15}b, q^{14}a, q^8a, q^6a, a, q^{14}b, q^{14}a$	-	-	-	-	0
$a, q^7b, q^{14}a, a, q^6a, q^8a, q^{14}b, q^6a$	-	-	-	-	0

A.8 Asymmetric Sequences

The following sequences which do not have the symmetry property conjectured in Section 5.4 were found for the 8-PSK and 16-PSK modulation schemes.

A.8.1 8-PSK

L	4	5
a, q^5a, qa	12	11.5
a, q^6a, q^3a	12.8	10.4
a, q^7a, q^5a	13.2	10.4
a, q^7a, q^7a	12	11.5
a, qa, qa	12	11.5
a, qa, q^3a	12.8	10.4
a, q^2a, q^5a	13.2	10.4
a, q^3a, q^7a	12	11.5
a, q^3a, q^7a, q^3a	-	1
a, q^2a, q^4a, q^7a	-	1
a, qa, q^3a, q^5a	-	1
a, qa, qa, qa	-	1
a, a, qa, q^2a	-	1
a, q^7a, q^7a, q^7a	-	1

L	4	5
a, q^7a, q^6a, q^6a	-	1
a, q^7a, q^5a, q^3a	-	1
a, q^6a, q^4a, qa	-	1
a, q^5a, qa, q^5a	-	1
a, q^4a, qa, q^6a	-	1
a, q^4a, q^7a, q^2a	-	1

A.8.2 16-PSK

L	3	4	5
$a, q^{15}b, q^{15}b, a$	0.2	0.1	0
$a, q^7b, q^{15}b, q^8a$	0.2	0.1	0
a, qb, q^2b, q^3a	0	0	-
$a, q^9b, q^2b, q^{11}a$	0	0	-
$a, q^{13}b, q^{10}b, q^7a$	0	0	-
$a, q^3b, q^7b, q^{12}a$	0	-	-
a, q^2a, q^3a	-	0	0
$a, q^{10}a, q^3a$	-	0	0
$a, q^7a, q^{13}a$	-	0	0
$a, q^{15}a, q^{13}a$	-	0	0
a, qa, qa	-	0	0
a, q^9a, qa	-	0	0
$a, q^7a, q^{15}a$	-	0	0
$a, q^{15}a, q^{15}a$	-	0	0
$a, q^5a, q^{11}a$	-	0	0
$a, q^{13}a, q^{11}a$	-	0	0
a, q^3a, q^5a	-	0	0
$a, q^{11}a, q^5a$	-	0	0
a, q^5a, q^9a	-	0	0
$a, q^{13}a, q^9a$	-	0	0
a, q^3a, q^7a	-	0	0
$a, q^{11}a, q^7a$	-	0	0
a, q^2a, q^2a, q^3a	-	0	0
$a, q^{10}a, q^2a, q^{11}a$	-	0	0
$a, q^7a, q^{15}a, q^5a$	-	0	0
$a, q^{15}a, q^{15}a, q^{13}a$	-	0	0
a, qa, q^3a, q^3a	-	0	-
$a, q^9a, q^3a, q^{11}a$	-	0	-
$a, q^7a, q^{13}a, q^5a$	-	0	-
$a, q^{15}a, q^{13}a, q^{13}a$	-	0	-
$a, q^5a, q^9a, q^{15}a$	-	0	0
$a, q^5a, q^{11}a, q^{15}a$	-	0	0
$a, q^{13}a, q^9a, q^7a$	-	0	0
$a, q^{13}a, q^{11}a, q^7a$	-	0	0
a, q^3a, q^5a, q^9a	-	0	0
a, q^3a, q^7a, q^9a	-	0	0
$a, q^{11}a, q^5a, qa$	-	0	0
$a, q^{11}a, q^7a, qa$	-	0	0
a, q^2a, q^5a, q^6a	-	0	0
$a, q^{10}a, q^5a, q^{14}a$	-	0	0
$a, q^7a, q^{12}a, q^2a$	-	0	0
$a, q^{15}a, q^{12}a, q^{10}a$	-	0	0
$a, q^5a, q^{12}a, q^2a$	-	0	0
$a, q^{13}a, q^{12}a, q^{10}a$	-	0	0
a, q^3a, q^4a, q^6a	-	0	0
$a, q^{11}a, q^4a, q^{14}a$	-	0	0
$a, q^5a, q^8a, q^{12}a$	-	0	0
$a, q^{13}a, q^8a, q^4a$	-	0	0
$a, q^{11}a, q^8a, q^4a$	-	0	0
$a, q^5b, q^{10}b, q^{15}a$	-	0	-
a, q^2b, q^5b, q^9a	-	0	-
$a, q^{10}b, q^5b, qa$	-	0	-
$a, q^6b, q^{13}b, q^5a$	-	0	-
$a, q^{14}b, q^{13}b, q^{13}a$	-	0	-
$a, q^{10}a, q^3a$	-	0	0
a, q^2a, q^3a	-	0	0
$a, q^{15}a, q^{13}a$	-	0	0
$a, q^7a, q^{13}a$	-	0	0
a, q^9a, qa	-	0	0
a, qa, qa	-	0	0
$a, q^{15}a, q^{15}a$	-	0	0
$a, q^7a, q^{15}a$	-	0	0
$a, q^{13}a, q^{11}a$	-	0	0

L	3	4	5
$a, q^5a, q^{11}a$	-	0	0
$a, q^{11}a, q^5a$	-	0	0
a, q^3a, q^5a	-	0	0
$a, q^{13}a, q^9a$	-	0	0
a, q^5a, q^9a	-	0	0
$a, q^{11}a, q^7a$	-	0	0
a, q^3a, q^7a	-	0	0
a, q^2a, qa, q^2a	-	-	0
$a, q^{10}a, qa, q^{10}a$	-	-	0
a, q^7a, a, q^6a	-	-	0
$a, q^{15}a, a, q^{14}a$	-	-	0
a, qa, qa, qa	-	-	0
a, q^9a, qa, q^9a	-	-	0
$a, q^7a, q^{15}a, q^7a$	-	-	0
$a, q^{15}a, q^{15}a, q^{15}a$	-	-	0
a, qa, q^2a, q^2a	-	-	0
$a, q^9a, q^2a, q^{10}a$	-	-	0
a, qa, q^4a, q^4a	-	-	0
$a, q^9a, q^4a, q^{12}a$	-	-	0
$a, q^5a, q^9a, q^{13}a$	-	-	0
$a, q^{13}a, q^9a, q^5a$	-	-	0
$a, q^7a, q^{12}a, q^4a$	-	-	0
$a, q^7a, q^{14}a, q^6a$	-	-	0
$a, q^7a, q^{13}a, q^3a$	-	-	0
$a, q^{15}a, q^{12}a, q^{12}a$	-	-	0
$a, q^{15}a, q^{14}a, q^{14}a$	-	-	0
$a, q^{15}a, q^{13}a, q^{11}a$	-	-	0
$a, q^3a, q^7a, q^{11}a$	-	-	0
$a, q^{11}a, q^7a, q^3a$	-	-	0
$a, q^5a, q^{11}a, qa$	-	-	0
$a, q^{13}a, q^{11}a, q^9a$	-	-	0
a, q^2a, q^4a, q^5a	-	-	0
$a, q^{10}a, q^4a, q^{13}a$	-	-	0
a, q^3a, q^5a, q^7a	-	-	0
$a, q^{11}a, q^5a, q^{15}a$	-	-	0
$a, qa, q^{15}a, q^{15}a$	-	-	0
$a, q^9a, q^{15}a, q^7a$	-	-	0
$a, q^5a, q^{10}a, q^{14}a$	-	-	0
$a, q^{13}a, q^{10}a, q^6a$	-	-	0
a, q^7a, qa, q^9a	-	-	0
$a, q^{15}a, qa, qa$	-	-	0
$a, q^3a, q^6a, q^{10}a$	-	-	0
$a, q^{11}a, q^6a, q^2a$	-	-	0
$a, q^5a, q^7a, q^{11}a$	-	-	0
$a, q^{13}a, q^7a, q^3a$	-	-	0
a, q^2a, q^3a, q^4a	-	-	0
$a, q^{10}a, q^3a, q^{12}a$	-	-	0
$a, q^7a, q^{14}a, q^4a$	-	-	0
$a, q^{15}a, q^{14}a, q^{12}a$	-	-	0
$a, q^3a, q^9a, q^{13}a$	-	-	0
$a, q^{11}a, q^9a, q^5a$	-	-	0
$a, q^5a, q^8a, q^{14}a$	-	-	0
$a, q^5a, q^{10}a, a$	-	-	0
$a, q^{13}a, q^8a, q^6a$	-	-	0
$a, q^{13}a, q^{10}a, q^8a$	-	-	0
$a, q^3a, q^8a, q^{12}a$	-	-	0
$a, q^3a, q^8a, q^{10}a$	-	-	0
a, q^3a, q^6a, q^8a	-	-	0
$a, q^{11}a, q^8a, q^2a$	-	-	0
$a, q^{11}a, q^6a, a$	-	-	0

References

[1] A. Abdi and M. Kaveh. A versatile spatio-temporal correlation function for mobile fading channels with non-isotropic scatterers. In *Proc. IEEE Workshop on Statistical Signal Processing*, pages 58–62, Pocono Manor, PA, 2000.

[2] K. Abed-Meraim, Y. Hua, and A. Belouchrani. Second-order near-field source localization: Algorithm and performance analysis. In *Proc. ASILOMAR Conf. Signal & Systems*, Nov 1996.

[3] T. D. Abhayapala, R. A. Kennedy, and R. C. Williamson. Spatial aliasing for near-field sensor arrays. *IEE Electronics Letters*, 35(10):764–765, May 1999.

[4] H. Akaike. A new look at the statistical model identification. *IEEE Trans. on Automatic Control*, 19(6):716–723, 1974.

[5] A. O. Allen. *Probability, Statistics and Queueing Theory.* Academic Press, Orlando, Florida, 1978.

[6] J. Bach Andersen, J. Jensen, S. H. Jensen, and F. Frederiksen. Prediction of future fading based on past measurements. In *Proc. IEEE Vehicular Technology Conference*, Fall, Delft, The Netherlands, Sep 1999.

[7] R. R. Anderson and G. J. Foschini. The minimum distance for MLSE digital data systems of limited complexity. *IEEE Trans. Information Theory*, 21:544–551, Sep 1975.

[8] E. Ban. Who funds mobile phone cancer study? *Nature Medicine*, 4(2), Feb 1998.

[9] W. J. Bangs. *Array Processing with Generalized Beamformers.* PhD thesis, Yale University, New Haven, CT, 1971.

[10] S. Barbarossa and A. Scaglione. Theoretical bounds on the estimation and prediction of multipath time-varying channels. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, volume V, pages 2545–2548, Istanbul, 2000.

[11] F. G. Bass and I. M Fuks. *Wave Scattering from Statistically Rough Surfaces*. Pergamon Press, Oxford, 1979. Translated and edited by C. B. Vesecky and J. F. Vesecky.

[12] P. Beckmann and A. Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Pergamon Press, Oxford, 1963.

[13] R. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961.

[14] P. A. Bello. Characterization of randomly time-variant linear channels. *IEEE Trans. Circuits & Systems*, 11:360–393, Dec 1963.

[15] W. R. Bennett. Distribution of the sum of randomly phased components. *Quart. Applied Math.*, 5:385–393, Jan 1948.

[16] L. L. Beranek. *Acoustics*. Acoustical Society of America, 1954.

[17] H. Bernhard. A tight upper bound on the gain of linear and nonlinear predictors for stationary stochastic processes. *IEEE Trans. Signal Processing*, 46(11):2909–2917, Nov 1998.

[18] H. Bernhard and G. Kubin. A fast mutual information calculation algorithm. In *Proc. 7th European Signal Processing Conference, EUSIPCO*, Edinburgh, Scotland, Sep 1994.

[19] G. E. Bottomley and S. Chennakeshu. Unification of MLSE receivers and extension to time-varying channels. *IEEE Trans. Communications*, 46(4):464–472, Apr 1998.

[20] M. S. Brandstein and H. Silverman. A practical methodology for speech source localisation with microphone arrays. *Computer, Speech and Language*, 11(2):91–126, 1997.

[21] W. R. Braun and U. Dersch. A physical mobile radio channel model. *IEEE Trans. Vehicular Technology*, 40(2):472–428, May 1991.

[22] R. J. C. Bultitude, T. J. Willink, M. H. A. J. Herben, and G. Brussaard. Detection of changes in the spectra of measured CW mobile radio data for space wave modeling applications. In *Proc. 20th Biennial Symposium on Communications*, Queen's University Kingston, Ontario, Canada, May 2000.

[23] R. J. C. Bultitude, T. J. Willink, M. H. A. J. Herben, and G. Brussaard. Radio channel modelling for terrestrial vehicular mobile applications. In *Proc. AP2000 Millenium Conf. on Antennas and Propagation*, Davos, Switzerland, Apr 2000.

[24] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[25] G. Castellini, F. Conti, E. Del Re, and L. Pierucci. A continuously adaptive MLSE receiver for mobile communications: Algorithm and performance. *IEEE Trans. Communications*, 45(1):80–89, Jan 1997.

[26] J. K. Cavers. Variable-rate transmission for Rayleigh fading channels. *IEEE Trans. Communications*, 20:15–22, Feb 1972.

[27] J. Chen, A. Paulraj, and U. Reddy. Multichannel maximum-likelihood sequence estimation (MLSE) equalizer for GSM using a parametric channel model. *IEEE Trans. Communications*, 47(1):53–63, Jan 1999.

[28] J. T. Chen. Multi-channel MLSE equalizer with parametric FIR channel identification. In *Proc. IEEE Vehicular Technology Conference*, pages 710–714, Phoenix, AZ, 1997.

[29] C. Chou and D. W. Lin. Training sequence and memory length selection for space-time Viterbi equalization. *Journal of Communications and Networks*, 2(4):361–366, Dec 2000.

[30] R. H. Clarke. A statistical theory of mobile-radio reception. *The Bell System Technical Journal*, 47:957–1000, July-August 1968.

[31] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer-Verlag, Berlin, 1992.

[32] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. D. Thompson, Jr. Measurement of correlation coefficients in reverberant sound fields. *J. Acoustical Soc. America*, 27(6):1072–1077, Nov 1955.

[33] C. Cortes and V. N. Vapnik. Support vector networks. *Machine Learning*, 20:273–297, 1995.

[34] A. J. Coulson. *Characterization of the Mobile Radio Multipath Fading Channel*. PhD thesis, University of Auckland, 1998.

[35] C. A. Coulson and A. Jeffrey. *Waves: A mathematical approach to the common types of wave motion*. Longman, London, second edition, 1977.

[36] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines*. Cambridge University Press, Cambridge, UK, 2000.

[37] U. Dersch and E. Zollinger. Multiple reflections of radio waves in a corridor. *IEEE Trans. Antennas & Propagation*, 42(11):1571–1574, Nov 1994.

[38] U. Dersch and E. Zollinger. Physical characteristics of urban-microcellular propagation. *IEEE Trans. Antennas & Propagation*, 42(11):1528–1539, Nov 1994.

[39] U. Dersch and E. Zollinger. Propagation mechanisms in microcell and indoor environments. *IEEE Trans. Vehicular Technology*, 43(4):1058–1066, Nov 1994.

[40] A. Duel-Hallen, S. Hu, and H. Hallen. Long-range prediction of fading signals. *IEEE Signal Proc. Mag.*, pages 62–75, May 2000.

[41] T. Ekman. *Prediction of Mobile Radio Channels.* Technical Licentiate Thesis, Signals and Systems,Uppsala University, Nov 2000.

[42] T. Ekman and G. Kubin. Nonlinear prediction of mobile radio channels: Measurements and MARS model designs. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, pages 2667–2670, Phoenix, AR, Mar 1999.

[43] T. Ekman, G. Kubin, M. Sternad, and A. Ahlén. Quadratic and linear filters for mobile radio channel prediction. In *Proc. IEEE Vehicular Technology Conference*, Fall, Amsterdam, Sep 1999.

[44] T. Eyceöz, A. Duel-Hallen, and H. Hallen. Prediction of fast fading parameters by resolving the interference pattern. In *Proc. ASILOMAR Conf. Signal & Systems*, volume 1, pages 167–171, 1997.

[45] T. Eyceöz, A. Duel-Hallen, and H. Hallen. Deterministic channel modeling and long range prediction of fast fading mobile radio channels. *IEEE Communications Letters*, 2(9):254–256, Sep 1998.

[46] T. Eyceöz, S. Hu, and A. Duel-Hallen. Performance analysis of long range prediction for fast fading channels. In *Proc. 33rd Annu. Conf. Inform. Sciences and Systems, ICC'99*, volume 2, pages 656–661, 1999.

[47] M. Failli. Digital land mobile radio communications. Technical Report COST 207, COST 207 Management Committee, Commission of the European Communities, Luxembourg, Sep 1988.

[48] G. D. Forney, Jr. Lower bounds on error probability in the presence of large intersymbol interference. *IEEE Trans. Communications*, 20:76–77, Feb 1972.

[49] G. D. Forney, Jr. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. Information Theory*, 18(3):363–378, May 1972.

[50] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6:311–335, 1998.

[51] A. M. Fraser. Information and entropy in strange attractors. *IEEE Trans. Information Theory*, 35(2):245–262, Mar 1989.

[52] J. M. Friedman. Multivariate adaptive regression splines. *The Annals of Statistics*, 19(1):1–141, 1991.

[53] J. Fuhl, A. F. Molisch, and E. Bonek. Unified channel model for mobile radio systems with smart antennas. *IEE Proc. — Radar, Sonar Navig.*, 145(1):32–41, Feb 1998.

[54] X. M. Gao, X. Z. Gao, J. M. A. Tanskanen, and S. J. Ovaska. Power prediction in mobile communication systems using an optimal neural-network structure. *IEEE Trans. Neural Networks*, 8(6):1446–1455, Nov 1997.

[55] W. Gautschi. Efficient computation of the complex error function. *SIAM J. Numerical Analysis*, 7:187–198, 1970.

[56] D. A. George and R. R. Bowen. Comments on "lower bounds to error probability in the presence of large intersymbol interference". *IEEE Trans. Communications*, 20:77–78, Feb 1972.

[57] A. J. Gersho. Characterization of time-varying linear systems. *Proc. IEEE*, 51:283, 1963.

[58] G. B. Giannakis and C. Tepedelenlioğlu. Basis expansion models and diversity techniques for blind identification and equalization of time varying channels. *Proc. IEEE*, 86(10):1969–1986, Oct 1998.

[59] R. N. Giere. Objective single-case probabilities and the foundations of statistics. In P. Suppes, L. Henkin, A. Joja, and G. C. Moisil, editors, *Proc. Logic, Methodology and Philosophy of Science*, pages 467–483, Amsterdam, North-Holland, 1973.

[60] A. Goldsmith and S. Chua. Variable-rate variable-power MQAM for fading channels. *IEEE Trans. Communications*, 45:1218–1230, Oct 1997.

[61] A. J. Goldsmith and S. G. Chua. Adaptive coded modulation for fading channels. *IEEE Trans. Communications*, 46:595–601, May 1998.

[62] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, 3rd edition, 1996.

[63] A. Gorokhov. On the performance of the Viterbi equalizer in the presence of channel estimation errors. *IEEE Signal Processing Letters*, 5(12):321–324, Dec 1998.

[64] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, San Diego, Fifth edition, 1994.

[65] W. Gröbner and N. Hofreiter. *Integraltafel, Erster Teil, Unbestimmte Integrale*. Springer-Verlag, Wein and Innsbruck, 1949.

[66] W. Gröbner and N. Hofreiter. *Integraltafel, Zweiter Teil, Bestimmte Integrale*. Springer-Verlag, Wein and Innsbruck, 1950.

[67] S. Grose. Algorithm gives milliseconds' warning of phone fade: it's enough. *The Canberra Times*, page 17, Feb 12 2001.

[68] J. Grouffaud, P. Larzabal, A. Ferréol, and H. Glergeot. Adaptive maximum likelihood algorithms for the blind tracking of time-varying multipath channels. *Int. J. Adapt. Control Signal Process.*, 12:207–222, Mar 1998.

[69] Y. Guo. *Aspects of Kernel Based Learning Algorithms*. PhD thesis, RSISE, The Australian National University, Canberra, Nov 2001.

[70] F. Hansen and F. I. Meno. Mobile fading — Rayleigh and lognormal superimposed. *IEEE Trans. Vehicular Technology*, 26:332–335, Nov 1977.

[71] H. Hashemi. The indoor radio propagation channel. *Proc. IEEE*, 81:943–968, Jul 1993.

[72] S. Haykin. *Communication Systems*. John Wiley, New York, 3rd edition, 1994.

[73] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[74] J. Hwang and J. H. Winters. Sinusoidal modeling and prediction of fast fading processes. In *Proc. GLOBECOM*, volume 2, pages 892–897, Sydney, Australia, Nov 1998.

[75] P. D. Inskip, R. E. Tarone, E. E. Hatch, T. C. Wilcosky, W. R. Shapiro, R. G. Selker, H. A. Fine, P. M. Black, J. S. Loeffler, and M. S. Linet. Cellular-telephone use and brain tumours. *New England J. Medicine*, 344(2):76–79, 2001.

[76] W. C. Jakes, editor. *Microwave Mobile Communications*. John Wiley, New York, 1974.

[77] D. H. Johnson and D. E. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Prentice Hall, Englewood Cliffs, NJ, 1993.

[78] H. M. Jones, R. A. Kennedy, and T. D. Abhayapala. On dimensionality of multipath fields: Spatial extent and richness. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, Orlando, Florida, 2002. Submitted.

[79] M. Kalkan and R. H. Clarke. Prediction of the space-frequency correlation function for base station diversity reception. *IEEE Trans. Vehic. Tech.*, 46(1):176–184, Feb 1997.

[80] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall, Englewood Cliffs, NJ, 1993.

[81] M. G. Kendall and A. Stuart. *The Advanced Theory of Statistics, Volume 1 Distribution Theory.* Charles Griffin & Company Limited, Second edition, 1963.

[82] A. Khenchaf. Bistatic reflection of electromagnetic waves from random rough surfaces: Application to the sea surface and snowy-covered. *The European Physical Journal – Applied Physics*, 14:45–62, 2001.

[83] E. Kreyszig. *Advanced Engineering Mathematics.* John Wiley & Sons, New York, 1988.

[84] R. Kumaresan. On the zeros of the linear prediction-error filter for deterministic signals. *IEEE Trans. Acoustics, Speech & Signal Processing*, 31(1):217–220, Feb 1983.

[85] R. Kumaresan and D. W. Tufts. Estimating the angles of arrival of multiple plane waves. *IEEE Trans. on Aerospace and Electronic Systems*, 19(1):134–139, Jan 1983.

[86] L. Landau. Monotonicity and bounds on Bessel functions. *Electronic Journ. Diff. Eqns*, Conf 04:147–154, 2000.

[87] T. Larsson. *A State-Space Partitioning Approach to Trellis Decoding.* PhD thesis, Chalmers University of Technology, Göteberg, Sweden, 1991. Technical Report No. 222.

[88] D. I. Laurenson. *Indoor Radio Channel Propagation Modelling by Ray Tracing Technqiues.* PhD thesis, The University of Edinburgh, 1994.

[89] H. Lee, Y. Chen, and C. Yeh. A covariance approximation method for near-field direction-finding using a uniform linear array. *IEEE Trans. Signal Processing*, 43(5):1293–1298, May 1995.

[90] W. C. Y. Lee. *Mobile Communications Engineering.* McGraw-Hill, New York, 1982.

[91] P. A. W. Lewis and J. G. Stevens. Nonlinear modelling of time series using multivariate adaptive regression splines (MARS). *Journal of the American Statistical Association*, 86:864–877, Dec 1991.

[92] J. P. Linnartz. *Narrowband Land-Mobile Radio Networks*. Artech House, Boston, 1993.

[93] F. R. Magee and J. G. Proakis. An estimate of the upper bound on error probability for maximum-likelihood sequence estimation on channels having a finite-duration impulse response. *IEEE Trans. Information Theory*, 19:699–702, Sep 1973.

[94] N. W. McLachlan. *Bessel Functions for Engineers*. Oxford University Press, second edition, 1961.

[95] M. Médard. The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel. *IEEE Trans. Information Theory*, 46(3):933–946, May 2000.

[96] D. H. Mellor. *The Matter of Chance*. Cambridge University Press, Cambridge, 1971.

[97] J. M. Mendel. *Lessons in Estimation Theory for Signal Processing, Communications and Control*. Prentice Hall, Englewood Cliffs, NJ, 1995.

[98] H. Meyr, M. Moeneclaey, and S. A. Fechtel. *Digital Communication Receivers*, volume 2, Synchronization, Channel Estimation, and Signal Processing. John Wiley & Sons, New York, 1997.

[99] J. E. Moulder, L. S. Erdreich, R. S. Malyapa, J. Merritt, W. F. Pickard, and Vijayalaxmi. Cell phones and cancer: What is the evidence for a connection? *Radiation Research*, 151(5):513–531, May 1999.

[100] M. Nakagami. *The m-distribution — a general formula of intensity distribution of rapid fading*. Pergamon Press, Oxford, 1960.

[101] Sir I. S. Newton. *PhilosophiæNaturalis Principia Mathematica*. University of California Press, Berkely, California, Cajori edition, 1946. (from the preface to the second (1713) edition by R. Cotes).

[102] O. Nørklit, P. D. Teal, and R. G. Vaughan. Measurement and evaluation of multi-antenna handsets in indoor mobile communication. *IEEE Trans. Antennas & Propagation*, 49(3):429–437, Mar 2001.

[103] H. Nyquist. Certain topics in telegraph transmission theory. *AIEE Trans.*, 47:617–644, 1928.

[104] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, 1965.

[105] A. Papoulis. Levinson's algorithm, Wold's decomposition and spectral estimation. *SIAM Review*, 27(3):405–441, Sep 1985.

[106] A. Papoulis. Predictable processes and Wold's decomposition: a review. *IEEE Trans. Acoustics, Speech & Signal Processing*, 33(4):933–938, Aug 1985.

[107] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, third edition, 1991.

[108] J. D. Parsons. *The Mobile Radio Propagation Channel*. Pantech Press, London, 1992.

[109] J. D. Parsons and J. G. Gardiner. *Mobile Communications Systems*. Blackie, London and Halstead Press, New York, 1989.

[110] M. Pätzold, U. Killat, and F. Laue. A deterministic digital simulation model for Suzuki processes with application to a shadowed Rayleigh land mobile radio channel. *IEEE Trans. Vehicular Technology*, 45(2):318–331, May 1996.

[111] M. Pätzold, U. Killat, F. Laue, and Y. Li. On the statistical properties of deterministic simulation models for mobile fading channels. *IEEE Trans. Vehicular Technology*, 47(1):254–269, Feb 1998.

[112] K. L. Pedersen, P. E. Mogensen, and B. H. Fleury. Power azimuth spectrum in outdoor environments. *IEE Electronics Letters*, 33(18):1583–1584, 1997.

[113] J. N. Pierce and S. Stein. Multiple diversity with non-independent fading. *Proc IRE*, 48:89–104, Jan 1960.

[114] K. R. Popper. The propensity interpretation of probability. *The British Journal for the Philosophy of Science*, 10:25–42, 1959.

[115] A. W. Preece, G. Iwi, A. Davies-Smith, K. Wesnes, S. Butler, E. Lim, and A. Varey. Effect of a 915-MHz simulated mobile phone signal on cognitive function in man. *International Journal of Radiation Biology*, 75(4):447–456, Apr 1999.

[116] J. G. Proakis. *Digital Communications*. McGraw-Hill, New York, 2nd edition, 1989.

[117] X. Qiu and K. Chawla. On the performance of adaptive modulation in cellular systems. *IEEE Trans. Communications*, 47:884–894, Jun 1999.

[118] G. G. Raleigh and J. M. Cioffi. Spatio-temporal coding for wireless communication. *IEEE Trans. Communications*, 46(3):357–366, Mar 1998.

[119] G. G. Raleigh and V. K. Jones. Multivariate modulation and coding for wireless communication. *IEEE J. Selected Areas in Communications*, 17(5):851–866, May 1999.

[120] T. S. Rappaport. *Wireless Communications, Principles and Practice*. Prentice Hall, Upper Saddle River, NJ, 1996.

[121] Lord Rayleigh. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *Phil. Mag. J. Sci*, 10(60):73–78, 1880.

[122] S. M. Redl, M. K. Weber, and M. W. Oliphant. *An Introduction to GSM*. Artech House, Boston, 1995.

[123] D. C. Rife and R. R. Boorstyn. Multiple tone parameter estimation from discrete-time observations. *The Bell System Technical Journal*, 55(9):1389–1410, Nov 1976.

[124] J. Rissanen. A universal prior for integers and estimation by minimum description length. *The Annals of Statistics*, 11(2):416–431, 1983.

[125] R. Roy and T. Kailath. ESPRIT–estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoustics, Speech & Signal Processing*, 37(7):984–995, 1989.

[126] A. J. Rustako, Jr., N. Amitay, G. J. Owens, and R. S. Roman. Radio propagation at microwave frequencies for line-of-sight microcellular mobile and personal communications. *IEEE Trans. Vehicular Technology*, 40(1):203–210, Feb 1991.

[127] W. C. Salmon. Propensities: A discussion review of D. H. Mellor's – "the matter of chance". *Erkenntnis*, 14:183–216, 1979.

[128] J. Salz and J. H. Winters. Effect of fading correlation on adaptive arrays in digital mobile radio. *IEEE Trans. Vehicular Technology*, 42(4):1049–1057, Nov 1994.

[129] R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas & Propagation*, 34(3):276–280, Mar 1986.

[130] M. Schwartz, W. R. Bennett, and S. Stein. *Communication Systems and Techniques*. McGraw-Hill, New York, 1966. (reprinted IEEE Press, 1996).

[131] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.

[132] W. Ser, K. Tan, and K. Ho. A new method for determining "unknown" worst-case channels for maximum-likelihood sequence estimation. *IEEE Trans. Communications*, 46(2):164–168, Feb 1998.

[133] M. Shinozuka and C. M. Jan. Digital simulation of random processes and its application. *Journal of Sound and Vibration*, 25(1):111–128, 1972.

[134] G. L. Sicuranza. Quadratic filters for signal processing. *Proc. IEEE*, 80(8):1263–1285, Aug 1992.

[135] M. Slack. The probability of sinusoidal oscillations combined in random phase. *J. IEEE*, 93, part III:76–86, 1946.

[136] M. R. Spiegel. *Mathematical Handbook of Formulas and Tables*. McGraw-Hill, New York, 1968.

[137] M. D. Springer. *The Algebra of Random Variables*. John Wiley, New York, 1979.

[138] H. Staras. Diversity reception with correlated signals. *J. Applied Physics*, 27:93–94, Jan 1956.

[139] D. Starer and A. Nehorai. Passive localization of near-field sources by path following. *IEEE Trans. Signal Processing*, 42(3):677–680, Mar 1994.

[140] M. Sternad, T. Ekman, and A. Ahlén. Power prediction on broadband channels. In *Proc. IEEE Vehicular Technology Conference*, Spring, Rhodes, Greece, May 2001. IEEE.

[141] P. Stoica, B. Ottersten, M. Viberg, and R. L. Moses. Maximum likelihood array processing for stochastic coherent sources. *IEEE Trans. Signal Processing*, 44(1):96–105, Jan 1996.

[142] F. G. Stremler. *Introduction to Communications Systems*. Addison Wesley, Reading, MA, second edition, 1982.

[143] P. Suppes. Propensity representations of probability. *Erkenntnis*, 26:335–358, 1987.

[144] H. Suzuki. A statistical model for urban radio propagation. *IEEE Trans. Communications*, 25:673–680, Jul 1977.

[145] T. Svantesson and A. Ranheim. Mutual coupling effects on the capacity of multielement antenna systems. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, 2001.

[146] P. Teal. Error performance of a channel of known impulse response. In *2nd Australian Communications Theory Workshop*, Adelaide, Feb 2001. http://www.itr.unisa.edu.au/~alex/AusCTW2001/talks/teal1.html.

[147] P. Teal. Performance bounds for real time prediction of the mobile multipath channel. In *2nd Australian Communications Theory Workshop*, Adelaide, Feb 2001.
http://www.itr.unisa.edu.au/~alex/AusCTW2001/talks/teal2.html.

[148] P. D. Teal. Rough surface scattering and prediction of the mobile channel. *IEEE Communications Letters*, 2001. To be submitted.

[149] P. D. Teal, T. A. Abhayapala, and R. A. Kennedy. Spatial correlation for general distributions of scatterers. *IEEE Signal Processing Letters*, 2002. (to appear).

[150] P. D. Teal, T. A. Abhayapala, and R. A. Kennedy. Spatial correlation in non-isotropic scattering scenarios. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, Orlando, Florida, 2002. (to appear).

[151] P. D. Teal, R. Raich, and R. G. Vaughan. Prediction of fading in the mobile multipath environment. *IEE Proceedings — Communications*, 2000. Submitted.

[152] P. D. Teal and R. G. Vaughan. Simulation and performance bounds for real-time prediction of the mobile multipath channel. In *Proc. IEEE Workshop on Statistical Signal Processing*, pages 548–551, Singapore, Aug 2001.

[153] P. D. Teal, R. C. Williamson, and R. A. Kennedy. Error performance of a channel of known impulse response. In *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 2733–2736, Istanbul, June 2000.

[154] C. W. Therrien. *Discrete Random Signals and Statistical Signal Processing*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[155] M. K. Tsatsanis and G. B. Giannakis. Equalization of rapidly fading channels: Self-recovering methods. *IEEE Trans. Communications*, 44(5):619–630, May 1996.

[156] M. K. Tsatsanis and G. B. Giannakis. Modelling and equalisation of rapidly fading channels. *International Journal of Adaptive Control and Signal Processing*, 10:159–176, Mar 1996.

[157] D. W. Tufts and R. Kumaresan. Estimation of frequencies of multiple sinusoids: Making linear prediction behave like maximum liklihood. *Proceedings of the IEEE*, 70(9):975–989, Sep 1982.

[158] D. W. Tufts and R. Kumaresan. Singular value decomposition and improved frequency estimation using linear prediction. *IEEE Trans. Acoustics, Speech & Signal Processing*, 30:671–675, Aug 1982.

[159] R. Vaughan, P. Teal, and R. Raich. Prediction of fading signals in a multipath environment. In *Proc. IEEE Vehicular Technology Conference*, volume 1, pages 751–758, Fall, Boston, Sep 2000.

[160] R. G. Vaughan. Pattern translation and rotation in uncorrelated source distributions. *IEEE Trans. Antennas & Propagation*, 46(7):982–990, Jul 1998.

[161] R. G. Vaughan. Spaced directive antennas for mobile communications by the Fourier transform method. *IEEE Trans. Antennas & Propagation*, 48(7):1025–1032, Jul 2000.

[162] R. G. Vaughan and J. Bach Andersen. *Prediction of Signals in Mobile Communications and Systems that use the Prediction for Signal Reception and Coding.* New Zealand patent application 280348, November 1995; WO 99/26423.

[163] R. G. Vaughan and J. Bach Andersen. *Wireless Channels, Propagation and Antennas.* Peregrinus, London. To be published.

[164] R. G. Vaughan and N. L. Scott. Terminated in-line monopoles for vehicular diversity. In *Proc. Int. Union of Radio Science (USRI) Trienniel Symposium on Electromagnetic Theory*, Sydney, Australia, Aug 1992.

[165] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. Acoustics, Speech & Signal Processing*, 33(2):387–392, Apr 1985.

[166] W. T. Webb and R. Steele. Variable rate QAM for mobile radio. *IEEE Trans. Communications*, 43:2223–2230, Jul 1995.

[167] W. Weibull. A statistical function of wide distribution. *J. Applied Mechanics*, 18:293–297, 1951.

[168] J. H. Wilkinson and C. Reinsch. *Handbook for Automatic Computation*, volume 2. Springer-Verlag, New York, 1971.

[169] H. Wold. *The Analysis of Stationary Time Series.* Almquist and Wicksell, Uppsala, Sweden, 2nd edition, 1954.

Index

Abdi, 23
Abed-Meraim, 53
Abhayapala, iii, 49, 55, 73, 94, 104
adaptation, 33
adaptive algorithm, 5
adaptive arrays, 16
adaptive modulation, 4, 68
adaptive techniques, 1
AGC (Automatic Gain Control), 45
Ahlén, 8, 87, 147
Akaike, 64
Akaike information criterion, 64
aliasing, 49
Allen, 72
Amitay, 28
Andersen, 3, 4, 8, 13, 58, 62, 67, 79
Anderson, 107
antenna gain, 9, 23
applications, 3
array processing, 10, 35
autocorrelation, 14
autoregressive model, 37, 51, 61, 74

Ban, 2
Bangs, 80
Barbarossa, 11, 81
Bass, 88, 90
battery, 2
beam width, 24
Beckmann, 88
Bellman, 93
Bello, 14, 15
Belouchrani, 53
Bennett, 6, 28, 29
BER, 16
Berendt, 17, 19
Bernhard, 68, 93, 94
Bessel function, 19, 61
    modified, 23, 104
Black, 2
Bonek, 25, 26
Boorstyn, 82
Bottomley, 5, 108
Bowen, 107
BPSK, 16
Brandstein, 52
Braun, 16
broadband, 8
Brussaard, 15
Bultitude, 15
Burges, 65
Butler, 2

Castellini, 108
Cavers, 4
channel
    avoidance, 3, 69, 107
    ensemble, 9
    gain, 34
    imperfect knowledge, 126
    inversion, 4
    model, 3, 9
    state information, 4
    worst case, 108
channel measurement, 75
characteristic function, 30
Chawla, 4
Chen, 5, 28, 53

Chennakeshu, 5, 108
chi-square distribution, 30
Chou, 131
Chua, 4
Cioffi, 4
Clarke, 18, 26, 59
co-ordinate transformation, 125
coherence bandwidth, 14, 17
Colton, 19, 21, 95
combining, 2, 6
    equal gain, 29
    maximal ratio, 29, 33
constellation, 4, 109
Conti, 108
convergence, 5, 126
convolution matrix, 127, 128
Cook, 17, 19
correlation, 17, 28, 41
    spatial, 16, 17, 19, 21, 23, 24, 27, 29, 90, 145
correlation coefficient, 17
Cortes, 65
Coulson, 15, 16, 94
covariance, 14, 46
covariance methods, 50
Cramer Rao bound, 68, 79, 145
Cristianini, 65

Davies-Smith, 2
demodulation, 34
derivative of likelihood function, 46
Dersch, 16, 28, 71
deterministic model, 9, 15, 33
DFE, 5, 107
difference alphabet, 108, 121
diffraction, 1, 39
diffuse field, 17, 28, 88, 92
dimensionality
    curse of, 93
    multipath, 94–104
distance of error event, 113
distribution
    chi-square, 30
    exponential, 72
    generalised gamma, 16
    Laplacian, 25
    lognormal, 16
    Nakagami-m, 16
    normal, 15
    Rayleigh, 15, 81
    Ricean, 15
    Suzuki, 16
    truncated normal, 26
    uniform, 81, 92
    von Mises, 23
    Weibull, 16
diversity, 2, 16, 29, 41, 42
    frequency, 2
    order of, 33
    polarisation, 2
    selection, 2, 29
    space, 2
    time, 2
    transmit, 4
Doppler frequency, 10, 12, 17, 62
Doppler spectrum, 37
Dudgeon, 34, 35
Duel-Hallen, 4, 8, 28, 51, 74, 87

Edelman, 17, 19
effective sources, 61
eigendecomposition, 31, 49, 51, 57, 58, 113, 125
Ekman, 8, 38, 39, 74, 87, 93, 147
ellipsoid, 113, 114, 125
ensemble, 15, 16
entropy, 93
equalisation, 1, 5, 16, 33
    parametric, 5
    Viterbi, 149

Erdreich, 2
error probability, 108
ESPRIT, 58, 74
Eyceöz, 8, 28, 51, 74

fade margin, 2
fade timing, 3
Failli, 28, 71
far field, 34
FDMA, 3, 107
Fechtel, 6
feedback channel, 2
Ferréol, 5
filter
    matched, 5
    transmit, 108
Fine, 2
Fisher information, 80, 82, 83
    singular, 82, 87
fixed sector prediction, 61
Forney, Jr, 107, 111
Foschini, 16, 107
Fourier Transform, 101
fractional sampling, 108
Fraser, 93
Frederiksen, 8, 58, 67, 79
frequency estimation, 62
frequency hopping, 4
Fresnel zone, 39, 88, 90
Friedman, 38
Fuhl, 25, 26
Fuks, 88, 90

gamma distribution, generalised, 16
Gans, 16
Gao, 4
Gardiner, 16
Gaussian model, 59
Gautschi, 26
George, 107
Giannakis, 5, 28, 71
Giere, 15
Glergeot, 5
Goldsmith, 4
Golub, 58, 119
Gorokhov, 126, 128, 131
Gröbner, 24, 26, 104
gradient, 8
Gradshteyn, 17, 23, 55, 134
grazing angle, 90
Grouffaud, 5
GSM, 4, 6, 7
Guo, 66

Hadamard product, 47
Hallen, 4, 8, 28, 87
Hansen, 16
harmonic expansion, 19, 26
Hashemi, 15, 16, 28, 71
Hatch, 2
Haykin, 15
Herben, 15
Ho, 113, 142
Hofreiter, 24, 26, 104
Horn, 115, 117, 119
Hu, 4, 51, 74, 87
Hua, 53
Hwang, 6, 51

imperfect channel knowledge, 126
impulse response, 14, 40, 107, 110
    notation, 11
Inskip, 2
interference, 1, 2, 36
interleaving, 2
interpolation, 79
irregular sampling, 52
iso-surface, 113
Iwi, 2

Jakes, 15, 16, 22, 28, 61, 71
Jan, v, 28

Jeffrey, 94
Jemima, v
Jensen, 8, 58, 67, 79
Johnson, 34, 35, 115, 117, 119
Jones, 4, 94, 104

Kailath, 58, 65
Kalkan, 18, 26
Kaveh, 23
Kay, 45, 80
Kendall, 84, 136
Kennedy, iii, v, 49, 55, 73, 94, 104
Khenchaf, 88
Killat, 28
Kolmogorov Smirnov test, 72, 128
Kress, 19, 21, 95
Kreyszig, 98, 100
Kubin, 8, 93, 94, 147
Kumaresan, 56, 57, 92

Lagrange multiplier, 57, 117
Landau, 96
Laplacian distribution, 25
Larsson, 107, 113, 115, 120, 156
Larzabal, 5
latency, 33
Laue, 28
Lee, 15, 53
Lewis, 38
Li, 28
likelihood, 45, 49, 55, 56, 66, 111
likelihood function, 43
Lim, 2
Lin, 131
linear prediction, 37, 57, 70
Linet, 2
Linnartz, 16
LINPACK, 83
Loan, 58, 119
Loeffler, 2
log-periodic antenna, 75
lognormal distribution, 16

Médard, 126
Magee, 107
Malyapa, 2
Marta, v
McLachlan, 95, 98
MDL, 61, 64, 69
measurement segment, 70
Mellin convolution, 127
Mellor, 15
Mendel, 45
Meno, 16
Merritt, 2
Meyr, 6
MIMO, 4
minimum norm, 56, 57
MLSE, 107, 114
modal analysis, 18
model
    deterministic, 8
    parsimonious, 8, 41
    syndetic, 7, 9, 33, 41, 145
model order, 64, 69
modified covariance method, 51
Moeneclaey, 6
Molisch, 25, 26
monopole, 32, 75
Moore-Penrose inverse, 45, 58, 60, 63
Moses, 80
Moulder, 2
MSE, 5
multipath
    acoustic, 40
    dimensionality, 94–104
    environment, 18
    fading, 1
    propagation, 9
multiple antennas, 16
multiple receivers, 108

multiplicity of error event, 112
MUSIC, 50, 52
    near field, 53
mutual impedance, 27, 32
mutual information, 93, 126

Nakagami, 16
Nakagami-m distribution, 16
narrowband, 8, 15, 17, 34, 45, 92
near field, 38, 40
    region, 55
Nehorai, 53
Newton, v
noise subspace, 52, 56
noise whitening, 35, 110
normal
    complex, 30, 34, 60, 69, 80, 81, 91, 110, 127
normal distribution, 15
    truncated, 26
normal equations, 57
normal process, 43
number of paths, 28, 40, 70, 71, 77, 104, 145, 147
number of samples, 73
Nyquist, 55
Nyquist sampling, 50, 55
Nørklit, iii, 4, 32

Oliphant, 4, 6
omni-directional field, 22
ordering
    positive semidefinite partial, 115
orthogonal basis functions, 21, 26
Ottersten, 80
Ovaska, 4
Owens, 28

pairwise error probability, 108, 112
Papoulis, 27, 37, 84, 87
parsimonious model, 8, 41
Parsons, 15, 16
Paulraj, 5, 28
PCLP, 57, 70, 92
PCS, 32
persymmetric matrix, 119, 120
perturbation theory, 88
phase-distortion, 1
Pickard, 2
Pierce, 6
Pierucci, 108
plane waves, 34
point sources, 70
polarisation, 9
polynomial, 56
    model, 61
    prediction, 70
Popper, 15
positive definite matrix, 115
power control, 4, 68
prediction, 4, 5, 43
    and correlation, 28
    concept, 2
    deterministic, 29
    filter, 57
    gain, 68
    length, 68, 74
    performance measures, 67, 74
Preece, 2
Proakis, 14, 16, 35, 107, 112
projection matrix, 45, 52
propagation
    multipath, 1, 9
propensity, 15
protocol, 6
pseudo-spectrum, 52, 55
Pätzold, 28

Qiu, 4

radiation exposure, 2
Raich, iii, v, 87

Rake receiver, 33
Raleigh, 4
Ranheim, 27
Rappaport, 6
Rayleigh, 15
    criterion, 88, 92
    distribution, 15, 28, 30, 81
    paper, 31
    parameter, 88
    quotient, 112, 119
real time, 9, 15, 33
Reddy, 5, 28
Redl, 4, 6
reflection, 1, 39
    multiple, 39
regression splines, 38
Reinsch, 83
residues, method of, 30
Ricean distribution, 15
Rife, 82
Rissanen, 64
RLS, 5
Roman, 28
rough surface scattering, 87–92, 146
Roy, 58
Rustako, Jr. , 28
Ryzhik, 17, 23, 55, 134

Salmon, 15
Salz, 22
Scaglione, 11, 81
scatterer distribution function, 18
scattering, 1, 91
    function, 12
    rough surface, 87–92, 146
    uncorrelated, 14
Schmidt, 52
Schwartz, 6, 29
Schwarz, 64
Scott, 27
Selker, 2
Ser, 113, 142
shadowing, 41
Shapiro, 2
Shawe-Taylor, 65
Shinozuka, 28
Sicuranza, 38
signal subspace, 52
Silverman, 52
singular value decomposition, 58
Slack, 28
slow wave, 27
snapshot, 50, 53, 146
SNR, 72
solid angle, 20
space loss, 9
space-time coding, 16
spatial correlation, 16–19, 21, 23, 24, 27–29, 90, 145
spatial frequency
    see Doppler frequency, 12
specular reflection, 88
    point of, 90
Spiegel, 98
Spizzichino, 88
Springer, 127
Staras, 6
Starer, 53
stationarity, 5, 14, 15
Steele, 4
steering matrix, 63
steering vector, 36, 47, 48, 54
Stein, 6, 29
Sternad, 8, 87, 147
Stevens, 38
stochastic model, 9, 15, 33
Stoica, 80
Stremler, 11
Stuart, 84, 136
subspace methods, 49

Suppes, 15
Suzuki, 16
Suzuki distribution, 16
Svantesson, 27
symbol alphabet, 108
symbol period, 6
synthetic array, 37, 43, 48–50, 63, 73, 79

Tan, 113, 142
Tanskanen, 4
Tarone, 2
Taylor series, 38
TDMA, 3, 55, 107
Teal, iii, 4, 32, 87
Tepedelenlioğlu, 28, 71
thermal noise, 44
Therrien, 11, 35, 50, 51
Thompson, Jr, 17, 19
time-dispersion, 1
Toeplitz, 112, 119
training sequence, 7, 126, 131
transfer function, 12, 14
transform
    Fourier, 12, 35, 38, 55, 64
    Laplace, 30
Tsatsanis, 5
Tufts, 56, 57, 92

uncorrelated noise, 44
uniform distribution, 81, 92
union, 114

Vandermonde matrix, 61
Vapnik, 65
Varey, 2
Vather, v
Vaughan, iii, v, 3, 4, 13, 24, 26, 27, 32, 62, 87
Viberg, 80
Vijayalaxmi, 2
visualisation, 122
Volterra series, 38
von Mises distribution, 23

Waterhouse, 17, 19
Wax, 65
Webb, 4
Weber, 4, 6
Weibull, 16
Weibull distribution, 16
Wesnes, 2
Wilcosky, 2
Wilkinson, 83
Williamson, iii, v, 49, 55, 73
Willink, 15
windowing, 6
Winters, 6, 22, 51
Wold, 57
Wold decomposition, 37, 57
worst case channel, 108
WSS, 14
WSSUS, 14

Yeh, 53

Zollinger, 28, 71