Signal-to-Noise Ratio and Signal-to-Impairment Ratio Estimation in AWGN and Wireless Channels

by

David Renato Pauluzzi

A thesis submitted to the Department of Electrical and Computer Engineering in conformity with the requirements for the degree of Master of Science (Engineering)

Queen's University
Kingston, Ontario, Canada
September, 1997

Abstract

As the need to support larger numbers of subscribers with higher-quality, wireless services increases, wireless technology must improve to keep up with the demand. Performance-enhancing techniques which historically have been too costly to implement are now gaining favour as digital hardware is becoming ever cheaper and more powerful. Signal-to-noise ratio (SNR) estimation and signal-to-impairment ratio (SIR) estimation algorithms are examples of techniques that may be used to improve receiver functions such as diversity combining and synchronization. Historically, the total signal-*plus*-noise power or the total signal-*plus*-impairment power often has been used in favour of the SNR or the SIR since estimators of the total power are easier to implement. Today, inexpensive, powerful digital hardware makes the implementation of SNR or SIR estimators practical so that the performance improvement effected by their use may be realized. In this thesis, various applications that may benefit from use of SNR or SIR estimation are identified.

SNR estimation is studied in the context of BPSK signaling in the real AWGN channel, and 8-PSK signaling in the complex AWGN channel. Several published SNR estimators are derived and adapted to the assumed system model. The performances of the SNR estimators are quantitatively compared by the analysis of SNR estimates produced by computer simulation, and a theoretical bound is derived for both the real and complex AWGN channels.

SIR estimation in wireless channels is studied by the statistical analysis of SIR estimates generated by the computer simulation of $\pi/4$-DQPSK-modulated signals in five "typical" mobile radio channels. A fading simulator that may be used to model fading processes with continuous Doppler and delay profiles is described and verified. Several SIR estimators are derived; some of the estimators are unmodified published algorithms, others are published algorithms adapted to the assumed system model, and others are original.

A few selected SIR estimators are used in the simulation of a postdetection, selec-

tion diversity combining application. The relative improvement in the performances of the various diversity combiner implementations due to the incorporation of the SIR estimators is compared. The simulated BER performances of "traditional" selection combiners using the total signal-plus-impairment power are provided for reference, and a simulated lower bound on the BER of postdetection selection diversity combining is also provided.

TO

My Parents.

Acknowledgments

I sincerely wish to thank my supervisor, Dr. Norman C. Beaulieu, for his guidance, support, and patience, particularly considering the challenges associated with supervising an off-campus student. Financial support is gratefully acknowledged from NSERC Grants STR0181058 and OGP0003986; the Telecommunications Research Institute of Ontario (TRIO); and the Ontario Ministry of Economic Development, Trade and Tourism: Technology and Training Development Ontario/Singapore Collaborative Research Grant. The staff of the Electrical and Computer Engineering and TRIO offices are thanked for their administrative help. Thanks also to Dr. David Nairn for kindly granting me access to the VLSI lab computer resources.

Thanks are given to R. Matzner of the Federal Armed Forces University Munich in Neubiberg, Germany; B. Shah of the Jet Propulsion Laboratory in Pasadena, California; P. Hoeher of the Institute for Communications Technology, German Aerospace Research Establishment, D-8031 Oberpfaffenhofen, Germany; and A. L. Brandão of the Department of Electronic and Electrical Engineering, The University of Leeds, Leeds, U.K. for their cooperation and for sharing their resources on this topic.

Thanks to MPR Teltech Ltd. (R.I.P.[1]) for granting me a leave of absence to pursue a master's degree, even though I was called back to work before I could complete my thesis. Thanks to B.C. Telecom, the former owners of the now defunct MPR Teltech Ltd., for divesting MPR and for laying off all of the members of the Rural Radio team (of which I was a part), thus providing me the opportunity to work on my thesis full-time with severance pay.

I would like to thank Dr. Andrew S. Wright and Mr. Laurence J. Wallace of DATUM Telegraphic, Inc. in Vancouver, B.C. for providing me with new employment, and for their patience and encouragement in the final months of my thesis.

Thanks to Aapool Biman and Chris Tan who were vital links to Queen's as I completed my thesis several thousands of kilometres away in B.C. I am also grateful

[1]Rest In Pieces.

for Chris' help with setting up my LaTeX text processing facilities on my PC.

Thanks to Glen Rempel for his technical insights and advice. I can empathize with your situation—completing your thesis part-time while working full-time!—I hope you celebrate a successful defence someday soon.

To my running partners (Chris Tan, Alex Seyoum, Joubin Karimi, and Leo Cordeau), work-out partners, and soccer and hockey teammates, I hope you always manage to balance work with a healthy lifestyle.

Best of luck to all of those I got to know in Kingston with whom I hope to stay in touch: Captain Aapoolçoyuz Biman, "Bouncing" Bo Tomiuk, Chris Tan, Dave "Pernacky" Paranchych, Joubin F. Karimi, Sasha Verma, Oguz Sunay, Radde Majeed, and the three "paesanos": Alex Seyoum, Andrew Parolin, and Gino Labinaz. A special note of appreciation goes to the MacDonald family who helped to make my stay in Kingston feel like home.

Thanks to all of my patient friends in B.C. who provided encouragement: Dave and Sun Mee Romalo, Raphael and Esther Smith, Bruce and Michiko Dow, Dave Neufeld, Jeff Chesko, Robert Weaver, Tony and Erika Palcich, Dave Mathur, Ali Motamedi, and Forrest Stokstad.

I also appreciate the encouragement from all of my former MPR colleagues, especially Bob Nicholson, Bob Wilson, Mario Vanin, Ron Scott, Dennis Rosenauer, Martin Katz, Lee Vishloff, Don Heywood, Hong Liang Zhang, Kartik Mahajan, Cuong Lai, Soon Shin, Bill Warner, Randy Chapman, Wayne Tressel, Amiee Chan, Paul Thiel, Steve Bennett, Elsa Willoner, Qiang Wang, and Charles Jeffrey.

I especially would like to thank my colleague and friend, Dave Romalo, who provided me with much technical assistance, advice on writing style, and moral support, and with whom I shared the experiences of the MPR divestiture and layoffs, and the subsequent search for employment.

Last, but certainly not least, I would like to thank my mother, father, sister and brother for their love, encouragement, and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Digital wireless technology has been incorporated into public telephony systems in many parts of the world, and it is poised to displace the analog cellular technology (AMPS) that has dominated the North American wireless telephone market for many years. The digital technology will deliver next to land-line voice quality and increased subscriber capacity. Every technical advantage is exploited to achieve the best bit error rate (BER) performance and the highest capacity with the least amount of power and expense. Complexity is less of an issue today than it has been in the past with the advent of cheap and powerful digital signal processor (DSP) and gate array devices, so that algorithms may be implemented in forms that are closer to ideal rather than simplified forms which typically exhibit inferior performance.

One example of a function that historically has been considered too costly to implement is the measurement of the signal-to-noise power ratio (SNR). Many algorithms in communications use the SNR as an input parameter but, in practice, the total signal-*plus*-noise power has been used instead since it is easier to measure. For wireless channels, BER performance depends not only on the SNR but also on other impairments such as fast, frequency-selective fading and cochannel interference (CCI) so that it is more appropriate to consider a signal-to-*impairment* power ratio (SIR).

1.2 Applications which may Benefit from Knowledge of the SNR or SIR

In many applications, the total received power is estimated for simplicity rather than the SNR or SIR. Goldsmith *et al* [3] discuss power measurement for time-varying cellular channels, and point out that real-time measurement of the received power is required for operations such as power control, handoff, and dynamic channel allocation. Examples of applications that use total power or received signal strength estimates are described in [4, 5, 6, 7]. In fact, the use of SNR or SIR estimates can improve the performances of the signal-strength-based algorithms used in these applications. Some references are listed below which describe applications that ideally require knowledge of the SNR or SIR.

1.2.1 Resource Management Algorithms

Measurement of the SIR is of great interest today as wireless service providers are finding that cochannel interference is the greatest factor limiting the extent to which cell sizes can be reduced in an effort to increase frequency re-use and system capacity [8, 9, 10, 11]. For this reason, methods to measure the SIR (where the impairment, in this case, is mainly CCI) are attracting much attention for use in resource management algorithms such as those used for handoff, dynamic channel allocation, and power control.

Handoff and Adaptive Dynamic Channel Allocation

The two concepts of handoff (or handover) and adaptive dynamic channel allocation (DCA) are closely related. Handoff refers to the event that occurs when an established link between a mobile unit and a base station is "seamlessly" broken and re-established, usually with another base station. A handoff event is typically caused by the physical movement of the mobile unit from one cell coverage area to another. Adaptive DCA refers to a dynamic channel allocation strategy that is adaptive both

to changing traffic conditions and interference levels.

Beck and Panzer [8] propose an algorithm called DYNINF which can be used for both handoff and DCA. The operation of the algorithm depends explicitly on the ability to measure the signal-to-interference level; however, no method to measure this quantity is offered. Hamabe *et al* [12] and Andersson *et al* [13] also present channel allocation strategies that require measurement of the SIR. In [13], the link quality is monitored by checking to see that known preamble bits are correctly detected by the receiver.

Other handoff and channel assignment schemes are discussed in [14, 15, 16, 17, 18].

Power Control

Zander [9] indicates that power control is important "to adjust the power of each transmitter for a given channel allocation such that the interference levels at the receiver locations are minimized." He points out that, in practice, power control algorithms typically keep the total received power at a constant level; however, he adds that keeping the signal-to-interference power ratio constant instead could improve system capacity. This view is supported by Jalali and Mermelstein in [19]. Zander admits that a practical implementation of the power control algorithm proposed in [9] would be difficult since the path gains of the desired signal and interferers are, in general, unknown. A means to estimate the SIR would facilitate the practical implementation of this power control algorithm.

Other relevant papers on power control include [20, 21, 22].

1.2.2 Diversity Combining

The classic pre-detection maximal-ratio combiner is described in many sources (see, for example, [23, 24, 25]). This combiner forms the weighted sum of two or more diversity branches where the weights are proportional to the amplitude of the signal, and inversely proportional to the noise variance. The weights for this diversity scheme

are often implemented using the signal-plus-noise envelope as opposed to the signal amplitude-to-noise variance ratio [2].

If an algorithm were available to estimate the signal and noise powers separately, the desired weightings for each of the branches of the maximal-ratio combiner could be formed trivially as the ratio of the square root of the signal power to the noise power. If the SNR is determined as an inseparable parameter, $\rho = S/N$, then the signal power and noise power could be computed by also estimating the total received power, $P = S + N$, so that simultaneous equations for ρ and P may be solved for S and N to yield $N = P/(1+\rho)$ and $S = P - N$. The branch weights are then formed trivially as $\sqrt{S}/N$.

In [26], Adachi presents an optimal postdetection diversity combiner that weights the differentially-detected symbols of each of the branches based on a formula which depends explicitly on both the SNR and the signal-to-interference ratio. The implementation of this formula actually requires two separate estimators—one to measure the SNR, the other to measure the signal-to-interference ratio.

A postdetection selection diversity combiner is described by Hladik, Chennakeshu, and Saulnier in [27] where the SIR of each diversity branch is measured on a symbol-by-symbol basis. Each symbol interval, the differentially-detected symbol corresponding to the branch with the largest instantaneous SIR is the one that is presented to the decision device. The specific SIR estimator is not given in any detail in [27], but is simply described as an approximation to the maximum-likelihood (ML) estimate of the SIR. It is highly probable that the SIR estimator used in [27] is the same as the SIR estimator described by Chennakeshu and Saulnier in Appendix II of [28].

1.2.3 Equalization

Balaban and Corrales [29] describe an equalizer for frequency-selective fading channels. The tap update algorithm of the equalizer requires both an estimate of the channel impulse response and an estimate of the SNR, thus illustrating another ap-

plication requiring some means to estimate the SNR. Note that channel impulse response estimation [30, 31] is a separate topic and is not considered in this thesis.

1.2.4 Synchronization

A maximum likelihood estimator of the bit timing is presented by Wintz and Luecke in equation (10) of [32] which is a function of the noise variance. For additive white Gaussian noise (AWGN), the noise variance drops from the estimator expression as shown by equation (11) of [32]. However, in time-varying channels, the noise variance cannot be assumed to be constant so it, or the SNR, must be estimated for optimal performance. Though a noise power estimator is required here, an SNR estimator could be used together with a total received power estimator to derive the noise power, as described in Section 1.2.2.

Chennakeshu and Saulnier [28] present a method to achieve timing and frequency synchronization by maximizing the SIR with respect to the timing and frequency offset. An approximation to a maximum likelihood SIR algorithm that can be used to implement this synchronization scheme is proposed in Appendix II of [28].

1.2.5 Adaptive Arrays

Adaptive arrays are used in wireless communications systems to cancel interference and mitigate fading effects by appropriately weighting and combining the output of two or more antennas. The optimal weight equation is given by Winters as (12) of [33] or, equivalently, (9) of [34], and is found to be a function of the noise variance. Again, using a technique such as that described in Section 1.2.2 an estimate of the noise power may be found from estimates of the SNR and the total received signal power.

1.2.6 Viterbi Equalization and Decoding

The path metric used in Viterbi equalization and decoding is shown by Hagenauer and Hoeher [35] to depend on what the authors call the "instantaneous" SNR, $E_s(k)/N_0$, where E_s is the energy per symbol, N_0 is the noise power spectral density, and k is the time index. The time dependence arises as a result of the time-varying nature of the multipath channel assumed in Section 3.2 of [35].

1.3 Literature Review

To the author's knowledge, no quantitative comparison of various SNR and SIR estimation techniques has ever been reported in the literature. Surveys of BER estimation techniques have been published by Scholz [36] and Newcombe and Pasupathy [37], but no quantitative comparisons of SNR or SIR estimators are provided. These BER estimation surveys include techniques such as those described by Hingorani and Chesler [38], Gooding [39], and Scholz, Cook, and Giles [40] that could be used to provide a relative measure of transmission quality. However, BER estimation techniques are not considered in this thesis as focus is placed on SNR and SIR estimation techniques that yield quantities not only useful as *relative* measures of quality, but also as *absolute* measures that can be used as input parameters in applications such as those described in Sections 1.2.2 to 1.2.6.

Interest in techniques to generate estimates of the SNR began in the mid- to late-1960's. The earliest recorded work on SNR estimation that could be found is a university report written by Nahi and Gagliardi in 1964 [41]. A subset of this work was published by Nahi and Gagliardi in 1967 [1]. The estimators described in [41, 1] form estimates of the SNR by measuring the power of a hard limited, received (noisy) signal at the output of a filter. Both the signal and noise are assumed to be Gaussian stochastic processes with correlation functions of known shape. An expression is given in [1] showing the output power as a function of the filter transfer function and the

SNR. The expression is not easily inverted so that, if the output power is known (measured), the SNR must be found implicitly using iterative techniques or a lookup table. It is not indicated in [1] whether or not this method is suitable for deterministic signals as well as stochastic signals.

Other early work on SNR estimation includes that of Benedict and Soong in 1967 [42]. The authors present three different methods to estimate separately the carrier strength and the noise level based on a finite number of samples. An ML estimator, an estimator based on amplitude moments, and an estimator based on square-law moments are presented along with plots of the bias and rms error of the simulated signal and noise estimates for the three estimation techniques. The SNR can be formed trivially as the ratio of the estimated carrier strength to the estimated noise level, but the performance of this parameter (which is of interest in this thesis) is not considered in [42]. Benedict and Soong refer to work done by Middleton published in 1962 [43] which predates that of Nahi and Gagliardi; however, the estimation method discussed in [43] assumes that the noise level is known and so is not applicable to this study.

The ML estimator derived by Benedict and Soong is complicated. The ML SNR estimation problem was formulated in a different manner by Kerr [44], Gagliardi and Thomas [45], and Gilchriest [46] to yield much simpler expressions. Kerr (July, 1966) proposes two different variations of a maximum likelihood SNR estimator where antipodal signaling in AWGN is assumed. In 1968, Gagliardi and Thomas [45] studied the ML estimator in more detail in a paper stemming from Thomas' PhD thesis [47]. The estimators derived by Kerr can be manipulated into the form of the SNR estimator derived by Gagliardi and Thomas. The pdf of the ML SNR estimator and analytical expressions for the bias and variance are offered in [45]. In the Jet Propulsion Laboratory report by Gilchriest dated June, 1966, a simple, intuitive SNR estimator is proposed based on the absolute mean and variance of an antipodal (BPSK) signal corrupted by AWGN. An analysis of the pdf of this estimator is

presented along with confidence intervals. This work was extended in 1967 by Layland [48] to study the performance of this SNR estimator at low levels of SNR. It is indicated in [45] that this intuitive SNR estimator is a type of ML estimator.

An analog method for determining the SNR of BPSK signals in AWGN was published by Edbauer [49] in 1977. The method is based on the processing of the in-phase and quadrature branches of a Costas demodulator. In this thesis, only discrete methods of SNR estimation are considered.

In more recent times (1986), Simon and Mileant [50] introduced an SNR estimator called the split symbol moments estimator (SSME) which is designed for BPSK signals in wideband AWGN channels. Shah and Hinedi [51] study the SSME in narrowband channels and provide plots of the means and inverse normalized variances of theoretical and simulated SSME estimates. In a Jet Propulsion Laboratory memo, Shah and Holmes [52] discuss a modification of the SSME designed to improve performance in narrowband channels. The channel models of [51] and [52] assume that all of the filtering occurs after noise is added which is different from the model developed in Section 2.2 of this thesis where filtering is assumed to be split between the transmitter and receiver.

In 1993, Matzner [53] presented an SNR estimator whose structure was first introduced by Benedict and Soong as the "square-law method" of carrier strength and noise level estimation. Matzner evaluates the performance of the SNR estimator (which is of interest in this thesis) as opposed to the performances of the separate estimators of carrier strength and noise level as treated in [42]. Matzner also provides more derivation details. The derivation assumes complex baseband signals in complex AWGN, but the estimator structures developed are also applicable, with relatively minor modifications, to real baseband signals in real AWGN. The mean square error (MSE) of the logarithm (dB) of simulated SNR estimates is plotted in [53] as a function of the block length and as a function of the true SNR. This estimator is derived using a different approach by Matzner and Engleberger [54], and a hardware imple-

mentation is described by Matzner, Engleberger, and Siewert [55]. The complex form of this SNR estimator may be modified to be used as a more general SIR estimator in fading channels with CCI and AWGN.

The "signal-to-variation" ratio (SVR) estimator proposed by Brandão [56] is an SIR estimator used to measure the quality of M-ary PSK signals in channels corrupted by multipath, CCI, and AWGN. This estimator may be modified to be used as an SNR estimator for complex signals in complex AWGN, or for real signals in real AWGN. The SVR estimator is identified as being of the "in-service" type which is a term sometimes used to refer to an estimator that forms estimates from the information-bearing received signal, thus avoiding the need to perform SIR or SNR measurements off-line. Plots are provided in [56] showing the theoretical and simulated SVR estimates as a function of signal power for three different fading channels.

Yoshida, Tan *et al* [57] and Yoshida, Hirai *et al* [58] describe an in-service estimator that, like the SVR estimator, also reflects the multipath spread and level of CCI in a wireless channel. This estimator can be used with any QPSK-like signal. The mean simulated SIR is plotted as a function of the delay spread of the channel, and as a function of the CCI level. Also included is a plot showing the correlation between the measured BER of the channel and the mean estimated SIR. This estimator can be used to measure SNR in complex AWGN but, as presented in [58], the actual estimated values do not correspond in an absolute sense to the true SNR of the channel; however, it is shown in Section 3.7.3 of this thesis how the estimator can be modified to yield true estimates of the SNR. Note that this estimator cannot be modified to operate with real signals in real AWGN since the algorithm requires a signal with both in-phase and quadrature components.

In Appendix II of [28], Chennakeshu and Saulnier derive an ML SIR estimator based on the pdf of the phases of differentially-detected samples. The estimator is specifically referred to as a "signal-to-impairment ratio" estimator and its application to timing and frequency synchronization is described. Simulation results of the

BER resulting from the incorporation of the SIR estimator into this synchronization application are provided.

Other recent examples of estimators designed to measure the SIR in wireless channels are presented by Andersin *et al* [59] and Austin and Stüber [60].

1.4 Contributions of the Thesis

The purpose of the thesis is to compare quantitatively various SNR and SIR estimation techniques in common channels using common performance metrics. Some performance results have been published in the literature, but the assumed channel conditions and the performance metrics are not consistent from one source to the next making a quantitative comparison difficult. Some of the estimators under study have been collected from the literature and are used without modification, others are modifications of published estimators that have been modified for the assumed channel conditions, and others are original. The specific contributions of the thesis are itemized below.

1. The SSME for real, wideband channels is derived and compared to the expression given in [51].

2. Two modified SSME algorithms are developed based on an approach similar to that described in [52] in attempts to improve the performance of the original SSME algorithm in real, narrowband channels.

3. The ML SNR estimator tailored to the real AWGN channel of Chapter 2 is derived and compared to the ML SNR estimator presented in [45]. The ML SNR estimator operates on the samples at the input to the matched filter in the receiver, and requires the transmitted data sequence to form SNR estimates. Two ML structures are identified: one that uses known transmitted data, and another that uses receiver decisions. Reduced-bias forms of these estimators are also provided.

4. The ML SNR estimator tailored to the complex AWGN channel of Chapter 2 is derived. This is a new contribution as the derivation for complex channels does not appear elsewhere in the literature. This estimator also operates on the samples at the input to the matched filter in the receiver and, as in the real case, two ML structures are identified: one that uses known transmitted data, and another that uses receiver decisions. Reduced-bias forms of these estimators are offered.

5. The intuitive SNR estimator of [46] is derived for real signals in real AWGN and is shown to be an ML SNR estimator for the baud-spaced samples after the matched filter in the receiver. One estimator structure using known transmitted data and another using receiver decisions are identified. Reduced-bias forms of these estimators are offered.

6. An intuitive SNR estimator for complex channels similar to the intuitive SNR estimator of [46] for real channels is derived using a modified version of the ML SNR estimator for complex channels. This SNR estimator structure does not appear elsewhere in the literature. Biased and reduced-bias forms of estimators using known transmitted data and using receiver decisions are presented.

7. The SNR estimator of [53] is derived for complex channels and compared to the published expression. A corresponding estimator for real channels is also derived. The derivation details for the real case do not appear elsewhere in the literature.

8. SNR estimator structures are derived for real and complex AWGN channels applying the approach used in [56] for the derivation of the SVR estimator. The resulting SNR estimator for complex channels is very similar to that presented in [56], but the SNR estimator for real channels does not appear elsewhere in the literature.

9. Cramér-Rao bounds for real and complex AWGN channels are derived using the approach described in [47] for real channels. The Cramér-Rao bound (CRB) for complex channels does not appear elsewhere in the literature.

10. The biases, variances, and MSE's of the simulated SNR estimates generated by the various SNR estimators under study are presented for BPSK signals in real AWGN, and 8-PSK signals in complex AWGN. The variance and MSE results are compared to the appropriate CRB's. A discussion of results is provided and strengths and weaknesses of the various SNR estimators are identified.

11. Issues to consider regarding the hardware implementation of an SNR estimator are discussed.

12. A brief survey of potential fading simulators which could be used to model the assumed channel characteristics is provided.

13. The estimator described in [53] is modified for use as an SIR estimator in general, wireless channels.

14. The SVR estimator of [56] is derived for use as an SIR estimator and compared with the published SVR expression.

15. The SIR estimator of [57, 58] is stated in a form consistent with the assumed channel environment of Chapter 3. In a complex AWGN channel, the estimator given in [57, 58] produces estimates that differ in an absolute sense from the true SNR of the channel. A modified form of this estimator is offered which generates estimates that correspond to the true SNR.

16. An ML SNR estimator for samples at the input to the matched filter is described as an *ad hoc* SIR estimator given the assumed system model of Chapter 3. Strengths and weaknesses of two forms of this estimator are identified.

17. The derivation of an ML SIR estimator presented in [28] based on the pdf of the phase of samples at the output of a differential detector is sketched. Two forms of this estimator are identified: one that uses known transmitted data, and another that uses receiver decisions.

18. The Pseudo-ML (PML) SIR estimator, the Signal-to-Impairment Variance (SIV) SIR estimator, and the Modified Signal-to-Impairment Variance (MSIV) SIR estimator are derived. All of these algorithms are original.

19. The statistical properties of the simulated SIR estimates generated by the various SIR estimators under study are presented for five different representative wireless channels.

20. A description of a postdetection selection diversity combiner is provided incorporating a few selected SIR estimators from Chapter 3. BER performances of diversity receivers employing SIR estimators and diversity receivers employing conventional received signal strength estimators are compared.

1.5 Thesis Outline

The main body of the thesis is broadly divided into three chapters. Chapter 2 is devoted to SNR estimation in the AWGN channel, SIR estimation in wireless channels is discussed in Chapter 3, and the application of SIR estimation to postdetection selection diversity combining is treated in Chapter 4. The problem of SNR estimation in AWGN merits a chapter all on its own since this is a classical problem which is well defined. The relatively simple Gaussian statistics of this channel often lead to tractable problems, and classical estimation techniques (such as the method of maximum likelihood) yield closed-form solutions.

In Chapter 2, the problem of SNR estimation in the AWGN channel and the concept of SNR is defined, the system model is described, the SNR estimators under

study are presented, measures of SNR estimator performance are given, the CRB's for real and complex channels are derived, simulation results are offered and compared to theory (where appropriate), and a discussion of results concludes the chapter.

The problem of SIR estimation in wireless channels[1] is discussed in Chapter 3 and compared to the problem of SNR estimation in the AWGN channel. Environment assumptions are stated and various potential fading simulators are identified to model the assumed channel characteristics. Details of the implementation of the chosen fading simulator are provided and statistical verification results are presented. The general system model is described and five specific, representative mobile radio test channels are identified in which the SIR estimators are tested. The SIR estimators under study are then described, measures of SIR estimator performance are given, simulation results are offered and compared, and a discussion of results concludes the chapter.

In Chapter 4, the postdetection selection diversity combining application is described, the BER of the software implementation of the combiner is compared to published data, simulation results are presented, and a discussion of results concludes the chapter.

Chapter 5 concludes the thesis with summary remarks and suggestions for further work.

[1] "Wireless channels" include, for example, high-capacity, land microwave digital radio links; low-data rate, point-to-point links; satellite/earth station links; indoor channels; and mobile radio channels. In Chapter 3, the mobile radio channel is chosen for study as all of the channel impairments typically observed in wireless channels are present, under certain conditions, in the widely-variable mobile radio channel. In the sequel, the terms "wireless channels" and "mobile radio channels" are used interchangeably.

Chapter 2

SNR Estimation in the AWGN Channel

2.1 Problem Definition

The goal is to find the "best" estimate of the signal-to-noise power ratio (SNR) in a digital receiver with the least cost. The SNR of interest is the ratio of discrete (sampled) signal power to discrete noise power at optimal sampling instants at the input to the decision device in the receiver. The estimators under consideration generate estimates by averaging observable properties of the received signal over many symbols.

The SNR of the discrete, optimally-sampled signal at the input to the decision device in the receiver is not to be confused with the analog SNR after the matched filter (MF) (in a system in which an analog MF is employed). For an analog signal at the output of the MF, an expression for the SNR is [61]

$$\mathrm{SNR}_{\mathrm{analog}} = \frac{E_s}{N_0}\frac{R}{B_{\mathrm{eff}}}$$

where E_s is the energy per symbol, N_0 is the noise power spectral density (PSD), B_{eff} is the effective noise bandwidth of the receiver filter, and R is the baud rate. For the discrete, binary signal in real AWGN (both signal and noise are one-dimensional),

the discrete SNR is (see [62] and (2.21) of Section 2.2)

$$\text{SNR}_{\text{discrete,real}} \leq 2\frac{E_s}{N_0} = 2\frac{E_b}{N_0}$$

where equality holds if a MF receiver is employed, and the energy per symbol, E_s, is equal to the energy per bit, E_b, for discrete, binary signaling. For a discrete signal constructed from a set of $M = 2^k$ k-bit symbols in complex AWGN (both signal and noise are two-dimensional), the discrete SNR is (see (2.19) of Section 2.2)

$$\text{SNR}_{\text{discrete,complex}} \leq \frac{E_s}{N_0}$$

where, again, equality holds if a MF is employed. Note the factor of two difference between $\text{SNR}_{\text{discrete,real}}$ and $\text{SNR}_{\text{discrete,complex}}$.

Note that both E_s (or E_b) and N_0 are analog quantities typically measured at the input to the receiver, but E_s/N_0, as a measure of discrete SNR, is best understood as a property of the sampled output of the MF. In the sequel, any reference to SNR implies $\text{SNR}_{\text{discrete,real}}$ or $\text{SNR}_{\text{discrete,complex}}$. Also, ρ is a variable often used here to represent the SNR (real or complex).

A few general SNR estimation strategies are introduced in the following sections.

2.1.1 SNR Estimation by the Physical Separation of Signal and Noise

Assume it is possible to isolate the message portion and noise portion of a signal corrupted by additive white Gaussian noise (AWGN). This isolation may be achieved by separation in time, frequency, or space. For example, space isolation may be achieved in the laboratory by connecting one of either a "clean" signal or the output of a noise generator to the input of a device such as a receiver or power meter. Let the sampled output of the MF be denoted by s_n when only the clean signal is presented to the receiver, and z_n when noise is the only input. The discrete SNR in this case is simply

$$\rho = \frac{E\{s_n^2\}}{E\{z_n^2\}} = \frac{S}{N}$$

where $S = E\{s_n^2\}$ is the average discrete signal power, $N = E\{z_n^2\}$ is the average discrete noise power, and $E\{\cdot\}$ denotes expectation.

This method obviously suffers from the fact that it is impossible to isolate the desired signal and noise except under very artificial, controlled conditions (for example, in the lab). Also, if the desired signal and noise are combined some time after the separate measurement of the discrete signal and noise powers, the claim that the SNR of the combined signal is still S/N is true only if both the signal power and noise power do not change with time.

2.1.2 SNR Estimation using a Reference Channel

A somewhat more practical approach is to consider two channels, one carrying a desired signal, s_n, plus noise, z_{n_1}, the other carrying only noise, z_{n_2}. As before, $S = E\{s_n^2\}$ and it is assumed that the noise power in both channels is equal so that $N = E\{z_{n_1}^2\} = E\{z_{n_2}^2\}$. Further, the signal and noise processes are assumed to have zero mean and are mutually independent.

Let the sampled output of the receiver in the first channel be denoted by r_n so that

$$r_n = s_n + z_{n_1},$$

and

$$E\{r_n^2\} = E\{s_n^2\} + 2E\{s_n\}E\{z_{n_1}\} + E\{z_{n_1}^2\} = S + N$$

where use is made of the assumed first- and second-order properties of the signal and noise. Taking the ratio of the discrete power of the two channels, one obtains

$$\frac{E\{r_n^2\}}{E\{z_{n_1}^2\}} = \frac{S+N}{N} = \rho + 1$$

so that

$$\rho = \frac{E\{r_n^2\}}{E\{z_{n_1}^2\}} - 1.$$

The drawback of this approach is that a channel isolated by time, frequency, or space must be set aside on which no data can be sent during the SNR estimation. This

resource overhead implies a throughput penalty since time, bandwidth, or space must be reserved as a reference channel in which no information may be sent. Note that it is assumed that the noise power in the reference channel and the noise power in the information channel are equal for all time.

2.1.3 Data-Aided SNR Estimation

SNR estimators which extract the SNR from a *single* channel on which both a desired signal and noise are present using knowledge of the transmitted message sequence are referred to as data-aided (DA) estimators. A DA SNR estimator performs best when the message sequence used by the estimator is identical to the true transmitted message sequence.

An example of this type of SNR estimator may be implemented using the correlation between the noisy signal and the known transmitted signal as suggested in [63, 64] and expressed in general terms as

$$\rho = \frac{\varphi^2}{1 - \varphi^2} \tag{2.1}$$

where φ is the correlation between the noisy received samples and known transmitted signal sequence. Equation (2.1) is an idealized form of a DA SNR estimator.

Various specific DA SNR estimators are discussed in more detail in later sections, but there are two general types classified according to whether the data used to aid the SNR estimation is known or estimated. An estimator that uses an exact, known copy of the transmitted message sequence will be referred to as a *TxDA* estimator. A DA SNR estimator that uses an *estimate* of the transmitted message sequence provided by receiver decisions will be referred to as an *RxDA* estimator. As a further classification, any SNR estimator that can generate SNR estimates from the unknown, information-bearing portion of the received signal is often referred to as an *in-service* estimator [56, 58]. RxDA SNR estimators are of the in-service type.

In TxDA SNR estimation, the fidelity of the message sequence used for SNR estimation is assured by making an exact copy of the transmitted message sequence

available to the receiver. As an example, short blocks of known data may be inserted periodically into the information-bearing sequence. DA equalization techniques [65] use so-called *training sequences* for a similar purpose. A throughput penalty is incurred since some channel capacity must be devoted to the transmission of non-information-bearing data (training sequences are not considered to carry information because the data is already known to the receiver). However, in systems which already employ training sequences for equalization or synchronization, there would be no *additional* throughput penalty since those same known sequences could be used to maximize the performance of a DA SNR estimator. Note that since TxDA SNR estimates can only be generated when known data is available at the receiver, the use of a TxDA estimator may not be appropriate in some situations where a continuous stream of SNR estimates is required.

Since receiver decisions are subject to error, the performance of RxDA SNR estimation is inferior to that of TxDA SNR estimation at low SNR where decision errors are more likely. An advantage of RxDA SNR estimation is that the SNR information is extracted directly from the information-bearing signal with no loss of throughput due to resource overhead. Since RxDA estimates may be generated whether the transmitted symbols are known or unknown, RxDA SNR estimators may be used in applications that require a continuous stream of SNR estimates.

2.1.4 SNR Estimation based on Statistics of the Received Signal

This class of SNR estimators generates estimates of the SNR assuming knowledge only of the statistics of the signal and channel. These estimation techniques are usually moment-based methods. Since no knowledge of the transmitted symbols is required, these techniques can derive SNR estimates from the information-bearing portion of the received signal and so are classified as in-service estimators.

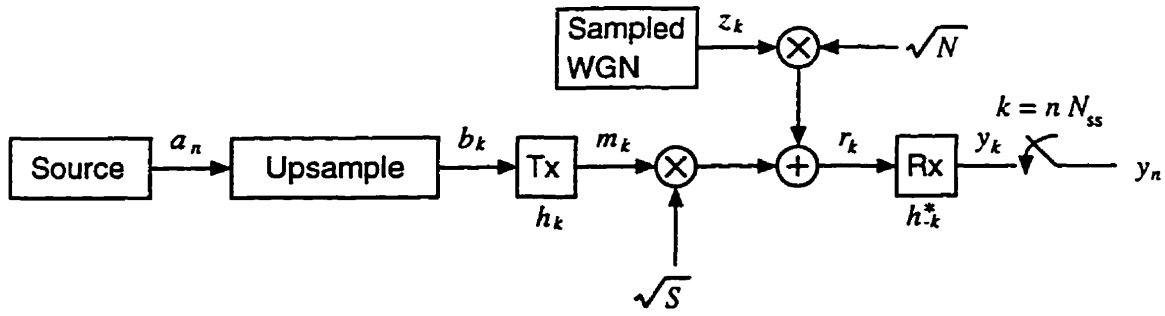

Figure 2.1: Discrete, baseband-equivalent, bandlimited model of coherent PSK in AWGN.

2.2 System Model

The search for the "best" in-service estimator is carried out by a performance comparison of several published techniques. The performance of each estimator is evaluated by the software simulation of a discrete, PSK signal in an AWGN channel.

Figure 2.1 illustrates a discrete, baseband-equivalent, bandlimited model of coherent PSK in AWGN. The model, as illustrated, applies equally well to coherent M-ary PSK in complex AWGN as to coherent BPSK in real AWGN. In the development that follows, the general complex case is assumed, but the results are equally applicable or easily adapted to the case of real signals. Perfect carrier and symbol timing recovery are assumed throughout.

A block of N_{sym} M-ary source symbols is upsampled to $N_{\text{ss}} = 16$ samples per symbol, shaped by a root raised-cosine (RRC) filter (with roll-off, $\alpha = 0.5$, and $L = 127$ tap coefficients), scaled by a constant attenuation factor, and corrupted by sampled, complex AWGN. The sequence of M-ary source symbols is represented by

$$a_n = e^{j\theta_n},\ n \in \{0, 1, \ldots, N_{\text{sym}} - 1\} \tag{2.2}$$

where θ_n is one of M phases spaced evenly around the unit circle. In the real, binary case, $a_n = \pm 1$. The upsampled, M-ary message sequence is

$$b_k = \sum_n a_n \delta_{k,nN_{ss}}$$

where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases}$$

is the Kronecker delta. The sampled, pulse-shaped, information-carrying signal is

$$\begin{aligned} m_k &= b_k \otimes h_k \\ &= \sum_n a_n h_{k-nN_{ss}} \end{aligned} \tag{2.3}$$

where h_k, $k \in \{-(L-1)/2, \ldots, -1, 0, 1, \ldots, (L-1)/2\}$ represents the RRC filter tap coefficients[1] for $|k| > (L-1)/2$, and $\otimes$ denotes discrete convolution. The RRC coefficients are derived from the analog transfer function of the full, raised-cosine filter using the frequency sampling technique of FIR digital filter design. A sketch of the derivation of the RRC coefficients is provided in Appendix B.

The signal presented to the receiver is

$$r_k = \sqrt{S}\, m_k + \sqrt{N}\, z_k \tag{2.4}$$

where z_k is sampled, zero-mean, complex AWGN of variance, σ_z^2, S is a signal power scale factor, and N is a noise power scale factor. The samples of the received signal after the MF can be expressed as

$$\begin{aligned} y_k &= (\sqrt{S}\, m_k + \sqrt{N}\, z_k) \otimes h^*_{-k} \\ &= \sqrt{S} \sum_l h_l\, m_{k-l} + \sqrt{N} \sum_l h_l\, z_{k-l} \end{aligned} \tag{2.5}$$

where $\otimes$ denotes discrete convolution, * in the exponent denotes the complex conjugate, and $h^*_{-k} = h_k$ because the RRC impulse response is real and even. Finally, the optimally-sampled, intersymbol-interference-free output of the MF (the decision variable) is

$$y_n = y_k|_{k=nN_{ss}} = \sqrt{S}\, a_n g_0 + \sqrt{N}\, w_n \tag{2.6}$$

[1]The range of k specified above implies a non-causal filter which is chosen for mathematical convenience. Though a non-causal filter is not realizable, this does not invalidate its use for analysis purposes since the only difference between a causal and a non-causal filter (having the same finite number of taps) is a shift in time. An odd number of filter taps is chosen for the analysis but an odd or even number of taps may be implemented affecting only the sampling time at the output of the MF in the receiver.

where g_0 is the peak of the full raised-cosine impulse response given by

$$g_k = h_k \otimes h^*_{-k} = \sum_l h_l h_{k-l},$$

and

$$w_n = z_k \otimes h^*_{-k}\Big|_{k=nN_{ss}} = \sum_l h_l z_{k-l}\Big|_{k=nN_{ss}}$$

represents the filtered AWGN at symbol-spaced sampling instants.

The first- and second-order statistics of the desired signal, AWGN, and filtered AWGN are expressed as

$$\begin{aligned}
E\{a_n\} &= 0 && (2.7a)\\
E\{a_{n+m}a^*_n\} &= \sigma_a^2\delta_{m,0} && (2.7b)\\
\mathrm{Var}\{a_n\} &= E\{|a_n|^2\} - (E\{a_n\})^2 = \sigma_a^2 && (2.7c)
\end{aligned}$$

$$\begin{aligned}
E\{z_k\} &= 0 && (2.8a)\\
E\{z_{k+l}z^*_k\} &= \sigma_z^2\delta_{l,0} && (2.8b)\\
\mathrm{Var}\{z_k\} &= E\{|z_k|^2\} - (E\{z_k\})^2 = \sigma_z^2 && (2.8c)
\end{aligned}$$

$$\begin{aligned}
E\{w_k\} &= \sum_i h_i E\{z_{k-i}\} = 0 && (2.9a)\\
E\{w_{k+l}w^*_k\} &= E\left\{\sum_i h_i z_{k+l-i}\sum_j h_j z^*_{k-j}\right\}\\
&= \sum_i\sum_j h_i h_j E\{z_{k+l-i}z^*_{k-j}\}\\
&= \sigma_z^2\sum_i h_i h_{i-l}\\
&= \sigma_z^2\sum_i h_i h_{l-i} = \sigma_z^2 g_l && (2.9b)\\
\mathrm{Var}\{w_k\} &= E\{|w_k|^2\} - (E\{w_k\})^2\\
&= \sigma_w^2\\
&= \sigma_z^2 g_0 && (2.9c)
\end{aligned}$$

where use has been made of the facts that

$$E\{z_{k+l-i}z^*_{k-j}\} = \sigma_z^2\delta_{j,i-l}$$

and

$$h_{j-l} = h_{l-j}$$

due to the symmetry of the RRC filter coefficients about the peak of the impulse response at $k = 0$. $\mathrm{Var}\{\cdot\}$ denotes variance. Equation (2.9b) shows that the auto-correlation of the noise taken at optimal sampling instants after the MF has a shape described by the impulse response of the full, raised-cosine filter which crosses zero every N_{ss} samples (the baud period). The implication is that the noise sampled at the baud rate after the MF is white. Also, from (2.7b), (2.8b), and (2.9b) it can be seen that the signal, AWGN, and filtered AWGN processes are wide sense stationary.

Further, the desired signal and noise are assumed independent (and thus uncorrelated) so that

$$E\{a^*_n z_k\} = E\{a^*_n\}E\{z_k\} = 0 \tag{2.10}$$

and

$$E\{a^*_n w_n\} = E\{a^*_n\}E\{w_n\} = 0. \tag{2.11}$$

Since y_n is the decision variable, the ratio of the discrete signal power component of y_n to the discrete noise power component of y_n is the SNR. That is,

$$\rho = \frac{E\{|\sqrt{S}a_n g_0|^2\}}{\mathrm{Var}\{\sqrt{N}w_n\}} = \frac{S\sigma_a^2 g_0^2}{N\sigma_w^2}. \tag{2.12}$$

The bit error probability may be stated in terms of this quantity. As an example, the bit error probability for coherent BPSK in real AWGN is

$$P_{b,\mathrm{BPSK}} = \mathrm{Q}(\sqrt{\rho}) = \mathrm{Q}\left(\frac{\sqrt{S}\sigma_a g_0}{\sqrt{N}\sigma_w}\right) \tag{2.13}$$

where

$$\mathrm{Q}(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\lambda^2/2}\,d\lambda$$

is the Q-function [62].

It will now be shown that, given the MF system model presented here, $\rho = E_s/N_0$ for complex signals in complex AWGN, and $\rho = 2E_s/N_0 = 2E_b/N_0$ for real, antipodal signals in real AWGN. The derivation of the symbol energy below assumes complex signals but the results are equally applicable to real signals. The main difference between the complex and real cases is in the treatment of the noise PSD which is explained later in this section.

The energy per symbol, E_s, is an analog quantity and can be expressed as [62]

$$E_s = P_{\text{avg,signal}} T \tag{2.14}$$

where T is the baud period and $P_{\text{avg,signal}}$ is the average (analog) desired signal power at the input to the receiver which may be found from the autocorrelation of the desired signal, $\sqrt{S}m_k$. It is convenient to derive the autocorrelation of m_k by an extension of the autocorrelation for continuous signals [65] to the discrete case.

Consider, for the moment, a continuous signal, $s(t)$, of the form

$$s(t) = \sum_{n=-\infty}^{\infty} a_n h(t - nT)$$

where the sequence of information symbols is denoted by a_n, as before, and $h(t)$ is the impulse response of some arbitrary pulse-shaping filter (note that $s(t)$ is analogous to the discrete signal, m_k). The autocorrelation of $s(t)$, $R_s(t+\tau, t)$, is evaluated as given in [66, page 298] as

$$\begin{aligned}
R_s(t+\tau, t) &= E\{s(t+\tau)s^*(t)\} \\
&= E_a\{s(t+\tau)s^*(t)\} \\
&= \sum_m \sum_n E_a\{a_m a_n^*\} h(t+\tau-mT)h(t-nT) \\
&= \sigma_a^2 \sum_m \sum_n \delta_{mn} h(t+\tau-mT)h(t-nT) \\
&= \sigma_a^2 \sum_n h(t+\tau-nT)h(t-nT)
\end{aligned}$$

where the expected value, $E_a\{\cdot\}$ is performed with respect to the information sequence, a_n.

The autocorrelation, $R_s(t+\tau,t)$, is a function of both t and τ so it is not stationary. It is, however, cyclostationary so that

$$R_s(t+\tau+T,t+T) = R_s(t+\tau,t).$$

The t dependence may be removed by taking the expected value of $R_s(t+\tau,t)$ with respect to the start time, t. Assume a uniform probability density function (pdf), $p_t(\lambda)$, for t such that

$$p_t(\lambda) = \begin{cases} \frac{1}{T} & 0 \le \lambda < T \\ 0 & \text{otherwise.} \end{cases}$$

Taking the expected value with respect to t one obtains

$$E_t\{R_s(t+\tau,t)\} = \sigma_a^2 \sum_n \frac{1}{T} \int_0^T h(\lambda+\tau-nT)h(\lambda-nT)\,d\lambda.$$

After some manipulation, one finds that

$$\begin{aligned} R_s(\tau) &= E_t\{R_s(t+\tau,t)\} \\ &= \frac{\sigma_a^2}{T} \int_{-\infty}^{\infty} h(\lambda)h(\lambda-\tau)\,d\lambda \\ &= \frac{\sigma_a^2}{T} h(\tau) * h(-\tau) \\ &= \frac{\sigma_a^2}{T} g(\tau) \end{aligned}$$

where $*$ denotes continuous-time convolution, and $h(\tau) * h(-\tau) = g(\tau)$ because $h(\tau)$ is even. By analogy with the continuous case just derived, the autocorrelation, $R_m(l)$, of m_k can be written as

$$\begin{aligned} R_m(l) &= \frac{\sigma_a^2}{N_{ss}T_S} \sum_\alpha h_\alpha h_{\alpha-l} T_S \\ &= \frac{\sigma_a^2}{N_{ss}} \sum_\alpha h_\alpha h_{l-\alpha} \\ &= \frac{\sigma_a^2}{N_{ss}} g_l. \end{aligned} \tag{2.15}$$

where $T = N_{ss}T_S$, and T_S is the sampling period. The average power of m_k may be stated simply as

$$E\{|m_k|^2\} = R_m(0) = \frac{\sigma_a^2}{N_{ss}} g_0, \tag{2.16}$$

and so (2.14) in terms of discrete signals becomes

$$\begin{aligned} E_s &= SR_m(0)N_{ss}T_S \\ &= \frac{S\sigma_a^2}{N_{ss}} g_0 N_{ss} T_S \\ &= S\sigma_a^2 g_0 T_S. \end{aligned} \tag{2.17}$$

Next, the noise PSD is related to the variance of the sampled noise. For this derivation, the noise PSD for the complex case is different from that of the real case by a factor of two (see Appendix A). Specifically, for the complex case it is N_0 (see Figure A.1(b) and (A.10) of Appendix A), and for the real case it is $N_0/2$ (see Figure A.1(c) and (A.29) of Appendix A). The complex case is treated first.

Sampled WGN may be conceptualized as continuous WGN filtered by an ideal lowpass filter centred at DC having a cutoff frequency of $\pm 1/2T_S$ [51]. Given the complex noise PSD of magnitude N_0, the variance of the sampled noise (the noise power) may be expressed as

$$\mathrm{Var}\{\sqrt{N} z_n\} = N\sigma_z^2 = \frac{N_0}{T_S}. \tag{2.18}$$

Putting together equations (2.9c), (2.12), (2.17), and (2.18) yields

$$\begin{aligned} \rho &= \frac{E\left\{\left|\sqrt{S} a_n g_0\right|^2\right\}}{\mathrm{Var}\{\sqrt{N} w_n\}} \\ &= \frac{S\sigma_a^2 g_0^2}{N\sigma_w^2} \\ &= \frac{S\sigma_a^2 g_0^2}{N\sigma_z^2 g_0} \\ &= \frac{S\sigma_a^2 g_0}{N\sigma_z^2} \\ &= \frac{S\sigma_a^2 (E_s / S\sigma_a^2 T_S)}{N(N_0/NT_S)} \\ &= \frac{E_s}{N_0}. \end{aligned} \tag{2.19}$$

In a similar manner, for the real case, the noise power may be expressed in terms of

the real noise PSD of magnitude $N_0/2$ as

$$\mathrm{Var}\{\sqrt{N}z_n\} = N\sigma_z^2 = \frac{N_0}{2T_S}. \tag{2.20}$$

Putting together equations (2.9c), (2.12), (2.17), and (2.20) yields

$$\begin{aligned}\rho &= \frac{S\sigma_a^2 g_0^2}{N\sigma_w^2} \\ &= \frac{S\sigma_a^2(E_s/S\sigma_a^2 T_S)}{N(N_0/2NT_S)} \\ &= 2\frac{E_s}{N_0} \\ &= 2\frac{E_b}{N_0}.\end{aligned} \tag{2.21}$$

From (2.2), it may be seen that $|a_n| = 1$ by design so that $\sigma_a^2 = 1$. If the noise variance is normalized such that $\sigma_z^2 = 1$, and if the RRC coefficients are scaled such that $g_0 = \sum_k h_k^2 = 1$, the decision variable (2.6) reduces to

$$y_n = \sqrt{S}\, a_n + \sqrt{N}\, w_n, \tag{2.22}$$

and the discrete SNR of (2.12) reduces to

$$\rho = \frac{S}{N}$$

so that the SNR may be set solely by the appropriate selection of S and N.

2.3 SNR Estimators under Study

All of the SNR estimation techniques under study estimate the discrete SNR by averaging one or more observable properties of the received signal over a number of received symbols. The SNR estimators under consideration fall into one of the two categories described in Sections 2.1.3 and 2.1.4.

Estimator algorithms for both real and complex channels are provided for each of the SNR estimators considered in this study except for the SSME's which are designed to operate exclusively with BPSK signals in real AWGN (see Section 2.3.1).

Where possible, estimator names and acronyms are those used by the original authors; otherwise, names and acronyms are arbitrarily assigned for ease of reference. All but one of the estimators are applied to the bandlimited channel model of PSK signaling in AWGN described in the previous section—the $SSME_0$ is the only algorithm designed specifically for wideband channels so it is treated separately.

2.3.1 The Split-Symbol Moments Estimator (SSME)

The SSME is based on the first- and second-order moments of the received signal and was developed at JPL in Pasadena, California for use in Deep Space Network receivers designed by NASA [51]. The original SSME algorithm, as formulated by Simon and Mileant [50], assumes a wideband channel since it was designed for low data rate transmission in the deep space channel where bandlimiting effects are minimal. Shah and Hinedi [51] studied the performance of the original SSME in narrowband channels. A modified SSME algorithm, presented in an unpublished, internal JPL memo [52], was developed to improve upon the performance of the original SSME in bandlimited channels.

Three different SSME algorithms are considered here. They are referred to as the $SSME_0$, $SSME_1$, and $SSME_2$ for ease of reference; their definitions are:

1. $SSME_0$: the original algorithm [50] for wideband channels.

2. $SSME_1$: a modification of $SSME_0$ for operation in narrowband channels motivated by the modified SSME algorithm for narrowband channels presented in [52].

3. $SSME_2$: an alternate SSME algorithm for narrowband channels proposed here in an attempt to improve upon the narrowband performance of the $SSME_1$.

All three of the SSME algorithms investigated use all of the N_{ss} samples per symbol available to the receiver. This is in contrast to SNR estimation techniques discussed in

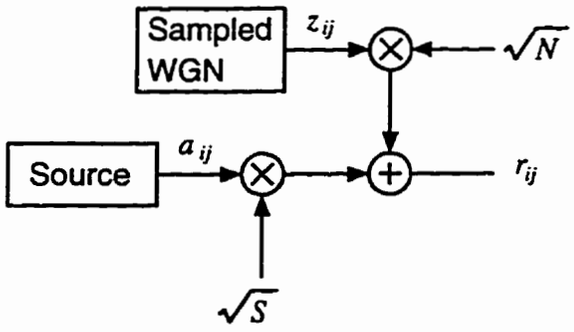

Figure 2.2: Baseband-equivalent model of wideband BPSK signaling in AWGN.

Section 2.3.3, Section 2.3.4, and Section 2.3.5 which use just one sample per symbol. All three SSME algorithms are of the in-service type. The published SSME techniques [50, 51, 52] are designed strictly for BPSK signals in real AWGN. None of the SSME algorithms is easily extended to higher orders of modulation.

The SSME Algorithm in Wideband Channels ($SSME_0$)

All of the SNR estimation techniques considered in this study except for the $SSME_0$ can be applied to the bandlimited channel model described in Section 2.2. In order to investigate the $SSME_0$ algorithm, a suitable wideband channel model is required. The wideband system model adopted here is illustrated in Figure 2.2 which is a pared down version of Figure 2.1. The development that follows borrows some notational conventions from [51].

The original formulation of the SSME algorithm [50, 51] assumes the transmitted message sequence is represented as a stream of square NRZ pulses, a_{ij}, to which an AWGN sequence, z_{ij}, is added to form the received signal,

$$r_{ij} = \sqrt{S}\, a_{ij} + \sqrt{N}\, z_{ij} \quad \begin{cases} i \in \{0, 1, \ldots, N_{ss} - 1\} \\ j \in \{0, 1, \ldots, N_{sym} - 1\} \end{cases} \tag{2.23}$$

where the subscripts (i, j) denote the i^{th} sample of the j^{th} symbol, N_{ss} is the number of samples per symbol, and N_{sym} is the number of symbols upon which an estimate

of the SNR is based. The binary source samples are given by

$$a_{ij} = \pm\frac{1}{\sqrt{N_{ss}}} \begin{cases} i \in \{0, 1, \ldots, N_{ss} - 1\} \\ j \in \{0, 1, \ldots, N_{sym} - 1\} \end{cases}$$

so that E_s is unity when $S = 1$. Also, as before, the noise source power is set to unity so that $E\{z_{ij}^2\} = 1$. Two parameters, m_p and m_{ss}, are used to compute the SNR estimate from observations of the received signal samples, r_{ij}. The SSME SNR estimator is expressed in terms of these two parameters as [51]

$$\hat{\rho}_{\text{SSME}} = \frac{m_p}{2\left(\frac{1}{4}m_{ss} - m_p\right)} \tag{2.24}$$

where m_p and m_{ss} are derived from operations performed on the first half and last half of each symbol. Let $Y_{\alpha j}$ and $Y_{\beta j}$ be the sum of the first $N_{ss}/2$ samples and last $N_{ss}/2$ samples, respectively, of the j^{th} symbol expressed as

$$Y_{\alpha j} = \sum_{i=0}^{N_{ss}/2-1} r_{ij} \tag{2.25}$$

and

$$Y_{\beta j} = \sum_{i=N_{ss}/2}^{N_{ss}-1} r_{ij}. \tag{2.26}$$

The quantities, $Y_{\alpha j}$ and $Y_{\beta j}$, are used to compute m_p and m_{ss} according to

$$m_p = \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} Y_{\alpha j} Y_{\beta j} \tag{2.27}$$

$$m_{ss} = \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} (Y_{\alpha j} + Y_{\beta j})^2 \tag{2.28}$$

so that m_p is the product of the sum of the samples over the first half and last half of every symbol averaged over a block of N_{sym} symbols, while m_{ss} is the square of the sum of the samples over the first and last half of every symbol, again averaged over N_{sym} symbols.

The following analysis shows that the SSME_0 is unbiased in the limit that $N_{\text{sym}} \to \infty$. Substitution of (2.23) into (2.25) yields

$$Y_{\alpha j} = \sum_{i=0}^{N_{ss}/2-1} (\sqrt{S}a_{ij} + \sqrt{N}z_{ij})$$

$$= \sqrt{S}\sum_{i=0}^{N_{ss}/2-1}\left(\pm\sqrt{\frac{S}{N_{ss}}}\right)+\sqrt{N}\sum_{i=0}^{N_{ss}/2-1} z_{ij}$$

$$= \pm\frac{1}{2}\sqrt{SN_{ss}}+\sqrt{N}\sum_{i=0}^{N_{ss}/2-1} z_{ij}. \tag{2.29}$$

Similarly, substitution of (2.23) into (2.26) gives

$$Y_{\beta j} = \sum_{i=N_{ss}/2}^{N_{ss}-1}(\sqrt{S}\,a_{ij}+\sqrt{N}\,z_{ij})$$

$$= \pm\frac{1}{2}\sqrt{SN_{ss}}+\sqrt{N}\sum_{i=N_{ss}/2}^{N_{ss}-1} z_{ij}. \tag{2.30}$$

Substituting (2.29) and (2.30) into (2.27) one obtains

$$m_p = \frac{1}{N_{sym}}\sum_{j=0}^{N_{sym}-1}\left[\left(\pm\frac{1}{2}\sqrt{SN_{ss}}+\sqrt{N}\sum_{i=0}^{N_{ss}/2-1} z_{ij}\right)\left(\pm\frac{1}{2}\sqrt{SN_{ss}}+\sqrt{N}\sum_{i=N_{ss}/2}^{N_{ss}-1} z_{ij}\right)\right].$$

Multiplying out terms and taking expected values gives

$$E\{m_p\} = \frac{1}{N_{sym}}\sum_{j=0}^{N_{sym}-1}\left[\frac{SN_{ss}}{4}\pm\frac{\sqrt{SN_{ss}N}}{2}\sum_{i=0}^{N_{ss}/2-1}E\{z_{ij}\}\right.$$
$$\pm\frac{\sqrt{SN_{ss}N}}{2}\sum_{i=N_{ss}/2}^{N_{ss}-1}E\{z_{ij}\}$$
$$\left.+N\sum_{i=0}^{N_{ss}/2-1}\sum_{k=N_{ss}/2}^{N_{ss}-1}E\{z_{ij}z_{kj}\}\right]. \tag{2.31}$$

The terms with $E\{z_{ij}\}$ vanish since the AWGN is assumed to be a zero-mean process. The term $E\{z_{ij}z_{kj}\}$ vanishes also since i is never equal to k in the double summation so that (2.31) reduces to

$$E\{m_p\} = \frac{1}{N_{sym}}\sum_{j=0}^{N_{sym}-1} SN_{ss}/4 = SN_{ss}/4. \tag{2.32}$$

Substitution of (2.29) and (2.30) into (2.28) yields

$$m_{ss} = \frac{1}{N_{sym}}\sum_{j=0}^{N_{sym}-1}\left[\left(\pm\frac{1}{2}\sqrt{SN_{ss}}+\sqrt{N}\sum_{i=0}^{N_{ss}/2-1} z_{ij}\right)\right.$$
$$\left.+\left(\pm\frac{1}{2}\sqrt{SN_{ss}}+\sqrt{N}\sum_{i=N_{ss}/2}^{N_{ss}-1} z_{ij}\right)\right]^2$$

$$= \frac{1}{N_{sym}}\sum_{j=0}^{N_{sym}-1}\left(\pm\sqrt{SN_{ss}}+\sqrt{N}\sum_{i=0}^{N_{ss}-1} z_{ij}\right)^2.$$

Expanding terms and taking expected values gives

$$E\{m_{ss}\} = \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} \left(SN_{\text{ss}} + 2\sqrt{SN_{\text{ss}}N} \sum_{i=0}^{N_{\text{ss}}-1} E\{z_{ij}\} + N \sum_{i=0}^{N_{\text{ss}}-1} E\{z_{ij}^2\} \right). \quad (2.33)$$

After simplification, (2.33) becomes

$$E\{m_{ss}\} = N_{\text{ss}}(S+N). \quad (2.34)$$

The expected value of this SSME SNR estimator may be written as

$$E\{\hat{\rho}^{*}_{\text{SSME}}\} = \frac{E\{m_p\}}{2\left(\frac{1}{4}E\{m_{ss}\} - E\{m_p\}\right)} \quad (2.35)$$

if N_{sym} is large enough [51]. Substituting (2.32) and (2.34) into (2.35) one obtains

$$E\{\hat{\rho}^{*}_{\text{SSME}}\} = \frac{\frac{SN_{\text{ss}}}{4}}{2\left(\frac{SN_{\text{ss}}}{4} + \frac{N_{\text{ss}}N}{4} + \frac{SN_{\text{ss}}}{4}\right)} = \frac{S}{2N} = \frac{E_s}{N_0}.$$

Evidently, $\hat{\rho}^{*}_{\text{SSME}}$ is an estimator of E_s/N_0, not S/N. With a trivial modification, however, an SSME estimator for S/N is obtained as

$$\hat{\rho}_{\text{SSME}_0} = 2\hat{\rho}^{*}_{\text{SSME}} = \frac{m_p}{\frac{1}{4}m_{ss} - m_p}.$$

The expected value of $\hat{\rho}_{\text{SSME}_0}$ is

$$E\{\hat{\rho}_{\text{SSME}_0}\} = 2\left(\frac{S}{2N}\right) = \frac{S}{N}$$

which demonstrates that the SSME_0 estimator is unbiased for large values of N_{sym}.

An SSME Algorithm for Narrowband Channels (SSME_1)

In [51], it is concluded that the *ad hoc* application of the SSME_0 algorithm (designed for wideband channels) to narrowband channels is not entirely satisfactory. In [52], a modified SSME is presented which takes into account the bandlimiting of a receiver

filter. The modified-SSME SNR estimator of [52] is presented here in notation consistent with that used in the previous discussion of the SSME_0 estimator. The modified SSME estimator is expressed as

$$\hat{\rho}^{-1}_{\text{SSME,modified}} = \frac{m'_p}{2(\frac{1}{4}m_{ss} - m'_p)} \tag{2.36}$$

where

$$m'_p = m_p - \alpha \sum_{i=0}^{N_{ss}-1} \sum_{k=0}^{N_{ss}-1} R_{w,i-k},$$

$R_{w,i-k} = E\{w_i w_k\}$ is the autocorrelation of the coloured noise, and the subscripts i and k denote samples taken at the sampling rate (N_{ss} samples/symbol). In [52], it is assumed that the shape of the autocorrelation of the noise is known (specified by $R_{w,i-k}$) but the amplitude is unknown, thus the scale factor, α.

The assumptions made for the derivation of (2.36) in [52] are not consistent with the model presented in Section 2.2, so an independent derivation of an SNR algorithm based on the modified-SSME of [52] is presented below. The main difference between the model assumed in [52] and the model presented in Section 2.2 is that, in the former, bandlimiting is due to a receiver filter only, while the latter considers bandlimiting due to both a transmitter and a receiver filter. The modified-SSME, tailored to the model of Section 2.2, will be referred to as SSME_1 to distinguish it from the modified-SSME of [52].

Starting with the output samples of the MF, y_k, given by (2.5), the substitution of the transmitted signal, m_k, by (2.3) gives

$$\begin{aligned} y_k &= \sqrt{S} \sum_l h_l m_{k-l} + \sqrt{N} \sum_l h_l z_{k-l} \\ &= \sqrt{S} \sum_n a_n \sum_l h_l h_{k-nN_{ss}-l} + \sqrt{N} w_k. \end{aligned} \tag{2.37}$$

Assume that the coefficients of the RRC filter are non-zero over the range $l \in \{0, 1, \ldots, L-1\}$, and let $k = k_0$ denote the start of valid data at the output of the causal MF. The double-subscript, ij, used in the derivation of the SSME_0 is dropped and a single subscript, k, is used instead since the ij notation becomes

awkward in the analysis of a pulse-shaped signal (whose amplitude is not constant over the duration of a symbol). Re-writing (2.25) and (2.26) in terms of the output of the matched filter and the single subscript, k, one obtains

$$\begin{aligned} Y_{\alpha j} &= \sum_{k=k_0+jN_{ss}}^{k_0+jN_{ss}+N_{ss}/2-1} y_k \\ Y_{\beta j} &= \sum_{k=k_0+jN_{ss}+N_{ss}/2}^{k_0+jN_{ss}+N_{ss}-1} y_k. \end{aligned} \tag{2.38}$$

Substituting (2.37) for y_k in (2.38) yields

$$\begin{aligned} Y_{\alpha j} &= \sum_{k=k_0+jN_{ss}}^{k_0+jN_{ss}+N_{ss}/2-1} \left(\sqrt{S} \sum_n a_n \sum_l h_l h_{k-nN_{ss}-l} + \sqrt{N} w_k \right) \\ Y_{\beta j} &= \sum_{k=k_0+jN_{ss}+N_{ss}/2}^{k_0+jN_{ss}+N_{ss}-1} \left(\sqrt{S} \sum_n a_n \sum_l h_l h_{k-nN_{ss}-l} + \sqrt{N} w_k \right). \end{aligned}$$

The product of $Y_{\alpha j}$ and $Y_{\beta j}$ may be expressed as

$$Y_{\alpha j} Y_{\beta j} = \sum_{k_1=k_0+jN_{ss}}^{k_0+jN_{ss}+N_{ss}/2-1} \sum_{k_2=k_0+jN_{ss}+N_{ss}/2}^{k_0+jN_{ss}+N_{ss}-1} x_1 x_2 \tag{2.39}$$

where

$$x_1 = \sqrt{S} \sum_{n_1} a_{n_1} \sum_{l_1} h_{l_1} h_{k_1-n_1N_{ss}-l_1} + \sqrt{N} w_{k_1}$$

and

$$x_2 = \sqrt{S} \sum_{n_2} a_{n_2} \sum_{l_2} h_{l_2} h_{k_2-n_2N_{ss}-l_2} + \sqrt{N} w_{k_2}.$$

Recall from (2.27) that

$$m_p = \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} Y_{\alpha j} Y_{\beta j}.$$

Substituting (2.39) into the expression for m_p above, and taking expected values gives

$$E\{m_p\} = \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} (A + B + C + D) \tag{2.40}$$

where

$$
\begin{aligned}
A &= S\sum_{k_1}\sum_{k_2}\sum_{n_1}\sum_{n_2} E\{a_{n_1}a_{n_2}\}\sum_{l_1}\sum_{l_2} h_{l_1}h_{l_2}h_{k_1-n_1N_{ss}-l_1}h_{k_2-n_2N_{ss}-l_2} \\
B &= \sqrt{SN}\sum_{k_1}\sum_{k_2}\sum_{n_1} E\{a_{n_1}w_{k_2}\}\sum_{l_1} h_{l_1}h_{k_1-n_1N_{ss}-l_1} \\
C &= \sqrt{SN}\sum_{k_1}\sum_{k_2}\sum_{n_2} E\{a_{n_2}w_{k_1}\}\sum_{l_2} h_{l_2}h_{k_2-n_2N_{ss}-l_2} \\
D &= N\sum_{k_1}\sum_{k_2} E\{w_{k_1}w_{k_2}\}.
\end{aligned}
$$

Using the real forms of (2.7), (2.9), and (2.11), one obtains

$$
\begin{aligned}
E\{a_{n_1}w_{k_2}\} &= E\{a_{n_1}\}E\{w_{k_2}\} = 0 \\
E\{a_{n_2}w_{k_1}\} &= E\{a_{n_2}\}E\{w_{k_1}\} = 0 \\
E\{w_{k_1}w_{k_2}\} &= g_{k_1-k_2} \\
E\{a_{n_1}a_{n_2}\} &= \sigma_a^2\delta_{n_1,n_2} = \delta_{n_1,n_2}.
\end{aligned}
$$

Equation (2.40) may then be written as

$$
E\{m_p\} = \alpha S + \beta N
$$

where

$$
\begin{aligned}
\alpha = \frac{1}{N_{\text{sym}}}\sum_{j=0}^{N_{\text{sym}}-1}\Bigg\{\sum_{n}\Bigg[&\sum_{k_1=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1}\sum_{l_1=0}^{L-1} h_{l_1}h_{k_1-n_1N_{ss}-l_1} \\
&\times\sum_{k_2=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1}\sum_{l_2=0}^{L-1} h_{l_2}h_{k_2-n_2N_{ss}-l_2}\Bigg]\Bigg\}
\end{aligned}
$$

and

$$
\beta = \sum_{k_1=0}^{N_{ss}/2-1}\sum_{k_2=N_{ss}/2}^{N_{ss}-1} g_{k_1-k_2}.
$$

The limits of the summation over n are determined by the limits of the summation over l as

$$
\begin{aligned}
0 \le k - nN_{\text{ss}} - l \le L-1 \\
\left\lceil\frac{k-l-L+1}{N_{\text{ss}}}\right\rceil \le n \le \left\lfloor\frac{k-l}{N_{\text{ss}}}\right\rfloor
\end{aligned}
\tag{2.41}
$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x, and $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

Next, recall from (2.28) that

$$m_{ss} = \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} (Y_{\alpha j} + Y_{\beta j})^2.$$

The squared sum of $Y_{\alpha j}$ and $Y_{\beta j}$ is expressed as

$$\begin{aligned}
&(Y_{\alpha j} + Y_{\beta j})^2 = \\
&\quad \left[\sum_{k_1=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1} \left(\sqrt{S} \sum_{n_1} a_{n_1} \sum_{l_1} h_{l_1} h_{k_1-n_1N_{ss}-l_1} + \sqrt{N} w_{k_1} \right) \right. \\
&\quad \left. + \sum_{k_2=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1} \left(\sqrt{S} \sum_{n_2} a_{n_2} \sum_{l_2} h_{l_2} h_{k_2-n_2N_{ss}-l_2} + \sqrt{N} w_{k_2} \right) \right] \\
&\quad \times \left[\sum_{k_3=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1} \left(\sqrt{S} \sum_{n_3} a_{n_3} \sum_{l_3} h_{l_3} h_{k_3-n_3N_{ss}-l_3} + \sqrt{N} w_{k_3} \right) \right. \\
&\quad \left. + \sum_{k_4=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1} \left(\sqrt{S} \sum_{n_4} a_{n_4} \sum_{l_4} h_{l_4} h_{k_4-n_4N_{ss}-l_4} + \sqrt{N} w_{k_4} \right) \right]. \qquad (2.42)
\end{aligned}$$

After substitution of (2.42) into (2.28), expanding, and taking expected values, one obtains after some simplification that

$$\begin{aligned}
E\{m_{ss}\} \;=\; & S \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} \left\{ \sum_n \left[\sum_{k_1=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1} \sum_{k_3=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1} x_1 x_3 \right. \right. \\
& + \sum_{k_1=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1} \sum_{k_4=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1} x_1 x_4 \\
& + \sum_{k_2=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1} \sum_{k_3=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1} x_2 x_3 \\
& \left. \left. + \sum_{k_2=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1} \sum_{k_4=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1} x_2 x_4 \right] \right\} \\
& + N \sum_{k_1=0}^{N_{ss}-1} \sum_{k_2=0}^{N_{ss}-1} g_{k_1-k_2} \qquad (2.43)
\end{aligned}$$

where

$$x_i = \sum_{l_i} h_{l_i} h_{k_i - n_i N_{ss} - l_i}.$$

The sums in (2.43) may be consolidated so that $E\{m_{ss}\}$ reduces to

$$E\{m_{ss}\} = \gamma S + \delta N$$

where

$$\gamma = \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} \sum_{n} \left[\sum_{k=k_0+jN_{ss}}^{k+jN_{ss}+N_{ss}-1} \sum_{l=0}^{L-1} h_l h_{k-nN_{ss}-l} \right]^2$$

and

$$\delta = \sum_{k_1=0}^{N_{ss}-1} \sum_{k_2=0}^{N_{ss}-1} g_{k_1-k_2}.$$

The limits on n are as specified in (2.41).

In summary, given the two observables, m_p and m_{ss}, the solution of the system of equations given by

$$\begin{aligned} m_p &= \alpha S + \beta N \\ m_{ss} &= \gamma S + \delta N \end{aligned}$$

yields an estimator for the signal power, $\hat{S}_{\text{SSME}_1}$, and an estimator for the noise power, $\hat{N}_{\text{SSME}_1}$, so that the SSME_1 SNR estimator is stated as

$$\hat{\rho}_{\text{SSME}_1} = \frac{\hat{S}_{\text{SSME}_1}}{\hat{N}_{\text{SSME}_1}}.$$

Note that the coefficients, α, β, γ, and δ are functions of the filter taps and so are not constant among systems with different channel impulse responses.

A Second SSME Algorithm for Narrowband Channels (SSME_2)

An alternate SSME-type algorithm for narrowband channels is provided below based on a modification of SSME_1. This extension of the SSME algorithm has not been found in the literature. This new estimator will be called SSME_2 for convenience.

Recall that the original SSME strategy sums samples over half-symbol intervals as expressed in (2.38) and reproduced below for convenience as

$$\begin{aligned} Y_{\alpha j} &= \sum_{k=k_0+jN_{ss}}^{k_0+jN_{ss}+N_{ss}/2-1} y_k \\ Y_{\beta j} &= \sum_{k=k_0+jN_{ss}+N_{ss}/2}^{k_0+jN_{ss}+N_{ss}-1} y_k. \end{aligned}$$

These half-symbol sums are averaged over N_{sym} symbols to form the observed quantities, m_p and m_{ss}, as given in (2.27) and (2.28) and reproduced below for convenience:

$$\begin{aligned} m_p &= \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} Y_{\alpha j} Y_{\beta j} \\ m_{ss} &= \frac{1}{N_{\text{sym}}} \sum_{j=0}^{N_{\text{sym}}-1} (Y_{\alpha j} + Y_{\beta j})^2. \end{aligned}$$

These observed quantities are used to compute an estimate of the SNR.

In bandlimited channels, the variance of the half-symbol sums, $Y_{\alpha j}$ and $Y_{\beta j}$, from symbol to symbol is relatively large. If the half-symbol sums are averaged over the entire block of symbols first and *then* used in the expressions for m_p and m_{ss}, perhaps this would make a better SNR estimator. Of course, if the half-symbol sums are performed over the whole block on the the samples themselves, the result of the sums will be zero since the expected values of the transmitted message sequence and the coloured WGN out of the MF are zero. However, consider the half-symbol sums performed on the *square* of the samples, and let m_p and m_{ss} be replaced by m_q and m_{sq} which are defined as

$$m_q = \left(\frac{1}{N_{\text{sym}}} \sum_{j_1=0}^{N_{\text{sym}}-1} \sum_{k_1=k_0+j_1N_{ss}}^{k_0+j_1N_{ss}+\frac{N_{ss}}{2}-1} y_{k_1}^2 \right) \left(\frac{1}{N_{\text{sym}}} \sum_{j_2=0}^{N_{\text{sym}}-1} \sum_{k_2=k_0+j_2N_{ss}+\frac{N_{ss}}{2}}^{k_0+j_2N_{ss}+N_{ss}-1} y_{k_2}^2 \right)$$

$$m_{sq} = \frac{1}{N_{\text{sym}}N_{\text{ss}}} \sum_{j=0}^{N_{\text{sym}}-1} \sum_{k=k_0+jN_{ss}}^{k_0+jN_{ss}+N_{ss}-1} y_k^4.$$

The expected values, $E\{m_q\}$ and $E\{m_{sq}\}$ are given by (the results are given without the lengthy derivations)

$$\begin{aligned} E\{m_q\} &= \alpha_1 S^2 + \beta_1 SN + \gamma_1 N^2 \\ E\{m_{sq}\} &= \alpha_2 S^2 + \beta_2 SN + \gamma_2 N^2 \end{aligned} \tag{2.44}$$

where

$$\alpha_1 = \frac{1}{N_{\text{sym}}^2}\left[\sum_{j_1=0}^{N_{\text{sym}}-1}\sum_{n}\sum_{k_1=k_0+j_1N_{ss}}^{k_0+j_1N_{ss}+N_{ss}/2-1}\left(\sum_{l_1=0}^{L-1}h_{l_1}h_{k_1-nN_{ss}-l_1}\right)^2\right]$$
$$\times\left[\sum_{j_2=0}^{N_{\text{sym}}-1}\sum_{n}\sum_{k_2=k_0+j_2N_{ss}+N_{ss}/2}^{k_0+j_2N_{ss}+N_{ss}-1}\left(\sum_{l_2=0}^{L-1}h_{l_2}h_{k_2-nN_{ss}-l_2}\right)^2\right]$$

$$\beta_1 = \beta_{11}+\beta_{12}+\beta_{13}$$

$$\beta_{11} = \frac{N_{\text{ss}}}{2N_{\text{sym}}}\sum_{j=0}^{N_{\text{sym}}-1}\sum_{n}\sum_{k=k_0+jN_{ss}}^{k_0+jN_{ss}+\frac{N_{ss}}{2}-1}\left(\sum_{l=0}^{L-1}h_lh_{k-nN_{ss}-l}\right)^2$$

$$\beta_{12} = \frac{N_{\text{ss}}}{2N_{\text{sym}}}\sum_{j=0}^{N_{\text{sym}}-1}\sum_{n}\sum_{k=k_0+jN_{ss}+\frac{N_{ss}}{2}}^{k_0+jN_{ss}+N_{ss}-1}\left(\sum_{l=0}^{L-1}h_lh_{k-nN_{ss}-l}\right)^2$$

$$\beta_{13} = \frac{4}{N_{\text{sym}}^2}\sum_{j_1=0}^{N_{\text{sym}}-1}\sum_{j_2=0}^{N_{\text{sym}}-1}\sum_{k_1=k_0+j_1N_{ss}}^{k_0+j_1N_{ss}+\frac{N_{ss}}{2}-1}\sum_{k_2=k_0+j_2N_{ss}+\frac{N_{ss}}{2}}^{k_0+j_2N_{ss}+N_{ss}-1}\Bigg[$$
$$\left(\sum_{n}\sum_{l_1=0}^{L-1}h_{l_1}h_{k_1-nN_{ss}-l_1}\right)\left(\sum_{n}\sum_{l_2=0}^{L-1}h_{l_2}h_{k_2-nN_{ss}-l_2}\right)g_{k_1-k_2}\Bigg]$$

$$\gamma_1 = \frac{1}{N_{\text{sym}}^2}\sum_{j_1=0}^{N_{\text{sym}}-1}\sum_{j_2=0}^{N_{\text{sym}}-1}\sum_{k_1=k_0+j_1N_{ss}}^{k_0+j_1N_{ss}+\frac{N_{ss}}{2}-1}\sum_{k_2=k_0+j_2N_{ss}+\frac{N_{ss}}{2}}^{k_0+j_2N_{ss}+N_{ss}-1}(1+2g_{k_1-k_2}^2).$$

$$\alpha_2 = \frac{1}{N_{\text{sym}}N_{\text{ss}}}\sum_{j=0}^{N_{\text{sym}}-1}\sum_{n}\sum_{k=k_0+jN_{ss}}^{k_0+jN_{ss}+N_{ss}-1}\left(\sum_{l=0}^{L-1}h_lh_{k-nN_{ss}-l}\right)^4$$

$$\beta_2 = \frac{6}{N_{\text{sym}}N_{\text{ss}}}\sum_{j=0}^{N_{\text{sym}}-1}\sum_{n}\sum_{k=k_0+jN_{ss}}^{k_0+jN_{ss}+N_{ss}-1}\left(\sum_{l=0}^{L-1}h_lh_{k-nN_{ss}-l}\right)^2$$

$$\gamma_2 = 3.$$

The bounds on n are given by (2.41).

The following outlines the method to generate an estimate of the SNR using the SSME_2 method:

1. Generate the two observables, m_q and m_{sq}, from a sequence of N_{sym} symbols.

2. Substitute the observed quantities, m_q and m_{sq}, for $E\{m_q\}$ and $E\{m_{sq}\}$, respectively, in the system of equations identified by (2.44) to get

$$m_q = \alpha_1 S^2 + \beta_1 SN + \gamma_1 N^2$$

$$m_{sq} = \alpha_2 S^2 + \beta_2 SN + \gamma_2 N^2$$

3. Solve the system of equations to obtain an estimate of the signal power, $\hat{S}_{\mathrm{SSME}_2}$, and an estimate of the noise power, $\hat{N}_{\mathrm{SSME}_2}$.

4. Using $\hat{S}_{\mathrm{SSME}_2}$ and $\hat{N}_{\mathrm{SSME}_2}$, the SSME_2 SNR estimator may be expressed as

$$\hat{\rho}_{\mathrm{SSME}_2} = \frac{\hat{S}_{\mathrm{SSME}_2}}{\hat{N}_{\mathrm{SSME}_2}}.$$

As in the case of the SSME_1, the coefficients of the SSME_2 are functions of the filter tap coefficients and so are different for channels with different impulse responses.

2.3.2 The Maximum-Likelihood SNR Estimator

An SNR estimator based on classical maximum-likelihood (ML) estimation theory [67] was introduced by Kerr [44], and Gagliardi and Thomas [45]. In the latter, more detailed treatment, based on work done by Thomas [47], the ML SNR estimator for coherent BPSK signals in AWGN is derived and studied.

In the this section, some general theory of ML estimation is first presented, followed by a derivation of the ML SNR estimator for coherent BPSK signals in real AWGN. Lastly, the ML SNR estimator is extended to the case of general, coherent M-ary PSK signals in complex AWGN.

Some General Estimation Theory

The following treatment is based on concepts developed in [67, 47, 68]. Consider some observable random variable, x, whose value is dependent on a set of L fixed, unknown parameters, $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_L)$. Assume K such observations are available to form the observation vector, $\mathbf{x} = (x_1, x_2, \ldots, x_K)$. The objective is to find an estimator, $\hat{\boldsymbol{\theta}}(\mathbf{x})$, which generates estimates of $\boldsymbol{\theta}$ based on the observations, $\mathbf{x}$.

The *a posteriori* pdf of $\boldsymbol{\theta}$ conditioned on $\mathbf{x}$ is written as $p(\boldsymbol{\theta}|\mathbf{x})$. The maximum value of this function occurs at the value of $\boldsymbol{\theta}$ which is most likely given the observa-

tions, $\mathbf{x}$. The estimator that generates estimates of $\boldsymbol{\theta}$ based on the maximization of $p(\boldsymbol{\theta}|\mathbf{x})$ is called the *maximum a posteriori* (MAP) estimator.

Since $p(\mathbf{x}|\boldsymbol{\theta})$ is often easier to find than $p(\boldsymbol{\theta}|\mathbf{x})$, consider the identity [66] expressed as

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{x})}, \tag{2.45}$$

where $p(\boldsymbol{\theta})$ represents a priori knowledge of the distribution of $\boldsymbol{\theta}$. Since the MAP estimator maximizes $p(\boldsymbol{\theta}|\mathbf{x})$ with respect to $\boldsymbol{\theta}$, and since $p(\mathbf{x})$ is not a function of $\boldsymbol{\theta}$, the value of $\boldsymbol{\theta}$ that maximizes (2.45) also maximizes $p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})$. The natural logarithm is a monotonic function so that the value of $\boldsymbol{\theta}$ that maximizes $p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ also maximizes

$$\ln p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}) = \ln p(\mathbf{x}|\boldsymbol{\theta}) + \ln p(\boldsymbol{\theta}). \tag{2.46}$$

It is often mathematically more convenient to work with the logarithm.

The value of $\boldsymbol{\theta}$ that maximizes (2.46) is found by taking partial derivatives with respect to each of the unknown parameters, θ_i, $i \in \{1, 2, \ldots, L\}$, and setting the results equal to zero as

$$\frac{\partial}{\partial \theta_i} \ln p(\mathbf{x}|\boldsymbol{\theta})\bigg|_{\theta_i = \hat{\theta}_i(\mathbf{x})} + \frac{\partial}{\partial \theta_i} \ln p(\boldsymbol{\theta})\bigg|_{\theta_i = \hat{\theta}_i(\mathbf{x})} = 0, \; i \in \{1, 2, \ldots, L\}. \tag{2.47}$$

Assume that all values of $\boldsymbol{\theta}$ are equally likely; that is, assume that $\boldsymbol{\theta}$ is uniformly distributed for all values of interest so that all partial derivatives of $\ln p(\theta_i)$ vanish and (2.47) becomes

$$\frac{\partial}{\partial \theta_i} \ln p(\mathbf{x}|\boldsymbol{\theta})\bigg|_{\theta_i = \hat{\theta}_{\mathrm{ML},i}(\mathbf{x})} = \frac{\partial}{\partial \theta_i} \Gamma(\boldsymbol{\theta})\bigg|_{\theta_i = \hat{\theta}_{\mathrm{ML},i}(\mathbf{x})} = 0, \; i \in \{1, 2, \ldots, L\} \tag{2.48}$$

where $\Gamma(\boldsymbol{\theta}) = \ln p(\mathbf{x}|\boldsymbol{\theta})$ is the *likelihood function* of $\boldsymbol{\theta}$. Solving (2.48) for $\boldsymbol{\theta}$ in terms of $\mathbf{x}$ yields the *maximum likelihood* (ML) estimator for $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_{\mathrm{ML}}(\mathbf{x})$.

Derivation of the ML SNR Estimator for Real Channels

The ML estimator described above is now applied to the model presented in Section 2.2 and illustrated in Figure 2.1, with complex signals replaced by real ones.

The received signal is given by (2.4) as

$$r_k = \sqrt{S}\, m_k + \sqrt{N}\, z_k.$$

The observation vector in this case is the sequence of $K = N_{\text{sym}} N_{\text{ss}}$ received real signal samples,

$$\mathbf{r} = \{r_0, r_1, \ldots, r_{K-1}\}.$$

The two unknown parameters to be estimated are S and N ($L = 2$). All that is required to find the ML estimators for S and N is the likelihood function,

$$\Gamma(S, N) = \ln p(\mathbf{r}|S, N).$$

Let $\nu_k = \sqrt{N}\, z_k$ so that the pdf of the noise term may be written as

$$p(\nu_k) = \frac{1}{\sqrt{2\pi N}} e^{-\nu_k^2/2N}. \tag{2.49}$$

Using (2.4) and (2.49), the pdf of r_k conditioned on S and N may be expressed as

$$p(r_k|S, N) = \frac{1}{\sqrt{2\pi N}} \exp\left(-\frac{(r_k - \sqrt{S}\, m_k)^2}{2N}\right).$$

Since r_k represents samples of the received signal *before* the MF, the noise samples are independent and the joint pdf for all K observed samples of r_k may be written as

$$\begin{aligned} p(\mathbf{r}|S, N) &= \prod_{k=0}^{K-1} p(r_k|S, N) \\ &= (2\pi N)^{-\frac{K}{2}} \exp\left\{-\frac{1}{2N} \sum_{k=0}^{K-1} (r_k - \sqrt{S}\, m_k)^2\right\}, \end{aligned}$$

so that the likelihood function is

$$\begin{aligned} \Gamma(S, N) &= \ln p(\mathbf{r}|S, N) \\ &= -\frac{K}{2} \ln(2\pi N) - \frac{1}{2N} \sum_{k=0}^{K-1} \left(r_k - \sqrt{S}\, m_k\right)^2. \end{aligned}$$

Strictly speaking, $\Gamma(S, N)$ also depends on the specific message sequence, m_k, that was transmitted. Let $m_k^{(i)}$ denote the i^{th} sequence out of $2^{N_{\text{sym}}}$ possible transmitted

binary sequences. This data dependence may be explicitly included in the likelihood function as

$$\Gamma(S,N,i) = -\frac{K}{2}\ln(2\pi N) - \frac{1}{2N}\sum_{k=0}^{K-1}\left(r_k - \sqrt{S}\,m_k^{(i)}\right)^2. \tag{2.50}$$

The estimators for S and N are the solutions of the system of equations

$$\left.\frac{\partial}{\partial S}\Gamma(S,N,i)\right|_{\substack{S=\hat{S}\\N=\hat{N}\\i=\hat{i}}} = 0 \tag{2.51a}$$

$$\left.\frac{\partial}{\partial N}\Gamma(S,N,i)\right|_{\substack{S=\hat{S}\\N=\hat{N}\\i=\hat{i}}} = 0. \tag{2.51b}$$

Taking the partial derivatives of $\Gamma(S,N,i)$ with respect to S, one obtains

$$\left.\frac{\partial}{\partial S}\Gamma(S,N,i)\right|_{\substack{S=\hat{S}\\N=\hat{N}\\i=\hat{i}}} = \frac{1}{\sqrt{\hat{S}}}\sum_{k=0}^{K-1} r_k m_k^{(i)} - \sum_{k=0}^{K-1}(m_k^{(i)})^2 = 0,$$

and solving for $\hat{S}$ gives the ML RxDA estimator for S for real channels as

$$\hat{S}_{\text{ML RxDA,real}} = \left[\frac{\sum_{k=0}^{K-1} r_k m_k^{(i)}}{\sum_{k=0}^{K-1}(m_k^{(i)})^2}\right]^2. \tag{2.52}$$

Similarly, taking the partial derivative with respect to N, one obtains

$$\left.\frac{\partial}{\partial N}\Gamma(S,N,i)\right|_{\substack{S=\hat{S}\\N=\hat{N}\\i=\hat{i}}} = -\frac{K}{2\hat{N}} + \frac{1}{2\hat{N}^2}\sum_{k=0}^{K-1}(r_k - \sqrt{S}m_k^{(i)})^2 = 0,$$

and solving for $\hat{N}$ gives the ML RxDA estimator for N for real channels as

$$\begin{aligned}\hat{N}_{\text{ML RxDA,real}} &= \frac{1}{K}\sum_{k=0}^{K-1} r_k^2 - 2\sqrt{\hat{S}}\frac{1}{K}\sum_{k=0}^{K-1} r_k m_k^{(i)} + \hat{S}\frac{1}{K}\sum_{k=0}^{K-1}(m_k^{(i)})^2 \\ &= \frac{1}{K}\sum_{k=0}^{K-1} r_k^2 - \hat{S}\frac{1}{K}\sum_{k=0}^{K-1}(m_k^{(i)})^2. \end{aligned}\tag{2.53}$$

Finding the ML estimate of i is a detection theory problem [67]. Consider the expansion of (2.50) as

$$\Gamma(S,N,i) = -\frac{K}{2}\ln(2\pi N) - \frac{1}{2N}\sum_{k=0}^{K-1} r_k^2 + \frac{\sqrt{S}}{N}\sum_{k=0}^{K-1} r_k m_k^{(i)} - \frac{S}{2N}\sum_{k=0}^{K-1}(m_k^{(i)})^2.$$

Since the objective is to maximize the likelihood function, $\hat{i}$ denotes the transmitted sequence that maximizes the inner product of r_k and $m_k^{(i)}$ [45]. This is precisely the task of a ML receiver [65, page 239] which consists of a device that selects the largest output of a bank of $2^{N_{sym}}$ correlators, where the i^{th} correlator performs the inner product of the received signal with a copy of the i^{th} possible binary transmitted sequence. The sequence associated with the correlator having the largest output is chosen as the most likely sequence and is denoted by $\hat{i}$.

Given (2.52), the ML estimator for S, and (2.53), the ML estimator for N, the ML SNR estimator may be found using the property [45] that the ML estimate of the ratio of two parameters is the same as the ratio of the two separate ML estimates. This fortunate property allows the in-service, ML RxDA SNR estimator for real channels to be written directly as

$$\hat{\rho}_{\text{ML RxDA,real}} = \frac{\hat{S}_{\text{ML RxDA,real}}}{\hat{N}_{\text{ML RxDA,real}}} = \frac{N_{ss}^2 \left[\frac{1}{K}\sum_{k=0}^{K-1} r_k m_k^{(\hat{i})}\right]^2}{\frac{1}{K}\sum_{k=0}^{K-1} r_k^2 - N_{ss}\left[\frac{1}{K}\sum_{k=0}^{K-1} r_k m_k^{(\hat{i})}\right]^2}, \tag{2.54}$$

where (2.16) has been used to evaluate the average power of the transmitted message sequence according to

$$\frac{1}{K}\sum_{k=0}^{K-1}(m_k^{(i)})^2 = R_m(0) = \frac{\sigma_a^2 g_0}{N_{ss}} = 1/N_{ss}. \tag{2.55}$$

Note that in [45], it is assumed that $R_m(0) = 1$.

The TxDA form of the ML SNR estimator is identical to (2.54) except that i does not need to be estimated because the transmitted sequence is known exactly by the receiver. The ML TxDA SNR estimator for real channels may be expressed as

$$\hat{\rho}_{\text{ML TxDA,real}} = \frac{N_{ss}^2 \left[\frac{1}{K}\sum_{k=0}^{K-1} r_k m_k\right]^2}{\frac{1}{K}\sum_{k=0}^{K-1} r_k^2 - N_{ss}\left[\frac{1}{K}\sum_{k=0}^{K-1} r_k m_k\right]^2}. \tag{2.56}$$

where m_k represents the true, pulse-shaped, transmitted samples given by (2.3). The ML TxDA SNR estimator given in [45, 47] appears different from (2.56) due to the assumption in [45, 47] that $R_m(0) = 1$.

Thomas [47] shows that the ML TxDA SNR estimator exhibits a small bias. It can be shown that the bias of the ML TxDA SNR estimator given by (2.56) is

$$E\{\hat{\rho}_{\text{ML TxDA,real}}\} = \frac{K}{K-3}\left(\rho + \frac{1}{N_{\text{sym}}}\right), \tag{2.57}$$

where $\rho = S/N$ represents the true SNR. As pointed out in [47], this bias can be minimized by scaling $\hat{N}_{\text{ML TxDA,real}}$ by $1/(K-3)$ instead of by $1/K$. Letting $\hat{\rho}'_{\text{ML,TxDA}}$ denote the "reduced-bias" ML TxDA SNR estimator, one has

$$\hat{\rho}'_{\text{ML TxDA,real}} = \frac{N_{ss}^2\left[\frac{1}{K}\sum_{k=0}^{K-1} r_k m_k\right]^2}{\frac{1}{K-3}\sum_{k=0}^{K-1} r_k^2 - \frac{N_{ss}}{K(K-3)}\left[\sum_{k=0}^{K-1} r_k m_k\right]^2}. \tag{2.58}$$

The expected value of the reduced-bias ML SNR estimator given by (2.58) is

$$E\{\hat{\rho}'_{\text{ML TxDA,real}}\} = \rho + \frac{1}{N_{\text{sym}}}.$$

The difference between the biased and reduced-bias estimates is small unless K is very small.

In [47], expressions are also given for the variances of the biased and reduced-bias ML TxDA SNR estimators, again assuming $R_m(0) = 1$. The expressions in [47] may be modified to represent the theoretical variances of the ML TxDA estimators given by (2.56) and (2.58) (for which $R_m(0) = 1/N_{ss}$) as

$$\text{Var}\{\hat{\rho}_{\text{ML TxDA,real}}\} = \frac{2K^2}{(K-3)^2(K-5)}\left[\rho^2 + 2N_{ss}\rho\left(1-\frac{2}{K}\right) + \frac{N_{ss}^2}{K}\left(1-\frac{2}{K}\right)\right] \tag{2.59}$$

$$\text{Var}\{\hat{\rho}'_{\text{ML TxDA,real}}\} = \frac{2}{K-5}\left[\rho^2 + 2N_{ss}\rho\left(1-\frac{2}{K}\right) + \frac{N_{ss}^2}{K}\left(1-\frac{2}{K}\right)\right]. \tag{2.60}$$

Analytical expressions for the bias and variance of the ML RxDA SNR estimator are difficult to derive since receiver decision errors must be taken into account. In [47], no derivation is given for the bias and variance for the RxDA case, but the assumption is made here that $\hat{N}_{\text{ML RxDA,real}}$ may be modified in the same manner as described above for $\hat{N}_{\text{ML TxDA,real}}$ in order to obtain a reduced-bias, in-service ML RxDA SNR estimator from (2.54) as

$$\hat{\rho}'_{\text{ML RxDA,real}} = \frac{N_{ss}^2\left[\frac{1}{K}\sum_{k=0}^{K-1} r_k m_k^{(i)}\right]^2}{\frac{1}{K-3}\sum_{k=0}^{K-1} r_k^2 - \frac{N_{ss}}{K(K-3)}\left[\sum_{k=0}^{K-1} r_k m_k^{(i)}\right]^2}. \tag{2.61}$$

The reduced-bias ML SNR estimator expressions given by (2.58) and (2.61) are the ones used in this study for real channels (coherent BPSK signals in real AWGN). Note that these expressions may be used unchanged in any coherent BPSK system using an arbitrary pulse-shaping filter as long as the sum of the squares of the filter coefficients is unity so that the average power of the transmitted message sequence is $1/N_{ss}$. In the case that the sum of the squares of the filter coefficients is not unity, modified expressions may be derived assuming the pulse-shaping filter coefficients are known.

Derivation of the ML SNR Estimator for Complex Channels

The ML estimator for complex channels is derived from samples of the complex, noisy received signal given by (2.4) which may be re-written explicitly in terms of real and imaginary parts as

$$r_k = r_{I_k} + jr_{Q_k} = \sqrt{S}(m_{I_k} + jm_{Q_k}) + \sqrt{N}(z_{I_k} + jz_{Q_k}). \tag{2.62}$$

Let $\nu_{I_k} = \sqrt{N}z_{I_k}$ and $\nu_{Q_k} = \sqrt{N}z_{Q_k}$ represent the in-phase and quadrature components of the noise, respectively, each having zero mean and variance, $N/2$. The in-phase and quadrature components of the noise are assumed independent so that their joint pdf can be written as

$$f(\nu_{I_k}, \nu_{Q_k}) = \frac{1}{\pi N} e^{-(\nu_{I_k}^2 + \nu_{Q_k}^2)/N}. \tag{2.63}$$

Using (2.62) and (2.63), the joint pdf of the in-phase and quadrature components of a received signal sample, given S is the signal power, N is the noise power, and i denotes the i^{th} sequence of $M^{N_{sym}}$ possible transmitted message sequences, may be expressed as

$$f(r_{I_k}, r_{Q_k}|S, N, i) = \frac{1}{\pi N} \exp\left(-\frac{\left(r_{I_k} - \sqrt{S}m_{I_k}^{(i)}\right)^2 + \left(r_{Q_k} - \sqrt{S}m_{Q_k}^{(i)}\right)^2}{N} \right). \tag{2.64}$$

The independence of the signal and noise sequences allows the joint pdf of the $K = N_{\text{sym}}N_{\text{ss}}$ received samples to be written as

$$\begin{aligned} f(\mathbf{r_I}, \mathbf{r_Q}|S, N, i) &= \prod_{k=0}^{K-1} f(r_{I_k}, r_{Q_k}|S, N, i) \\ &= (\pi N)^{-K} \exp\left[-\frac{1}{N}\left(\sum_{k=0}^{K-1}\left(r_{I_k} - \sqrt{S}m_{I_k}^{(i)}\right)^2 + \sum_{k=0}^{K-1}\left(r_{Q_k} - \sqrt{S}m_{Q_k}^{(i)}\right)^2\right)\right] \end{aligned} \quad (2.65)$$

where $\mathbf{r_I} = \{r_{I_0}, r_{I_1}, \ldots, r_{I_{K-1}}\}$ and $\mathbf{r_Q} = \{r_{Q_0}, r_{Q_1}, \ldots, r_{Q_{K-1}}\}$. The likelihood function, $\Gamma(S, N, i)$, is given by

$$\begin{aligned} \Gamma(S, N, i) &= \ln f(\mathbf{r_I}, \mathbf{r_Q}|S, N, i) \\ &= -K\ln(\pi N) - \frac{1}{N}\left[\sum_{k=0}^{K-1}\left(r_{I_k} - \sqrt{S}m_{I_k}^{(i)}\right)^2 + \sum_{k=0}^{K-1}\left(r_{Q_k} - \sqrt{S}m_{Q_k}^{(i)}\right)^2\right]. \end{aligned} \quad (2.66)$$

Solving the system of equations given by (2.51) using (2.66), one obtains the ML RxDA signal power estimator for complex channels as

$$\hat{S}_{\text{ML RxDA,complex}} = \left[\frac{\frac{1}{K}\sum_{k=0}^{K-1}(r_{I_k}m_{I_k}^{(i)} + r_{Q_k}m_{Q_k}^{(i)})}{\frac{1}{K}\sum_{k=0}^{K-1}[(m_{I_k}^{(i)})^2 + (m_{Q_k}^{(i)})^2]}\right]^2, \quad (2.67)$$

and the ML RxDA noise power estimator for complex channels as

$$\hat{N}_{\text{ML RxDA,complex}} = \frac{1}{K}\sum_{k=0}^{K-1}(r_{I_k}^2 + r_{Q_k}^2) - \hat{S}\frac{1}{K}\sum_{k=0}^{K-1}\left[\left(m_{I_k}^{(i)}\right)^2 + \left(m_{Q_k}^{(i)}\right)^2\right]. \quad (2.68)$$

In a similar manner to that shown by (2.55) for the real case, the average power of the complex transmitted message sequence simplifies to

$$\frac{1}{K}\sum_{k=0}^{K-1}\left[\left(m_{I_k}^{(i)}\right)^2 + \left(m_{Q_k}^{(i)}\right)^2\right] = R_m(0) = \frac{\sigma_a^2 g_0}{N_{\text{ss}}} = 1/N_{\text{ss}}. \quad (2.69)$$

As mentioned earlier, the ML estimate of the ratio of two parameters is the ratio of the ML estimates of the two parameters so that, in a similar manner to that shown by (2.54), the ML RxDA SNR estimator for complex channels may be written as

$$\hat{\rho}_{\text{ML RxDA,complex}} = \frac{\hat{S}_{\text{ML RxDA,complex}}}{\hat{N}_{\text{ML RxDA,complex}}} = \frac{N_{\text{ss}}^2\left[\frac{1}{K}\sum_{k=0}^{K-1}\text{Re}\{r_k^* m_k^{(i)}\}\right]^2}{\frac{1}{K}\sum_{k=0}^{K-1}|r_k|^2 - N_{\text{ss}}\left[\frac{1}{K}\sum_{k=0}^{K-1}\text{Re}\{r_k^* m_k^{(i)}\}\right]^2} \quad (2.70)$$

where $\hat{\imath}$ corresponds to the sequence of received symbols estimated by a ML receiver, as before, and $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex quantity.

The ML TxDA SNR estimator for complex channels is obtained by replacing the estimated symbol sequence, denoted by $\hat{\imath}$, with the known, transmitted pulse-shaped samples, denoted by m_k, to yield the desired estimator as

$$\hat{\rho}_{\mathrm{ML\,TxDA,complex}} = \frac{\hat{S}_{\mathrm{ML\,TxDA,complex}}}{\hat{N}_{\mathrm{ML\,TxDA,complex}}} = \frac{N_{\mathrm{ss}}^2 \left[\frac{1}{K}\sum_{k=0}^{K-1} \mathrm{Re}\{r_k^* m_k\}\right]^2}{\frac{1}{K}\sum_{k=0}^{K-1} |r_k|^2 - N_{\mathrm{ss}}\left[\frac{1}{K}\sum_{k=0}^{K-1} \mathrm{Re}\{r_k^* m_k\}\right]^2}. \tag{2.71}$$

It was shown earlier that multiplying $\hat{N}_{\mathrm{ML\,TxDA,real}}$ by $K/(K-3)$ yields a reduced-bias ML SNR estimator for the real case. It was found by simulation that the *ad hoc* application of this factor to $\hat{N}_{\mathrm{ML\,TxDA,complex}}$ does not improve the performance of the ML SNR estimator in the complex case as the bias becomes worse. In order to obtain the correct multiplier factor, the pdf of $\hat{\rho}_{\mathrm{ML\,TxDA,complex}}$ must be derived so that the bias of the complex ML TxDA SNR estimator may be evaluated as Thomas did for the real case in [47]. This derivation is complicated.

Instead, in a heuristic attempt to find a reduced-bias estimator, consider computing an estimate of the noise power as the unbiased sample variance of the noisy received signal[2] so that an expression for a reduced-bias ML TxDA SNR estimator for complex channels may be written as

$$\hat{\rho}'_{\mathrm{ML\,TxDA,complex}} = \frac{N_{\mathrm{ss}}^2 \left[\frac{1}{K}\sum_{k=0}^{K-1} \mathrm{Re}\{r_k^* m_k\}\right]^2}{\frac{1}{K-1}\sum_{k=0}^{K-1} |r_k|^2 - \frac{N_{\mathrm{ss}}}{K(K-1)}\left[\sum_{k=0}^{K-1} \mathrm{Re}\{r_k^* m_k\}\right]^2}. \tag{2.72}$$

By simulation it was observed that this form of the ML TxDA SNR estimator in a system with 8-PSK-modulated signals corrupted by complex AWGN performs marginally better than the form given by (2.71).

Similarly, a reduced-bias, in-service ML RxDA SNR estimator for complex chan-

[2]The idea to try the unbiased sample variance in an attempt to find a reduced-bias ML SNR estimator for complex channels was inspired by the SNR estimator proposed by Gilchriest [46] described in Section 2.3.3.

nels may be written as

$$\hat{\rho}'_{\text{ML RxDA,complex}} = \frac{N_{ss}^2 \left[\frac{1}{K}\sum_{k=0}^{K-1} \text{Re}\left\{r_k^* m_k^{(i)}\right\}\right]^2}{\frac{1}{K-1}\sum_{k=0}^{K-1}|r_k|^2 - \frac{N_{ss}}{K(K-1)}\left[\sum_{k=0}^{K-1} \text{Re}\left\{r_k^* m_k^{(i)}\right\}\right]^2}. \tag{2.73}$$

The reduced-bias ML SNR expressions given by (2.72) and (2.73) are the ones used in this study for complex channels. These expressions may be used unchanged in any system where coherent M-ary PSK modulation is employed in a complex AWGN channel, and where the sum of the squares of the pulse-shaping filter coefficients is unity. These estimators are easily extended to other coherent modulation schemes, such as QAM, and other systems using arbitrary pulse-shaping filters as long as the impulse response of the pulse-shaping filter in the transmitter is known.

The ML estimators and the SSME series of SNR estimators are the only SNR algorithms in this study that operate on more than one sample per symbol. The remaining algorithms discussed in subsequent sections operate on the one optimal sample per symbol at the output of the MF.

2.3.3 The Squared Signal-to-Noise Variance (SNV) Estimator

This estimator is based on the first and second moments of the optimally-sampled output of the MF, and is probably the simplest and most straightforward approach to SNR estimation. It was studied by Gilchriest in 1966 at the Jet Propulsion Laboratory (JPL) in Pasadena, California, and reported in an internal company document [46]. Benedict and Soong [42] describe an estimation procedure based on first and second moments, but it is computationally more complicated (the signal power must be solved implicitly), and only the separate estimation of carrier level and noise level is considered. The SNV estimator studied by Gilchriest [46] estimates explicitly the ratio of signal power to noise power. Coherent BPSK signaling in the real AWGN channel is assumed.

In this section, the original formulation of the SNV RxDA estimator for coherent

BPSK signals in real AWGN is first presented, followed by a description of the TxDA counterpart. The SNV SNR estimator is then extended to the case of general, coherent M-ary PSK signals in complex AWGN. Reduced-bias forms of the SNV estimators are provided for both the real and complex cases.

Derivation of the SNV SNR Estimator for Real Channels

The SNV RxDA SNR estimator for real channels proposed by Gilchriest [46] is simply the ratio of the square of the mean of the absolute value of the signal to the noise variance. Recall from equation (2.22) in Section 2.2 that the sampled output of the MF may be expressed as

$$y_n = \sqrt{S}\, a_n + \sqrt{N}\, w_n.$$

Let

$$x_n = |y_n| = y_n\, a_n^{(i)} \tag{2.74}$$

where i denotes one of the $2^{N_{sym}}$ possible symbol sequences estimated by the receiver. The latter equality holds because a_n, given by (2.2), is either 1 or -1 for BPSK signals in real channels. Due to the independence of the signal and the noise sequences, the mean of x_n is evaluated as

$$E\{x_n\} = \sqrt{S}\, E\{a_n a_n^{(i)}\} + \sqrt{N}\, E\{w_n\} E\{a_n^{(i)}\} = \sqrt{S} \tag{2.75}$$

where the assumption is made that there are no receiver errors so that $a_n^{(i)} = a_n$ and, consequently,

$$E\{a_n a_n^{(i)}\} = E\{a_n^2\} = 1.$$

The variance of x_n is evaluated as

$$\begin{aligned} \mathrm{Var}\{x_n\} &= \mathrm{Var}\{|y_n|\} \\ &= E\{|y_n|^2\} - E\{|y_n|\}^2 \\ &= E\{y_n^2\} - E\{|y_n|\}^2. \end{aligned} \tag{2.76}$$

Using the real forms of (2.7), (2.9), and (2.11), one has

$$E\{y_n^2\} = S\,E\{a_n^2\} + 2\sqrt{SN}\,E\{a_n w_n\} + N\,E\{w_n^2\}$$
$$= S + N. \tag{2.77}$$

Substituting (2.75) and (2.77) into (2.76), one obtains

$$\mathrm{Var}\{x_n\} \approx (S+N) - S = N,$$

where the result is approximate due to the assumption of no receiver errors. The SNV RxDA SNR estimator for coherent BPSK in real AWGN may then be expressed as

$$\hat{\rho}_{\text{SNV RxDA,real}} = \frac{E\{x_n\}^2}{\mathrm{Var}\{x_n\}} = \frac{E\{|y_n|\}^2}{\mathrm{Var}\{|y_n|\}}. \tag{2.78}$$

In practice, ergodicity is assumed to hold so that the statistical mean, $E\{x_n\}$, and the statistical variance, $\mathrm{Var}\{x_n\}$, in (2.78) may be replaced with the sample mean, $\bar{x}$, and the unbiased sample variance, s_x^2, respectively. For a finite number of received symbols, N_{sym}, the sample mean is evaluated as

$$\bar{x} = \frac{1}{N_{\text{sym}}} \sum_{n=0}^{N_{\text{sym}}-1} x_n,$$

and the unbiased sample variance is

$$s_x^2 = \frac{1}{N_{\text{sym}}-1} \sum_{n=0}^{N_{\text{sym}}-1} x_n^2 - \frac{N_{\text{sym}}}{N_{\text{sym}}-1}\bar{x}^2$$

so that the in-service, SNV RxDA SNR estimator for coherent BPSK in real AWGN based on sample statistics may be expressed as

$$\hat{\rho}_{\text{SNV RxDA,real}} = \frac{\bar{x}^2}{s_x^2} = \frac{\left[\frac{1}{N_{sym}}\sum_{n=0}^{N_{sym}-1}|y_n|\right]^2}{\frac{1}{N_{sym}-1}\sum_{n=0}^{N_{sym}-1} y_n^2 - \frac{1}{N_{sym}(N_{sym}-1)}\left[\sum_{n=0}^{N_{sym}-1}|y_n|\right]^2}. \tag{2.79}$$

This is essentially the result of [46].

Consider now SNV *TxDA* SNR estimation where a copy of the transmitted data sequence, a_n, is made available to the receiver. In a similar manner as shown previously by (2.74), define a variable, x_n, as

$$x_n = y_n\,a_n \tag{2.80}$$

where a_n represents the true transmitted symbols. Again invoking the independence of the signal and noise sequences allows the mean of (2.80) to be evaluated as

$$E\{x_n\} = \sqrt{S}\,E\{a_n^2\} + \sqrt{N}\,E\{w_n\}E\{a_n\} = \sqrt{S}.$$

The variance of x_n is evaluated as

$$\begin{aligned}
\mathrm{Var}\{x_n\} &= E\{x_n^2\} - E\{x_n\}^2 \\
&= E\{y_n^2 a_n^2\} - S \\
&= E\{y_n^2\} - S \\
&= (S+N) - S \;=\; N.
\end{aligned}$$

Note that $E\{y_n^2 a_n^2\} = E\{y_n^2\}$ since $a_n = \pm 1$ for real BPSK signals. The SNV TxDA SNR estimator may be expressed in terms of the received sequence, y_n, and the known transmitted sequence, a_n, as

$$\hat{\rho}_{\mathrm{SNV\,TxDA,real}} = \frac{E\{x_n\}^2}{\mathrm{Var}\{x_n\}} = \frac{E\{y_n\,a_n\}^2}{\mathrm{Var}\{y_n\,a_n\}}.$$

In terms of sample statistics, the SNV TxDA SNR estimator is expressed as

$$\hat{\rho}_{\mathrm{SNV\,TxDA,real}} = \frac{\left[\frac{1}{N_{sym}}\sum_{n=0}^{N_{sym}-1} y_n\,a_n\right]^2}{\frac{1}{N_{sym}-1}\sum_{n=0}^{N_{sym}-1} y_n^2 - \frac{1}{N_{sym}(N_{sym}-1)}\left[\sum_{n=0}^{N_{sym}-1} y_n\,a_n\right]^2}. \tag{2.81}$$

This SNV TxDA SNR estimator is a simple extension of the SNV RxDA SNR estimator presented in [46]. There is no mention of it in the literature, but its performance is of interest and it can be used to advantage in systems where known sequences are inserted into the transmitted data.

The SNV expressions given by (2.79) and (2.81) are remarkably similar to the ML expressions given by (2.61) and (2.58), respectively. In fact, the structure of the intuitive SNV estimator is identical to the structure of the ML estimator operating on the samples *after* the MF. The ML method may be applied to those baud-spaced samples, given by (2.22), because in the context of the system model described in

Section 2.2, the information symbols are ISI-free and the noise is white at those instants.

Replacing K with N_{sym}, r_k (the received signal before the MF) with y_n (the baud-spaced received signal after the MF), and m_k with a_n in the ML expression given by (2.58), and acknowledging that $N_{\text{ss}} = 1$ in the SNV case, one obtains a reduced-bias, SNV TxDA SNR estimator for coherent BPSK in real AWGN as

$$\hat{\rho}'_{\text{SNV TxDA,real}} = \frac{\left[\frac{1}{N_{\text{sym}}} \sum_{n=0}^{N_{\text{sym}}-1} y_n\, a_n\right]^2}{\frac{1}{N_{\text{sym}}-3} \sum_{n=0}^{N_{\text{sym}}-1} y_n^2 - \frac{1}{N_{\text{sym}}(N_{\text{sym}}-3)} \left[\sum_{n=0}^{N_{\text{sym}}-1} y_n\, a_n\right]^2}. \tag{2.82}$$

This expression differs from (2.81) only in that the factor $1/(N_{\text{sym}} - 3)$ appears in the denominator of (2.82), whereas the factor $1/(N_{\text{sym}} - 1)$ appears in the denominator of (2.81). In a similar manner, one obtains a reduced-bias form of the SNV RxDA SNR estimator for coherent BPSK in real AWGN as

$$\hat{\rho}'_{\text{SNV RxDA,real}} = \frac{\left[\frac{1}{N_{\text{sym}}} \sum_{n=0}^{N_{\text{sym}}-1} |y_n|\right]^2}{\frac{1}{N_{\text{sym}}-3} \sum_{n=0}^{N_{\text{sym}}-1} y_n^2 - \frac{1}{N_{\text{sym}}(N_{\text{sym}}-3)} \left[\sum_{n=0}^{N_{\text{sym}}-1} |y_n|\right]^2}. \tag{2.83}$$

The reduced-bias expressions given by (2.82) and (2.83) are the expressions corresponding to the SNV TxDA SNR estimator and the in-service, SNV RxDA SNR estimator, respectively, used in this study for coherent BPSK in real AWGN. These expressions can be used unchanged in any coherent BPSK system that employs a root-Nyquist filter (not necessarily a RRC filter) in the transmitter and receiver.

Derivation of the SNV SNR Estimator for Complex Channels

In the above derivation of the SNV SNR estimator for real channels, it was found that the SNV estimator is actually the ML SNR estimator operating on the baud-spaced samples given by (2.22). Using this fact allows one to write the reduced-bias, SNV TxDA SNR estimator for coherent M-ary PSK in complex AWGN as

$$\hat{\rho}'_{\text{SNV TxDA,complex}} = \frac{\left[\frac{1}{N_{\text{sym}}} \sum_{n=0}^{N_{\text{sym}}-1} \text{Re}\left\{y_n^* a_n\right\}\right]^2}{\frac{1}{N_{\text{sym}}-1} \sum_{n=0}^{N_{\text{sym}}-1} |y_n|^2 - \frac{1}{N_{\text{sym}}(N_{\text{sym}}-1)} \left[\sum_{n=0}^{N_{\text{sym}}-1} \text{Re}\left\{y_n^* a_n\right\}\right]^2} \tag{2.84}$$

which is an adaptation of the reduced-bias, complex ML TxDA SNR estimator given by (2.72).

Similarly, an in-service, reduced-bias, SNV RxDA SNR estimator for coherent M-ary PSK in complex AWGN may be expressed as

$$\hat{\rho}'_{\text{SNV RxDA,complex}} = \frac{\left[\frac{1}{N_{sym}} \sum_{n=0}^{N_{sym}-1} \text{Re}\{y_n^* a_n^{(\hat{i})}\}\right]^2}{\frac{1}{N_{sym}-1} \sum_{n=0}^{N_{sym}-1} |y_n|^2 - \frac{1}{N_{sym}(N_{sym}-1)} \left[\sum_{n=0}^{N_{sym}-1} \text{Re}\{y_n^* a_n^{(\hat{i})}\}\right]^2} \tag{2.85}$$

which is an adaptation of the in-service, reduced-bias, complex ML RxDA SNR estimator given by (2.73). In (2.85), $\hat{i}$ denotes the $\hat{i}^{\text{th}}$ sequence of $M^{N_{sym}}$ possible transmitted symbol sequences formed from receiver decisions.

Equations (2.84) and (2.85) are the SNV estimators used in this study for complex channels. These expressions can be used unchanged in any coherent M-ary PSK system that employs a root-Nyquist filter (not necessarily a RRC filter) in the transmitter and receiver, and they can be extended to other forms of coherent modulation such as QAM.

2.3.4 The Second- and Fourth-Order Moments (M_2M_4) Estimator

The earliest mention of the application of the second and fourth moments to SNR estimation was in 1967 by Benedict and Soong [42] who consider only the separate estimation of carrier strength and noise level (as opposed to the estimation of the *ratio* of carrier strength to noise level) in AWGN. No derivation details are provided. The estimation of signal power to noise power using the method of second and fourth moments is treated in papers by Matzner [53], and Matzner and Engleberger [54] (both published over two decades after the work of Benedict and Soong) in which more derivation details are provided. A patent application has been submitted as a result of the modern-day work [69]. Mathematically, the modern-day derivation yields the same expressions as those given in [42].

The results given in [42] and [53] are based on the assumption of complex signals in complex AWGN, while in [54] an overview is given of the real case using a different approach. The detailed derivation of the M_2M_4 estimator for real channels is offered here first, followed by a sketch of the derivation of the M_2M_4 estimator for complex channels. The development is very similar to that given in [53]. No precise name or acronym has previously been assigned to this technique, so "M_2M_4" is used for easy reference.

Derivation of the M_2M_4 SNR Estimator for Real Channels

Consider the sampled output of the MF receiver previously given by (2.22) as

$$y_n = \sqrt{S}\, a_n + \sqrt{N}\, w_n$$

where all quantities are assumed real. Let M_2 denote the second moment of y_n as

$$\begin{aligned} M_2 &= E\{y_n^2\} \\ &= E\{(\sqrt{S}\, a_n + \sqrt{N}\, w_n)^2\} \\ &= S\, E\{a_n^2\} + 2\sqrt{SN}\, E\{a_n w_n\} + N\, E\{w_n^2\}. \end{aligned}$$

Since the desired signal and noise each have zero mean and are mutually independent, M_2 reduces to

$$M_2 = S + N. \tag{2.86}$$

Similarly, let M_4 denote the fourth moment of y_n as

$$\begin{aligned} M_4 &= E\{y_n^4\} \\ &= S^2\, E\{a_n^4\} + 4S\sqrt{SN}\, E\{a_n^3 w_n\} + 6SN\, E\{a_n^2 w_n^2\} \\ &\quad + 4N\sqrt{SN}\, E\{a_n w_n^3\} + N^2 E\{w_n^4\}. \end{aligned}$$

Again, due to the statistical independence of the desired signal and noise, the expected values of the products of signal and noise may be written as the products of the

expected values. Thus

$$\begin{aligned} E\{a_n^3 w_n\} &= E\{a_n^3\}\, E\{w_n\} = 0 \\ E\{a_n w_n^3\} &= E\{a_n\}\, E\{w_n^3\} = 0 \\ E\{a_n^2 w_n^2\} &= E\{a_n^2\}\, E\{w_n^2\} = 1. \end{aligned}$$

Let k_a be the kurtosis of a_n, and k_w be the kurtosis of w_n expressed as

$$k_a = \frac{E\{a_n^4\}}{E\{a_n^2\}^2} \tag{2.87}$$

$$k_w = \frac{E\{w_n^4\}}{E\{w_n^2\}^2}, \tag{2.88}$$

then M_4 may be written simply as

$$M_4 = k_a S^2 + 6SN + k_w N^2. \tag{2.89}$$

Equations (2.86) and (2.89) form a system of two equations in two unknowns, S and N. The solution of the system of equations yields an expression for the signal power estimator for real channels as

$$\hat{S}_{M_2M_4,\text{real}} = \frac{M_2(k_w - 3) \pm \sqrt{(9 - k_a k_w){M_2}^2 + M_4(k_a + k_w - 6)}}{(k_a + k_w - 6)}, \tag{2.90}$$

and an expression for the noise power estimator for real channels as

$$\hat{N}_{M_2M_4,\text{real}} = M_2 - \hat{S}_{M_2M_4,\text{real}}. \tag{2.91}$$

The M_2M_4 SNR estimator for real channels is the ratio

$$\hat{\rho}_{M_2M_4,\text{real}} = \frac{\hat{S}_{M_2M_4,\text{real}}}{\hat{N}_{M_2M_4,\text{real}}}. \tag{2.92}$$

For BPSK in real AWGN, $k_a = 1$ and $k_w = 3$ so that (2.92) simplifies to

$$\hat{\rho}_{M_2M_4} = \frac{\frac{1}{2}\sqrt{6{M_2}^2 - 2M_4}}{M_2 - \frac{1}{2}\sqrt{6{M_2}^2 - 2M_4}}, \tag{2.93}$$

where the negative root of (2.90) is chosen so that $\hat{\rho}_{M_2M_4,\text{real}}$ is positive.

Derivation of the M_2M_4 SNR Estimator for Complex Channels

For the complex case, M_2 and M_4 must be re-derived assuming the signal given by (2.22) is complex. The second moment in this case is expressed as

$$\begin{aligned} M_2 &= E\{y_n y_n^*\} \\ &= S\,E\{|a_n|^2\} + \sqrt{SN}\,E\{a_n w_n^*\} + \sqrt{SN}\,E\{a_n^* w_n\} + N\,E\{|w_n|^2\}, \end{aligned} \tag{2.94}$$

and the fourth moment is given by

$$\begin{aligned} M_4 &= E\{(y_n y_n^*)^2\} \\ &= S^2\,E\{|a_n|^4\} + 2S\sqrt{SN}(E\{|a_n|^2 a_n w_n^*\} + E\{|a_n|^2 a_n^* w_n\}) \\ &\quad + SN(E\{(a_n w_n^*)^2\} + 4E\{|a_n|^2|w_n|^2\} + E\{(a_n^* w_n)^2\}) \\ &\quad + 2N\sqrt{SN}(E\{|w_n|^2 a_n w_n^*\} + E\{|w_n|^2 a_n^* w_n\}) + N^2\,E\{|w_n|^4\}. \end{aligned} \tag{2.95}$$

Again, assuming the signal and noise are zero-mean, independent random processes, and assuming the in-phase and quadrature components of the noise are mutually independent, (2.94) and (2.95) reduce to

$$M_2 = S + N \tag{2.96}$$

(as in the real case), and

$$M_4 = k_a\,S^2 + 4SN + k_w\,N^2, \tag{2.97}$$

respectively, where $k_a = E\{|a_n|^4\}/E\{|a_n|^2\}^2$ and $k_w = E\{|w_n|^4\}/E\{|w_n|^2\}^2$ are the kurtosis of the complex signal and the kurtosis of the complex noise, respectively. Solving for S and N, one obtains the M_2M_4 estimators for the signal power and noise power for complex channels as

$$\hat{S}_{M_2M_4,\text{complex}} = \frac{M_2(k_w - 2) \pm \sqrt{(4 - k_a k_w){M_2}^2 + M_4(k_a + k_w - 4)}}{k_a + k_w - 4} \tag{2.98}$$

and

$$\hat{N}_{M_2M_4,\text{complex}} = M_2 - \hat{S}_{M_2M_4,\text{complex}}. \tag{2.99}$$

Equation (2.99) is equivalent in form to (2.91).

For any M-ary PSK signal, $k_a = 1$, and for complex noise, $k_w = 2$, so that the M_2M_4 estimator for this special case reduces to

$$\hat{\rho}_{M_2M_4,\text{complex}} = \frac{\hat{S}_{M_2M_4,\text{complex}}}{\hat{N}_{M_2M_4,\text{complex}}} = \frac{\sqrt{2M_2{}^2 - M_4}}{M_2 - \sqrt{2M_2{}^2 - M_4}}, \tag{2.100}$$

where the negative root in (2.98) is selected so that $\hat{\rho}_{M_2M_4,\text{complex}}$ is positive.

For the purpose of implementation, M_2 and M_4 are approximated by their respective time averages [54] as

$$M_2 = \frac{1}{N_{\text{sym}}} \sum_{n=0}^{N_{\text{sym}}-1} |y_n|^2, \tag{2.101}$$

and

$$M_4 = \frac{1}{N_{\text{sym}}} \sum_{n=0}^{N_{\text{sym}}-1} |y_n|^4, \tag{2.102}$$

which, as written, are applicable to both the real and complex cases.

The M_2M_4 estimator is of the in-service type, and has the advantage that knowledge of the transmitted data need not be known or estimated (which is why the TxDA and RxDA labels are not applied to this estimator). The M_2M_4 estimator given by (2.100) is applicable to any coherent or differentially coherent M-PSK modulation, and is easily extended to systems employing other modulations such as QAM. Note that carrier phase is not typically recovered in differentially coherent systems. The M_2M_4 estimator has the advantage that carrier phase recovery is not required (since the second and fourth moments of the received signal carry no phase information) which is why it is applicable to both coherent and differentially coherent systems. The expressions given by (2.93) and (2.100) may also be applied to systems in which the system impulse response is different from that specified in Section 2.2 as long as the transmit and receive filters are root-Nyquist.

2.3.5 The Signal-to-Variation Ratio (SVR) Estimator

In 1994, Brandão *et al* [56] presented the SVR estimator as a means to measure the quality of a received signal in a multipath fading channel. Though intended

for (complex) fading channels, this estimator may be adapted to any coherent or differentially coherent M-ary PSK modulated signal, including BPSK in real AWGN. It is not applicable, in general, to other modulation schemes such as QAM. This method is based on the autocorrelation properties of the square of the sampled output of the MF. The derivation of the SVR SNR estimator for BPSK in real AWGN is presented first, followed by an extension to M-ary PSK in complex AWGN. The derivation presented here is similar to the development given in [56].

Derivation of the SVR SNR Estimator for Real Channels

Recall the expression for the sampled output of the MF from (2.22) written as

$$y_n = \sqrt{S}\, a_n + \sqrt{N}\, w_n,$$

assuming real quantities. The SVR estimator is a function of the parameter, β, which is expressed for real signals as

$$\beta = \frac{E\{y_n^2 y_{n-1}^2\}}{E\{y_n^4\} - E\{y_n^2 y_{n-1}^2\}}. \tag{2.103}$$

After expansion of the right-hand side of (2.103) using (2.22), an expression may be found for $\rho = S/N$ as a function of β. This expression is the SVR estimator for real AWGN channels, $\hat{\rho}_{\mathrm{SVR,real}}$, which generates SNR estimates from measurements of β.

The term, $E\{y_n^4\}$, may be recognized as M_4 of the M_2M_4 estimator derived in Section 2.3.4, and given by (2.89) as

$$M_4 = k_a S^2 + 6SN + k_w N^2.$$

In order to expand the other term in (2.103), consider the product

$$\begin{aligned} y_n^2 y_{n-1}^2 &= (Sa_n^2 + 2\sqrt{SN}a_n w_n + Nw_n^2)(Sa_{n-1}^2 + 2\sqrt{SN}a_{n-1}w_{n-1} + Nw_{n-1}^2) \\ &= S^2 a_n^2 a_{n-1}^2 + 2S\sqrt{SN}\, a_n^2 a_{n-1} w_{n-1} \\ &\quad + SNa_n^2 w_{n-1}^2 + 2S\sqrt{SN}\, a_n a_{n-1}^2 w_n \\ &\quad + 4SNa_n a_{n-1} w_n w_{n-1} + 2N\sqrt{SN}\, a_n w_n w_{n-1}^2 \\ &\quad + SNa_{n-1}^2 w_n^2 + 2N\sqrt{SN}\, a_{n-1} w_{n-1} w_n^2 + N^2 w_n^2 w_{n-1}^2. \end{aligned} \tag{2.104}$$

In order to evaluate the expected value of (2.104), note that considerable simplification results from the fact that the expected value of any product with a single signal or noise term vanishes due to the assumed mutual independence of the signal and noise. Exploiting the independence properties of the signal and noise further, one has

$$E\{a_n^2 a_{n-1}^2\} = E\{a_n^2\}\, E\{a_{n-1}^2\} = \sigma_a^4 = 1$$
$$E\{a_n^2 w_{n-1}^2\} = E\{a_{n-1}^2 w_n^2\} = \sigma_a^2 \sigma_w^2 = 1$$
$$E\{w_n^2 w_{n-1}^2\} = E\{w_n^2\}\, E\{w_{n-1}^2\} = \sigma_w^4 = 1$$
$$E\{a_n a_{n-1} w_n w_{n-1}\} = E\{a_n a_{n-1}\}\, E\{w_n w_{n-1}\} = 0.$$

Using these relations allows the expected value of (2.104) to be written as

$$E\{y_n^2 y_{n-1}^2\} = S^2 + 2SN + N^2. \tag{2.105}$$

Substituting (2.89) and (2.105) into (2.103), one obtains

$$\begin{aligned} \beta &= \frac{S^2 + 2SN + N^2}{(k_a - 1)S^2 + 4SN + (k_w - 1)N^2} \\ &= \frac{\rho^2 + 2\rho + 1}{(k_a - 1)\rho^2 + 4\rho + (k_w - 1)}. \end{aligned} \tag{2.106}$$

The SVR estimator is then found by solving (2.106) for ρ in terms of β yielding the general expression for real channels as

$$\hat{\rho}_{\text{SVR,real}} = \frac{(2\beta - 1) \pm \sqrt{(2\beta - 1)^2 - [1 - \beta(k_a - 1)][1 - \beta(k_w - 1)]}}{1 - \beta(k_a - 1)}. \tag{2.107}$$

Specifically, recall from Section 2.3.4 that $k_a = 1$ for BPSK signals and $k_w = 3$ for real AWGN so that, taking the positive root, (2.107) simplifies to

$$\hat{\rho}_{\text{SVR,real}} = (2\beta - 1) + \sqrt{2\beta(2\beta - 1)}. \tag{2.108}$$

Derivation of the SVR SNR Estimator for Complex Channels

In the complex case, the SVR estimator is a function of the parameter β expressed as

$$\beta = \frac{E\{y_n y_n^* y_{n-1} y_{n-1}^*\}}{E\{(y_n y_n^*)^2\} - E\{y_n y_n^* y_{n-1} y_{n-1}^*\}} \tag{2.109}$$

where y_n is given by (2.22). The term, $E\{(y_n y_n^*)^2\}$, is recognized as M_4 of the complex form of the M_2M_4 estimator of Section 2.3.4 given by (2.97) as

$$M_4 = k_a S^2 + 4SN + k_w N^2.$$

The other term in (2.109) simplifies to

$$E\{y_n y_n^* y_{n-1} y_{n-1}^*\} = S^2 + 2SN + N^2, \tag{2.110}$$

again exploiting the independence properties of the signal and noise samples. Substituting (2.97) and (2.110) into (2.109), one obtains

$$\beta = \frac{\rho^2 + 2\rho + 1}{(k_a - 1)\rho^2 + 2\rho + (k_w - 1)} \tag{2.111}$$

which may be solved for ρ to yield the general SVR estimator for complex channels as

$$\hat{\rho}_{\text{SVR, complex}} = \frac{(\beta - 1) \pm \sqrt{(\beta - 1)^2 - [1 - \beta(k_a - 1)][1 - \beta(k_w - 1)]}}{1 - \beta(k_a - 1)}. \tag{2.112}$$

Specifically, recall from Section 2.3.4 that $k_a = 1$ for M-ary PSK signals and $k_w = 2$ for complex AWGN so that, taking the positive root, (2.112) simplifies to

$$\hat{\rho}_{\text{SVR, complex}} = \beta - 1 + \sqrt{\beta(\beta - 1)}. \tag{2.113}$$

In practice, β is computed in terms of time averages as

$$\beta = \frac{\frac{1}{N_{sym}} \sum_{n=0}^{N_{sym}-1} |y_n|^2 |y_{n-1}|^2}{\frac{1}{N_{sym}} \sum_{n=0}^{N_{sym}-1} |y_n|^4 - \frac{1}{N_{sym}} \sum_{n=0}^{N_{sym}-1} |y_n|^2 |y_{n-1}|^2} \tag{2.114}$$

which, as written, could be applied to both real and complex channels.

Just as the M_2M_4 estimator, the SVR estimator is of the in-service type, and has the advantage that the transmitted data need not be known or estimated. The SVR estimator given by (2.113) is applicable to any coherent or differentially coherent M-PSK modulation. As mentioned in Section 2.3.4, in differentially coherent systems, the carrier phase typically is not recovered, but this is not a problem for the SVR

estimator since the phase information of the signal is removed in the estimation process. As mentioned earlier in this section, the SVR estimator is not easily extended to systems employing other modulations. The expressions given by (2.108) and (2.113) may be applied to systems in which the system impulse response is different from that specified in Section 2.2, as long as the transmit and receive filters are root-Nyquist.

2.4 Other SNR Estimators of Interest but not Included in this Study

This section lists a few SNR estimation strategies which are mentioned for completeness, but which were judged to be unsuitable for this study. The reasons for excluding these estimators are given in the brief descriptions that follow.

2.4.1 Use of Limiters for Estimating SNR

Nahi and Gagliardi [1] discuss an unusual method of SNR estimation which is accomplished with a hard limiter, a filter and a power meter. The method exploits the autocorrelation properties of the desired signal and noise. In [1], the assumption is made that both the desired signal and noise are Gaussian stochastic processes. There is no mention in [1] or elsewhere in the literature of the application of this method to discrete, random signals corrupted by AWGN. Although the method is attractive since the measurement of SNR is accomplished by a simple power measurement, it was judged unsuitable for this study of SNR estimators based on the results reported in [1] as explained below.

In [1], the desired signal and noise are Gaussian stochastic processes having different autocorrelation functions (the autocorrelation functions cannot be identical in order for this method to work). The SNR is related to the average power, P_{out}, of the filtered output of the corrupted signal after hard limiting. An expression for P_{out} may be derived explicitly in terms of the SNR, but the expression is not easily inverted. That is, the SNR cannot be expressed explicitly as a function of P_{out} so that the SNR

must either be solved implicitly by iterative techniques, or found with the use of a look-up table.

A more serious problem is the fact identified in [1] that the relationship between P_{out} and SNR/(SNR + 1) is linear, at least for the case in which both the desired signal and noise are Gaussian stochastic processes. At a first glance, this linear relationship seems encouraging since it is easy to estimate empirically a straight line which can then be used to convert measurements of P_{out} to SNR. However, consider Figure 2.3. Figure 2.3(a) displays some arbitrary linear relationship between P_{out} and SNR/(SNR+1). Figure 2.3(b) shows P_{out} plotted as a function of SNR. It can be seen that for SNR $>$ 10 dB (almost the entire SNR range of interest), P_{out} is essentially "saturated"; that is, even a small inaccuracy in the measurement of P_{out} can result in a very large error in the estimation of the SNR, as the 5% error bars demonstrate.

Though these results are based on Gaussian stochastic transmitted signals, simulation results indicate a similar sensitivity in the case of binary, discrete random signals. With this degree of sensitivity, a very large number of samples is required to obtain reasonable SNR estimates—the number of required samples is orders of magnitude more than the number of samples required by the other estimators described to this point for the same quality estimate. This method was not considered further since it was judged to be too impractical.

2.4.2 Autocorrelation Method of SNR Estimation

This SNR estimation method is also based on the autocorrelation properties of the desired signal and noise. A discussion of this particular estimator has not been found in the literature.

The estimate of the SNR is derived from the samples, y_k, of the output of the MF which, from (2.5), are expressed as

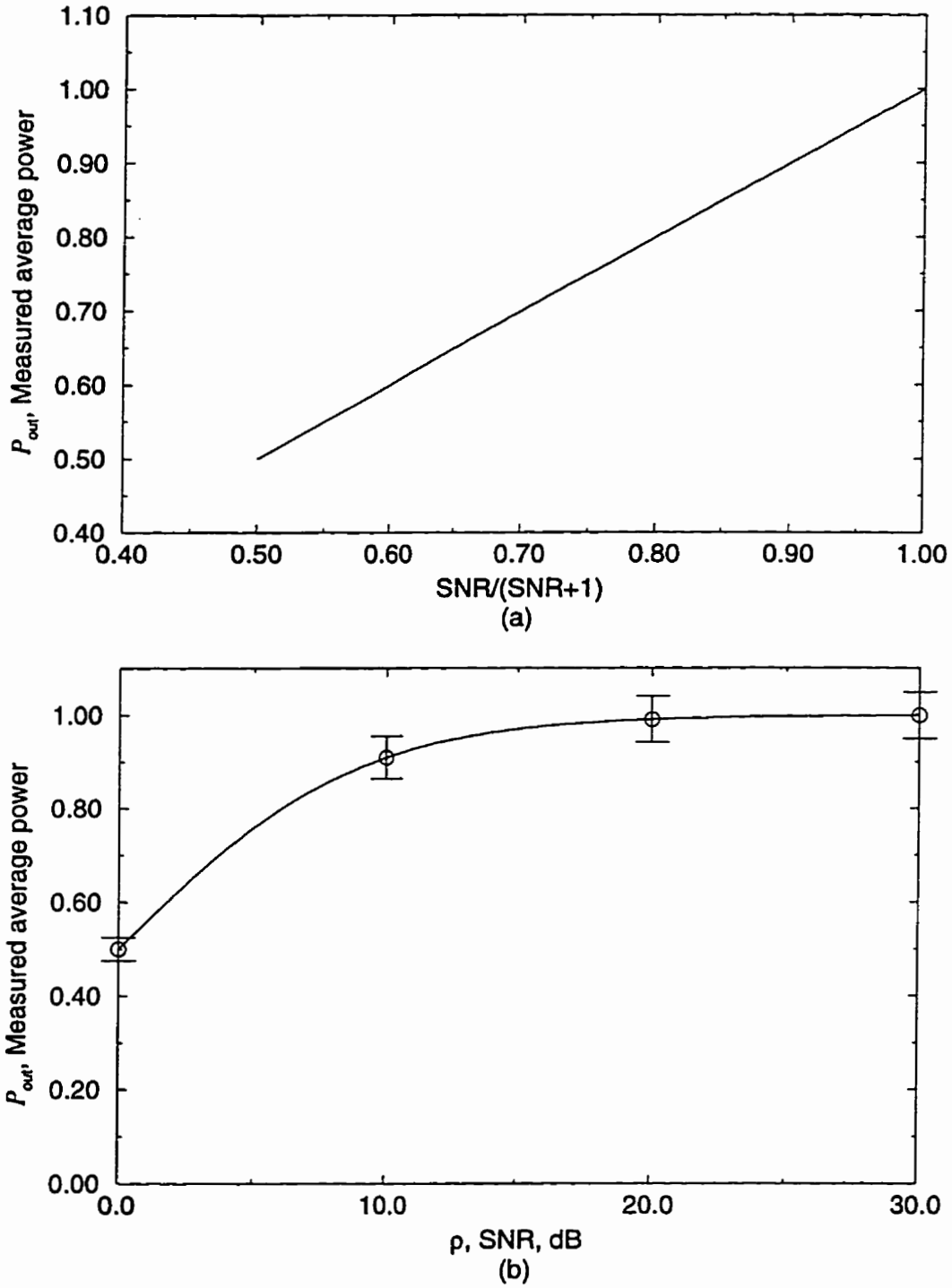

Figure 2.3: Performance of the SNR estimator described in [1]: (a) measured average power as a function of SNR/(SNR+1), (b) measured average power as a function of SNR.

$$\begin{aligned} y_k &= (\sqrt{S}\, m_k + \sqrt{N}\, z_k) \otimes h^*_{-k} \\ &= \sqrt{S} \sum_l h_l m_{k-l} + \sqrt{N} \sum_l h_l z_{k-l}. \end{aligned}$$

Let the filtered samples of the desired signal be denoted by ζ_k and, as before, let w_k denote the coloured noise samples so that the output of the MF may be expressed as

$$y_k = \sqrt{S}\, \zeta_k + \sqrt{N}\, w_k.$$

The normalized autocorrelation of y, denoted by φ_y, is required in order to proceed with this method. An expression for φ_y is derived below by considering the autocorrelations of the desired signal and noise separately.

First, consider the autocorrelation of the desired signal. Recall from (2.15) that the autocorrelation of the transmitted signal, m_k, can be expressed as

$$R_m(l) = \frac{\sigma_a^2}{N_{ss}} g_l = \frac{1}{N_{ss}} g_l.$$

Since ζ is the result of the discrete convolution of m_k with h_k, the autocorrelation of ζ is evaluated as

$$\begin{aligned} R_\zeta(l) &= h_l \otimes h_{-l} \otimes R_m(l) \\ &= g_l \otimes \left(\frac{1}{N_{ss}} g_l \right) \\ &= \frac{\alpha}{N_{ss}} \left(\frac{g_l \otimes g_l}{\alpha} \right) \\ &= \frac{\alpha}{N_{ss}} \varphi_\zeta(l) \end{aligned} \tag{2.115}$$

where $\varphi_\zeta(l)$ is the normalized autocorrelation of ζ, and

$$\begin{aligned} \alpha &= g_l \otimes g_l|_{l=0} \\ &= \sum_k g_k g_{l-k}|_{l=0} \\ &= \sum_k g_k^2. \end{aligned}$$

From (2.9b), the normalized autocorrelation of the coloured noise can be written as

$$\varphi_w(l) = g_l. \tag{2.116}$$

Using (2.115) and (2.116), the normalized autocorrelation of y may be written as

$$\begin{aligned}\varphi_y(l) &= \frac{\frac{S\alpha}{N_{ss}}\varphi_\zeta(l) + N\varphi_w(l)}{\frac{S\alpha}{N_{ss}} + N} \\ &= \frac{\rho\frac{\alpha}{N_{ss}}\varphi_\zeta(l) + \varphi_w(l)}{\rho\frac{\alpha}{N_{ss}} + 1}\end{aligned}$$

where, as always, $\rho = S/N$. Solving for ρ, the SNR is obtained as

$$\rho = \frac{N_{ss}}{\alpha}\left(\frac{\varphi_w(l) - \varphi_y(l)}{\varphi_y(l) - \varphi_\zeta(l)}\right). \tag{2.117}$$

Both $\varphi_\zeta(l)$ and $\varphi_w(l)$ are known, as are N_{ss} and α, so that a measurement of the autocorrelation of the output of the MF may be used to find an estimate of the SNR. Consider the special case of $l = nN_{ss}$; that is, consider the optimally-sampled output of the MF. In this case,

$$\varphi_w(nN_{ss}) = g_{nN_{ss}} = 0, \; n \neq 0$$

so that (2.117) simplifies to

$$\rho = \frac{N_{ss}}{\alpha}\left(\frac{\varphi_y(nN_{ss})}{\varphi_\zeta(nN_{ss}) - \varphi_y(nN_{ss})}\right), \; n \neq 0.$$

It is immediately apparent that this SNR estimator will not work well for high SNR since $\varphi_y(nN_{ss})$ approaches $\varphi_\zeta(nN_{ss})$ implying that the measurement of $\varphi_y(nN_{ss})$ would have to be extremely accurate in order to obtain meaningful (finite and positive) SNR estimates. Computer simulation tests of this estimator at moderate values of SNR (around 10 dB) revealed that the number of symbols must be very large (on the order of 10^4 symbols or more) for meaningful SNR estimates to be generated. Apart from this estimator, the one described in Section 2.4.1, and the SSME_1 and SSME_2, all of the other estimators described in Section 2.3 require *many* fewer symbols (on the order of just 10 to 100 symbols) to generate meaningful SNR estimates. Given this fact, this method was not considered further for extensive simulation testing.

2.4.3 Implicit Methods of SNR Estimation

All of the methods discussed so far have been treated explicitly as methods of SNR estimation. Some SNR methods arise implicitly in the discussion of other subject areas. Two examples are provided here.

In the context of blind identification and equalization, Tong *et al* [70, page 346] show how the noise variance may be found from the singular value decomposition of the covariance matrix of the received signal. Since the authors had no need to compute the signal power, no method for signal power estimation is offered in [70], but it is presumed that one may simply estimate the total signal-plus-noise power (by averaging the squares of the noise-corrupted received samples), and then subtract from this the estimated noise variance (obtained using the method presented in [70]). In this study, only explicit SNR estimators are considered so this implicit method is omitted.

Another example of an implicit SNR estimator is given by Proakis in [65] in the context of equalization. This SNR estimation method is not considered in this study since it is not an explicit method. On page 564 of [65], an expression is offered to compute the SNR at the output of a linear adaptive equalizer based on the MSE criterion. This implicit method assumes the signal power is normalized to unity. No method is offered to compute the SNR if the signal power is not unity and, most importantly, no expression is offered to relate the SNR after the equalizer to the unequalized SNR to facilitate a comparison with all of the other SNR estimation methods studied here.

2.5 Measures of Estimator Performance

The performance of the SNR estimators may be characterized by an analysis of the sample statistics. A few tests based on sample statistics are listed below, followed by the description of a bound on estimator performance.

2.5.1 Tests Based on Sample Statistics

The SNR estimators are assessed by statistical tests of the mean, variance, and mean-squared-error (MSE) of the SNR estimates. Let $\hat{\rho}$ represent estimates of the SNR for a particular estimator, let $f(\hat{\rho})$ be the pdf of the estimates, and let ρ be the true SNR, then the statistical mean is given by

$$E\{\hat{\rho}\} = \int \hat{\rho} f(\hat{\rho})\, d\hat{\rho}, \tag{2.118}$$

the statistical variance is given by

$$\mathrm{Var}\{\hat{\rho}\} = E\{(\hat{\rho} - E\{\hat{\rho}\})^2\} = \int (\hat{\rho} - E\{\hat{\rho}\})^2 f(\hat{\rho})\, d\hat{\rho}, \tag{2.119}$$

and the statistical MSE is given by

$$\mathrm{MSE}\{\hat{\rho}\} = E\{(\hat{\rho} - \rho)^2\} = \int (\hat{\rho} - \rho)^2 f(\hat{\rho})\, d\hat{\rho}. \tag{2.120}$$

The bias of an estimator is defined as

$$\mathrm{Bias}\{\hat{\rho}\} = E\{\hat{\rho}\} - \rho.$$

Note that the MSE reflects both the bias and the variance of the SNR estimates.

The expressions given by (2.118), (2.119), and (2.120) above can be used to generate analytical expressions for the mean, variance, and MSE, respectively, for a given estimator, but it is not always possible to obtain analytical results because either a closed-form expression cannot be obtained for $f(\hat{\rho})$, or the integrals are too unwieldy to evaluate.

On the other hand, the performance of *all* estimators may be characterized by computer simulation to an arbitrary degree of accuracy. The processes underlying the generation of the SNR estimates are assumed ergodic so that the sample mean, sample variance, and sample MSE may be computed rather than attempting to evaluate closed-form expressions for the statistical mean, statistical variance, and statistical MSE. This assumption allows estimator performance to be evaluated by the analysis of a statistically significant number of SNR estimates.

Let $\hat{\rho}_i$ denote one of N_t SNR estimates generated by a particular estimator, then the unbiased sample mean is given by

$$\bar{\hat{\rho}} = \frac{1}{N_t}\sum_{i=1}^{N_t}\hat{\rho}_i,$$

the unbiased sample variance is given by

$$s^2_{\hat{\rho}} = \frac{1}{N_t - 1}\sum_{i=1}^{N_t}(\hat{\rho}_i - \bar{\hat{\rho}})^2,$$

and the unbiased sample MSE is given by

$$\mathrm{MSE}\{\hat{\rho}\} = \frac{1}{N_t}\sum_{i=1}^{N_t}(\hat{\rho}_i - \rho)^2.$$

The sample bias in this case is expressed as

$$\mathrm{Bias}\{\hat{\rho}\} = \bar{\hat{\rho}} - \rho.$$

The "best" estimator is the one that is unbiased (or has the smallest bias) and has the smallest variance and MSE.

2.5.2 The Cramér-Rao Bound (CRB)

The CRB is an analytical lower bound on estimator variance [67, 47]. A general expression for the CRB of the variance of an SNR estimator is presented in Appendix C, and is used below to find the CRB for SNR estimators operating in the channel described in Section 2.2. The derivation of the CRB for real channels is presented first, following a development similar to that provided by Thomas in [47]. Subsequently, the derivation of the CRB for complex channels is presented which does not appear in [47] and has not been presented elsewhere in the literature.

Derivation of the CRB for Real Channels

Recall the likelihood function for real channels appropriate to the model of Section 2.2 given by (2.50) as

$$\Gamma(S, N, i) = -\frac{K}{2}\ln(2\pi N) - \frac{1}{2N}\sum_{k=0}^{K-1}\left(r_k - \sqrt{S}\,m_k^{(i)}\right)^2.$$

Re-write (2.50) in terms of $\rho = S/N$ and N, instead of S and N, to obtain

$$\Gamma(\rho, N, i) = -\frac{K}{2}\ln(2\pi N) - \frac{1}{2N}\sum_{k=0}^{K-1}\left(r_k - \sqrt{\rho N}m_k^{(i)}\right)^2.$$

This form of the likelihood function is substituted into (C.4) of Appendix C so that the CRB for $\hat{\rho}$ may be written as

$$\mathrm{Var}\{\hat{\rho}\} \geq \frac{-\left(\frac{\partial}{\partial\rho}E\{\hat{\rho}\}\right)^2 E\left\{\frac{\partial^2\Gamma}{\partial N^2}\right\}}{E\left\{\frac{\partial^2\Gamma}{\partial\rho^2}\right\}E\left\{\frac{\partial^2\Gamma}{\partial N^2}\right\} - E\left\{\frac{\partial^2\Gamma}{\partial\rho\partial N}\right\}^2}. \tag{2.121}$$

The required second partial derivatives of (2.121) are

$$\begin{aligned}
\frac{\partial^2\Gamma}{\partial\rho^2} &= -\frac{1}{4\sqrt{\rho^3 N}}\sum_{k=0}^{K-1} r_k m_k^{(i)} \\
\frac{\partial^2\Gamma}{\partial N^2} &= \frac{K}{2N^2} - \frac{1}{N^3}\sum_{k=0}^{K-1} r_k^2 + \frac{3}{4}\sqrt{\frac{\rho}{N^5}}\sum_{k=0}^{K-1} r_k m_k^{(i)} \\
\frac{\partial^2\Gamma}{\partial\rho\partial N} &= -\frac{1}{4\sqrt{\rho N^3}}\sum_{k=0}^{K-1} r_k m_k^{(i)},
\end{aligned}$$

where $K = N_{\mathrm{sym}}N_{\mathrm{ss}}$ is the total number of received samples. Using (2.16), the expected values of the partial derivatives may be expressed as

$$\begin{aligned}
E\left\{\frac{\partial^2\Gamma}{\partial\rho^2}\right\} &= -\frac{K}{4\rho N_{\mathrm{ss}}} \\
E\left\{\frac{\partial^2\Gamma}{\partial N^2}\right\} &= -\frac{K(2+\rho/N_{\mathrm{ss}})}{4N^2} \\
E\left\{\frac{\partial^2\Gamma}{\partial\rho\partial N}\right\} &= -\frac{K}{4N\,N_{\mathrm{ss}}}.
\end{aligned}$$

Substituting these results into (2.121), one obtains the CRB of the variance of an SNR estimator operating in real channels as

$$\mathrm{Var}(\hat{\rho}) \geq 2\left(\frac{2\rho}{N_{\mathrm{sym}}} + \frac{\rho^2}{N_{\mathrm{ss}}N_{\mathrm{sym}}}\right)\left(\frac{\partial}{\partial\rho}E\{\hat{\rho}\}\right)^2. \tag{2.122}$$

If $\hat{\rho}$ is unbiased, then $\partial E\{\hat{\rho}\}/\partial\rho = 1$ so that (2.122) reduces to

$$\mathrm{Var}(\hat{\rho}) \geq 2\left(\frac{2\rho}{N_{\mathrm{sym}}} + \frac{\rho^2}{N_{\mathrm{ss}}N_{\mathrm{sym}}}\right). \tag{2.123}$$

Normalizing (2.123) by ρ^2 shows explicitly the asymptotic behaviour of the estimators with increasing SNR. The normalized variance and MSE for an unbiased SNR estimator operating in real channels are bounded by

$$\frac{\mathrm{MSE}\{\hat{\rho}\}}{\rho^2} = \frac{\mathrm{Var}\{\hat{\rho}\}}{\rho^2} \geq 2\left(\frac{2}{\rho N_{\mathrm{sym}}} + \frac{1}{N_{\mathrm{ss}} N_{\mathrm{sym}}}\right) \tag{2.124}$$

where $\mathrm{MSE}\{\hat{\rho}\} = \mathrm{Var}\{\hat{\rho}\}$ because $\hat{\rho}$ is assumed unbiased. At high SNR, the normalized variance (or normalized MSE) is ultimately determined by the number of received independent samples, $K = N_{\mathrm{sym}} N_{\mathrm{ss}}$, used to generate estimates of the SNR. At low SNR, the lower bound on the normalized variance is inversely proportional both to the SNR and the number of symbols used to form an SNR estimate.

No unique expression exists for the CRB of *biased* estimators since it is dependent on the bias of a particular estimator. Consider the biased ML TxDA SNR estimator of Section 2.3.2 for real channels whose expected value is given by (2.57) as

$$E\{\hat{\rho}_{\mathrm{ML\,TxDA,real}}\} = \frac{K}{K-3}\left(\rho + \frac{1}{N_{\mathrm{sym}}}\right).$$

Noting that $\partial E\{\hat{\rho}_{\mathrm{ML\,TxDA,real}}\}/\partial\rho = K/(K-3)$, the normalized CRB for this biased estimator may be written as

$$\frac{\mathrm{Var}\{\hat{\rho}_{\mathrm{ML\,TxDA,real}}\}}{\rho^2} \geq \frac{2K^2}{(K-3)^2}\left(\frac{2}{\rho N_{\mathrm{sym}}} + \frac{1}{N_{\mathrm{ss}} N_{\mathrm{sym}}}\right). \tag{2.125}$$

The ratio of the normalized, biased CRB (using the bias of the ML estimator) given by (2.125) to the normalized, unbiased CRB given by (2.124) is plotted in Figure 2.4 for values of K in the range from 64 to 1024 samples. It can be seen that the difference between the two expressions becomes small very quickly with increasing K.

Consider briefly the issue of estimator efficiency[3]. An "efficient" estimator is one that meets the CRB for any sample size. If an efficient estimator can be proven to exist, then it is the ML estimator. Comparison of (2.125) and (2.124) with the analytical expressions for the biased and reduced-bias ML TxDA SNR estimators for

[3]The following definitions and concepts regarding efficiency and other properties of ML estimators may be found in [67].

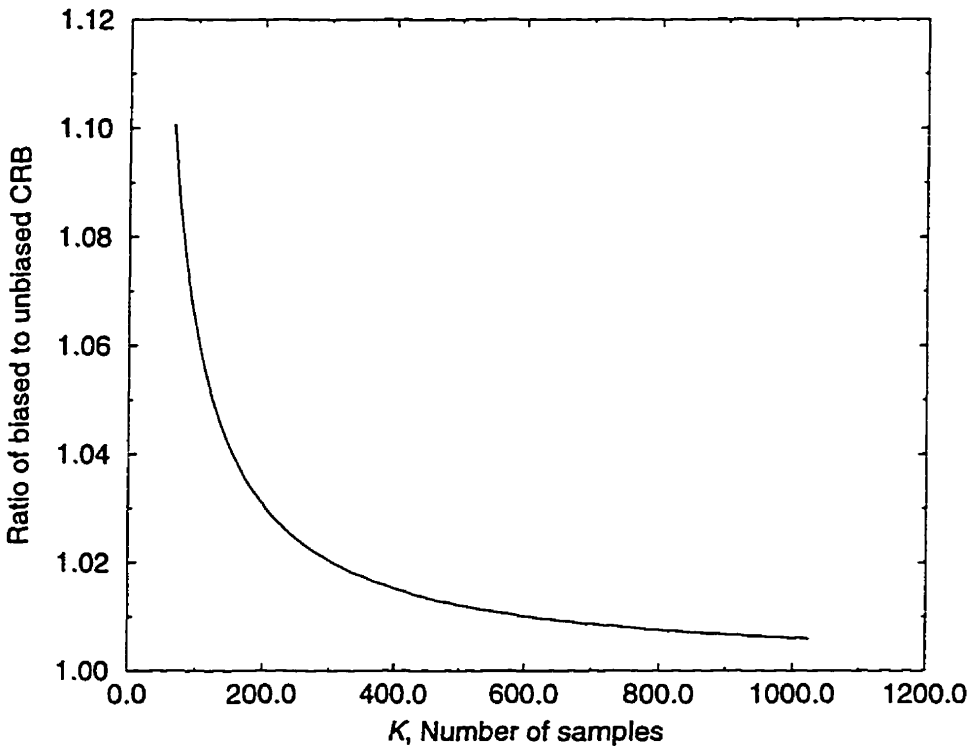

Figure 2.4: Ratio of biased CRB (based on bias of ML TxDA estimator) to unbiased CRB.

real channels given by (2.59) and (2.60), respectively, reveals that the CRB is not met in either case, so an efficient estimator does not exist (otherwise, it would be the ML estimator). Since an efficient estimator does not exist, there may exist some other estimator which has a variance smaller than that of the ML estimator; however, since the CRB and the analytical variance are so close, it seems that the search for this hypothetical, "better" estimator would have little practical value. Under reasonably general conditions [67], ML estimators are at least asymptotically efficient which means that as the number of independent observed samples increases, the variance of the ML estimator asymptotically approaches the CRB.

To illustrate the asymptotic efficiency of the ML estimator, the normalized, unbi-

ased CRB given by (2.124) and the variance of the reduced-bias ML estimator given by (2.60), normalized to the square of the true SNR, are plotted for $K = 64$ and $K = 1024$ in Figure 2.5. It is clear that, as the sample size increases, the variance of the ML estimator approaches the CRB.

Derivation of the CRB for Complex Channels

The likelihood function applicable to this complex case is given by (2.66) which is re-written as a function of ρ and N to obtain

$$\begin{aligned}\Gamma(\rho, N, i) &= \ln f(\mathbf{r_I}, \mathbf{r_Q}|\rho, N, i) \\ &= -K\ln(\pi N) - \frac{1}{N}\left[\sum_{k=0}^{K-1}\left(r_{I_k} - \sqrt{\rho N}m_{I_k}^{(i)}\right)^2 + \sum_{k=0}^{K-1}\left(r_{Q_k} - \sqrt{\rho N}m_{Q_k}^{(i)}\right)^2\right].\end{aligned} \tag{2.126}$$

The second partial derivatives of (2.126) required for the evaluation of (2.121) are

$$\begin{aligned}\frac{\partial^2\Gamma}{\partial\rho^2} &= -\frac{1}{2\sqrt{\rho^3 N}}\sum_{k=0}^{K-1}(r_{I_k}m_{I_k}^{(i)} + r_{Q_k}m_{Q_k}^{(i)}) \\ \frac{\partial^2\Gamma}{\partial N^2} &= \frac{K}{N^2} - \frac{2}{N^3}\sum_{k=0}^{K-1}(r_{I_k}^2 + r_{Q_k}^2) + \frac{3}{2}\sqrt{\frac{\rho}{N^5}}\sum_{k=0}^{K-1}(r_{I_k}m_{I_k}^{(i)} + r_{Q_k}m_{Q_k}^{(i)}) \\ \frac{\partial^2\Gamma}{\partial\rho\partial N} &= -\frac{1}{2\sqrt{\rho N^3}}\sum_{k=0}^{K-1}(r_{I_k}m_{I_k}^{(i)} + r_{Q_k}m_{Q_k}^{(i)}).\end{aligned}$$

Using (2.16), the expected values of the partial derivatives are expressed as

$$E\left\{\frac{\partial^2\Gamma}{\partial\rho^2}\right\} = -\frac{K}{2\rho N_{ss}} \tag{2.127}$$

$$E\left\{\frac{\partial^2\Gamma}{\partial N^2}\right\} = -\frac{K(2+\rho/N_{ss})}{2N^2} \tag{2.128}$$

$$E\left\{\frac{\partial^2\Gamma}{\partial\rho\partial N}\right\} = -\frac{K}{2NN_{ss}}. \tag{2.129}$$

Substituting these results into (2.121), one obtains the CRB of the variance of an unbiased SNR estimator operating in complex channels as

$$\mathrm{Var}\{\hat{\rho}\} \geq \frac{2\rho}{N_{sym}} + \frac{\rho^2}{N_{ss}N_{sym}}. \tag{2.130}$$

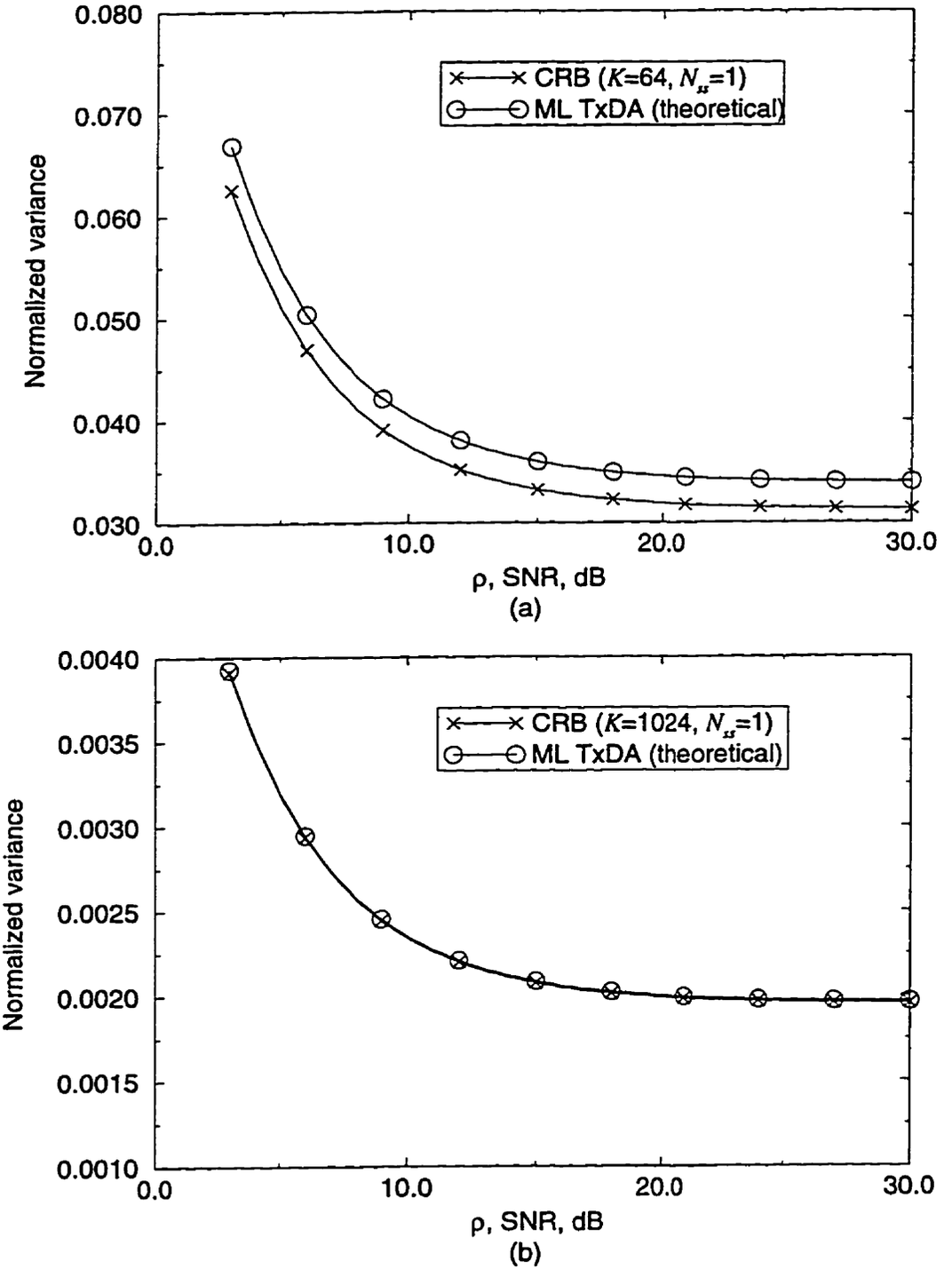

Figure 2.5: Comparison of CRB with theoretical variance of reduced-bias ML TxDA (a) $K = 64$, (b) $K = 1024$.

Therefore, the normalized variance and normalized MSE for an unbiased SNR estimator operating in complex channels are bounded by

$$\frac{\mathrm{MSE}\{\hat{\rho}\}}{\rho^2} = \frac{\mathrm{Var}\{\hat{\rho}\}}{\rho^2} \geq \frac{2}{\rho N_{\mathrm{sym}}} + \frac{1}{N_{\mathrm{ss}} N_{\mathrm{sym}}}, \tag{2.131}$$

where, as mentioned in the derivation of the CRB for the real case, $\mathrm{MSE}\{\hat{\rho}\} = \mathrm{Var}\{\hat{\rho}\}$ because $\hat{\rho}$ is assumed unbiased. Comparing (2.131) with (2.124), it is apparent that the lower bound on SNR estimator variance is a factor of two smaller in complex channels as compared to real channels.

2.6 Simulation Results

The model described in Section 2.2 is implemented in *C* for Monte Carlo simulation. A hard limited uniform random number generator provides the source data and a Gaussian random number generator supplies independent WGN samples. The random number generator algorithms are taken from Chapter 7 of [71].

For a given SNR, an SNR estimate is generated from a block of $N_{\mathrm{sym}} = 64$ or 1024 received symbols. Many such estimates are generated and the bias, variance, and MSE are computed as described in Section 2.5.1 (the plotted bias, variance, and MSE curves are shown normalized to the true SNR). Each estimator from Section 2.3 is tested over an SNR range spanning 3–30 dB in discrete steps. Enough estimates are generated in each case to ensure that the error in the variance is less than 20% with 95% confidence (see Appendix D).

Both real and complex channels are modeled. Specifically, BPSK-modulated signals in real AWGN, and 8-PSK-modulated signals in complex AWGN are simulated. A verification of the simulators is provided in Figure 2.6 by a comparison of the theoretical probabilities of error, P_b, for coherent BPSK and coherent 8-PSK to the simulated results. The solid curve represents the theoretical probability of error for coherent BPSK given by (2.13). The dashed curve represents the theoretical proba-

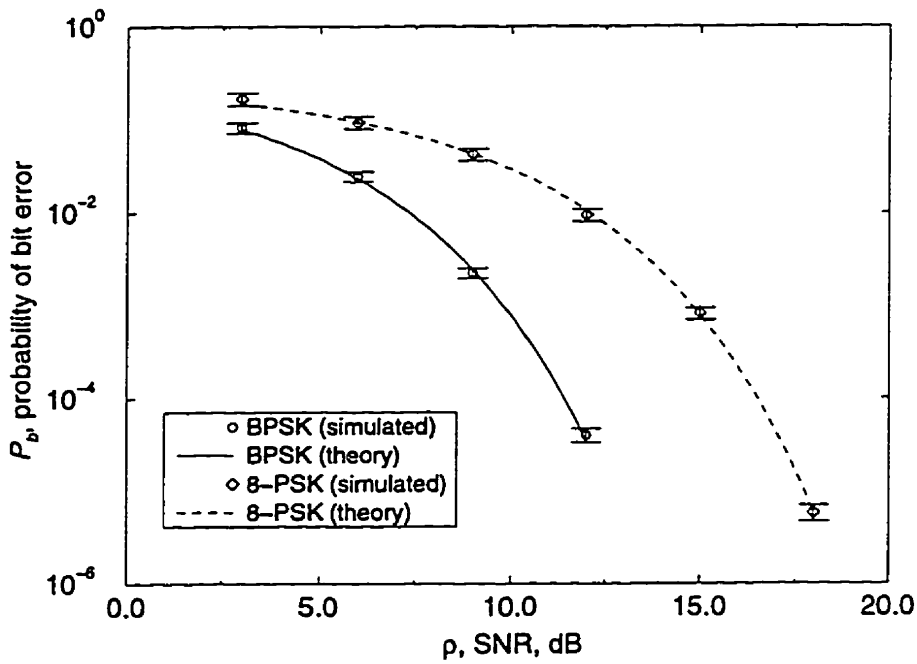

Figure 2.6: Comparison of theoretical and simulated probability of error curves for coherent BPSK in real AWGN and coherent 8-PSK in complex AWGN.

bility of error for coherent 8-PSK approximated by [65]

$$P_{b,8\text{-PSK}} \approx \frac{1}{3}\text{erfc}\left(\sqrt{3\frac{E_b}{N_0}}\sin\frac{\pi}{8}\right).$$

All simulated points are accurate to within at least 20% error with 95% confidence. The SNR, ρ, plotted on the abscissa is equal to $2E_b/N_0$ for BPSK (in real AWGN) and $E_s/N_0 = 3E_b/N_0$ for 8-PSK (in complex AWGN). The simulated points show good agreement with theory.

The unbiased expressions for the CRB, given by (2.124) for real channels and (2.131) for complex channels, are included in the plots of normalized variance and normalized MSE for comparison with the simulated results. The unbiased forms

of the CRB are selected to avoid having the CRB's be estimator specific since the bias for each estimator is different. The CRB is included for reference in the plots of normalized MSE with the understanding that, strictly speaking, the CRB expressions as given in Section 2.5.2 are really bounds on estimator *variance*.

For the ML estimators, the constants to be used in (2.124) and (2.131) are $N_{ss} = 16$ and $N_{sym} = 1024$ or 64. A CRB expression that uses $N_{ss} = 16$ will be referred to here as a *pre-MF CRB* since the ML estimator operates on the $N_{ss} = 16$ samples per symbol at the input to the MF. The SSME algorithms also use 16 samples per symbol, except they are taken at the *output* of the MF. Strictly speaking, the variance and MSE of the SSME algorithms should be compared to a different CRB that takes into account the correlation of the noise samples after the MF. It is more interesting, however, to compare the SSME results to the ML estimator and pre-MF CRB, so the more complicated, multiple-sample per symbol CRB with correlated noise samples is not presented here.

For the estimators that operate on the one sample per symbol (taken at the optimal sampling instant) at the output of the MF (or, equivalently, at the input to the decision-device), the constants to be used in (2.124) and (2.131) are $N_{ss} = 1$ and $N_{sym} = 1024$ or 64. A CRB expression that uses $N_{ss} = 1$ will be referred to here as a *post-MF CRB*. The expressions given by (2.124) and (2.131) are applicable to the post-MF samples in this case because the baud-spaced noise samples at the output of the root-Nyquist receiver filter are uncorrelated (see (2.9b)).

The simulation results are presented in the next three sections. The first section gives results for estimators operating with BPSK signals in the real form of the narrowband, AWGN channel described in Section 2.2. The second section gives performance results for estimators operating with 8-PSK signals in the complex form of the narrowband, AWGN channel. The third section presents additional results of some interest: the performances of a couple of selected estimators are simulated in the wideband, real AWGN channel described in Section 2.3.1, and one-sample per

symbol ML SNR estimation before and after the MF is compared.

2.6.1 Performance Results for BPSK in Real AWGN

Performance Results for Estimators using Sixteen Samples per Symbol

The estimators that use sixteen samples per symbol to generate SNR estimates in the narrowband channel of Section 2.2 are the $SSME_1$, $SSME_2$, ML TxDA, and ML RxDA SNR estimators. Plots of the normalized bias, normalized variance, and normalized MSE are presented in Figures 2.7 to 2.9 for BPSK signals in real AWGN. Note that curves for the $SSME_1$ and $SSME_2$ are absent from Figures 2.7(b), 2.8(b), and 2.9(b). These curves are omitted because the SNR estimates generated by $SSME_1$ and $SSME_2$ using a block length of 64 symbols are poor—the estimates are not always real, positive numbers since the systems of equations used to generate the SNR estimates are ill-conditioned with so few samples.

It is immediately apparent from Figures 2.8 and 2.9 that the normalized variance and MSE curves of the ML TxDA SNR estimator are almost indistinguishable from the corresponding pre-MF CRB's. The variance of the ML RxDA SNR estimator is almost as good as its TxDA counterpart, but its MSE diverges from the MSE of the ML TxDA estimator at low SNR due to the bias caused by receiver decision errors. The low-SNR gap between the MSE of the ML TxDA estimator and that of the ML RxDA estimator narrows as the block length, N_{sym}, is shortened. This observation may be explained by observing in Figure 2.7 that the bias of the ML RxDA estimator at low SNR is relatively unchanged as the block length is decreased from 1024 to 64 symbols, whereas Figure 2.8 reveals that the variances of the ML RxDA and ML TxDA estimators (and the CRB) increase by an order of magnitude. Recall that the MSE reflects both the bias and variance of an estimator. As the block length decreases, the effect of the inherent variance of the estimator on the MSE becomes comparable to the effect of the error-induced bias.

Judging by the plots of the variance of the ML TxDA estimator in Figure 2.8,

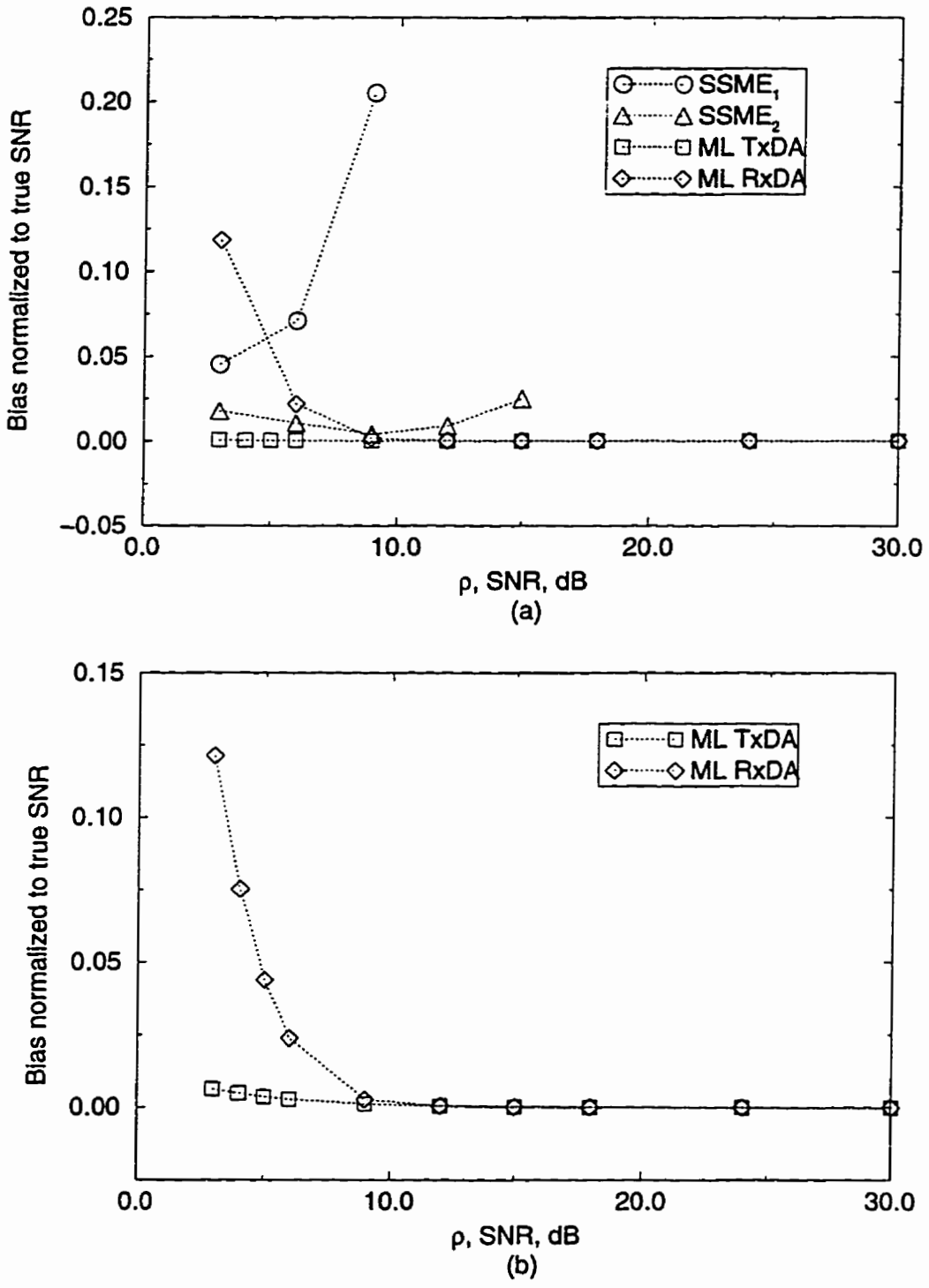

Figure 2.7: Normalized bias with BPSK signals in real AWGN with $N_{ss} = 16$ and (a) $N_{sym} = 1024$, (b) $N_{sym} = 64$.

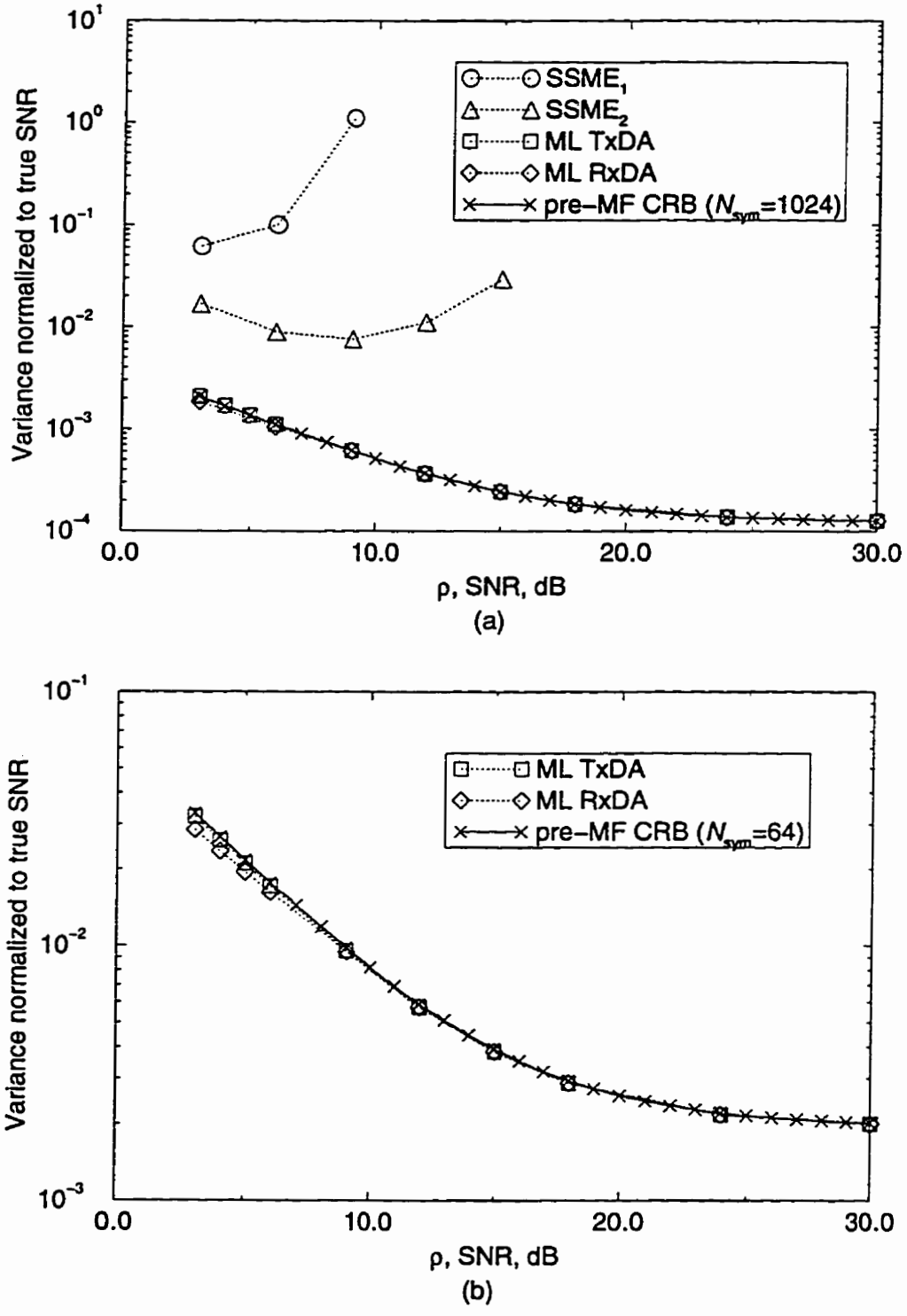

Figure 2.8: Normalized variance with BPSK signals in real AWGN with $N_{ss} = 16$ and (a) $N_{sym} = 1024$, (b) $N_{sym} = 64$.

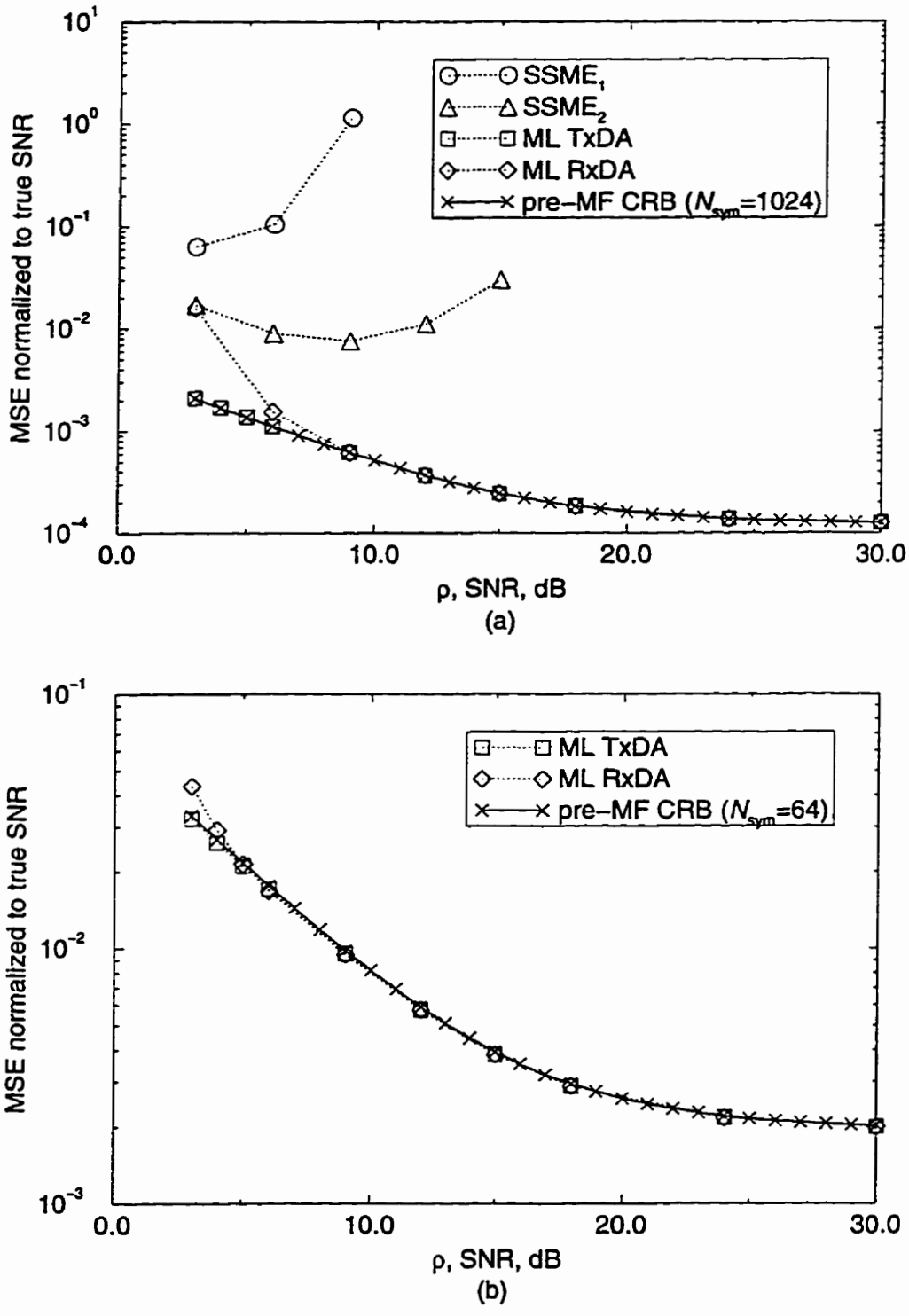

Figure 2.9: Normalized MSE with BPSK signals in real AWGN with $N_{ss} = 16$ and (a) $N_{sym} = 1024$, (b) $N_{sym} = 64$.

it appears, at first glance, that this estimator is efficient; however, it was noted in Section 2.3.2 that the ML TxDA SNR estimator is not efficient since the CRB for unbiased estimators (2.124) does not equal the variance of the reduced-bias ML estimator (2.60). It was also noted that the difference between the variance of the ML TxDA SNR estimator and the corresponding CRB is so small that it would seem of little practical value to search for a more efficient estimator.

The performances of the $SSME_1$ and $SSME_2$ algorithms are generally not as good as those of the ML TxDA and ML RxDA estimators. The performance of the $SSME_2$ is slightly better than that of the $SSME_1$, as is evident from Figures 2.7, 2.8, and 2.9. The bias of the $SSME_2$ at low SNR (less than about 9 dB) is better than that of the ML RxDA estimator, but its variance and MSE are worse over the entire tested range of SNR. The $SSME_1$ and $SSME_2$ are examples of estimators based on higher-order moments that tend to have poor convergence properties in bandlimited channels (other examples are the autocorrelation-based methods discussed in Sections 2.4.1 and 2.4.2). The $SSME_2$ was also adapted to operate on the samples at the input to the MF rather than the output, but similar poor performance was observed and those results are not plotted.

Performance Results for Estimators using One Sample per Symbol

The estimators that use one sample per symbol to generate SNR estimates are the SVR, M_2M_4, SNV TxDA, and SNV RxDA SNR estimators. Plots of the normalized bias, normalized variance, and normalized MSE are presented in Figures 2.10 to 2.12. Results are shown for block lengths of 1024 and 64 symbols for each of the three performance metrics.

The performances of the SNV estimators parallel those of the ML estimators. The SNV TxDA estimator is asymptotically efficient (the number of samples is small enough in Figure 2.11(b) to see that the SNV TxDA estimator is not perfectly efficient). The variance of the SNV RxDA estimator is almost identical to that of

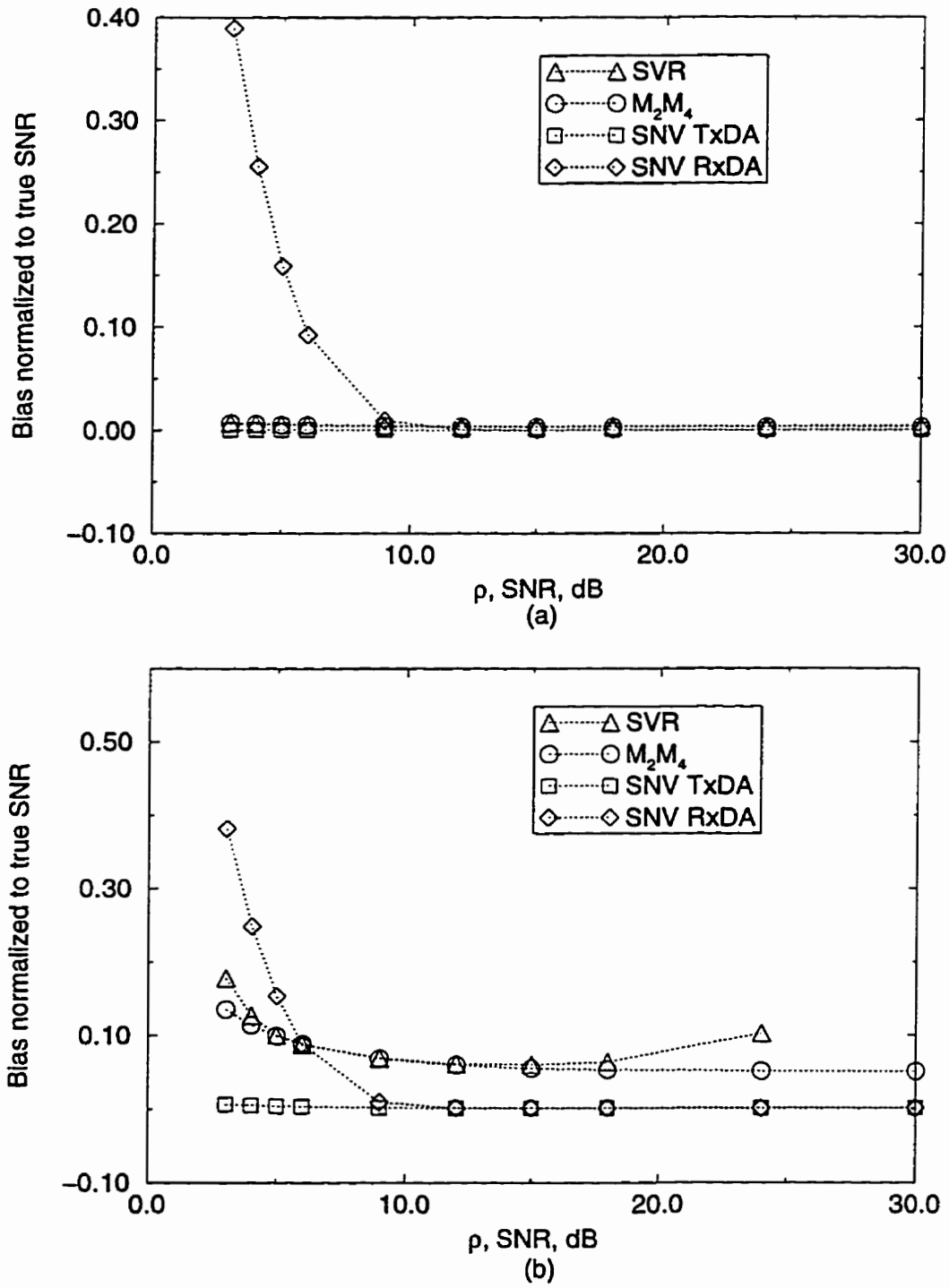

Figure 2.10: Normalized bias with BPSK signals in real AWGN with $N_{ss} = 1$ and (a) $N_{sym} = 1024$, (b) $N_{sym} = 64$.

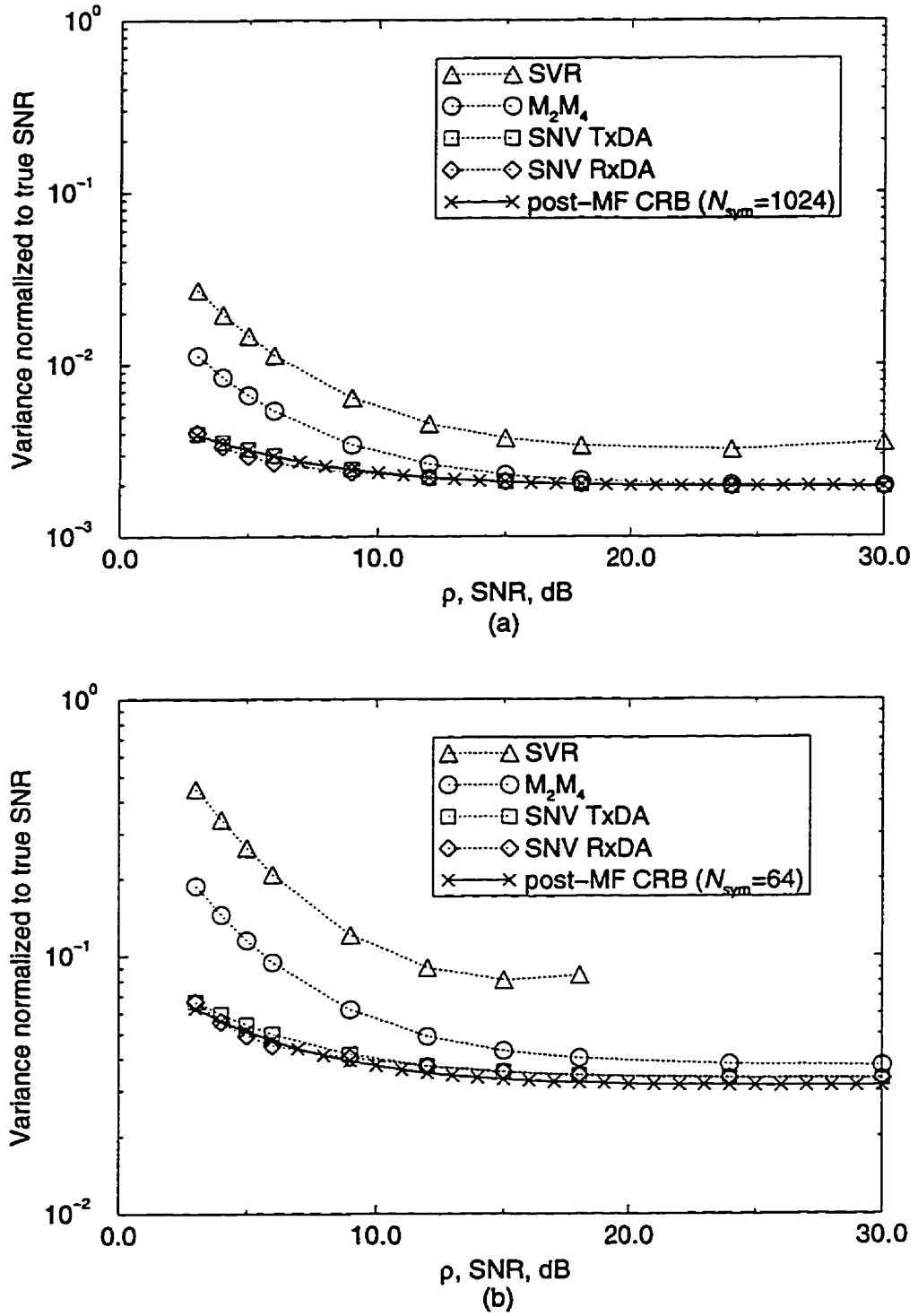

Figure 2.11: Normalized variance with BPSK signals in real AWGN with $N_{ss} = 1$ and (a) $N_{sym} = 1024$, (b) $N_{sym} = 64$.

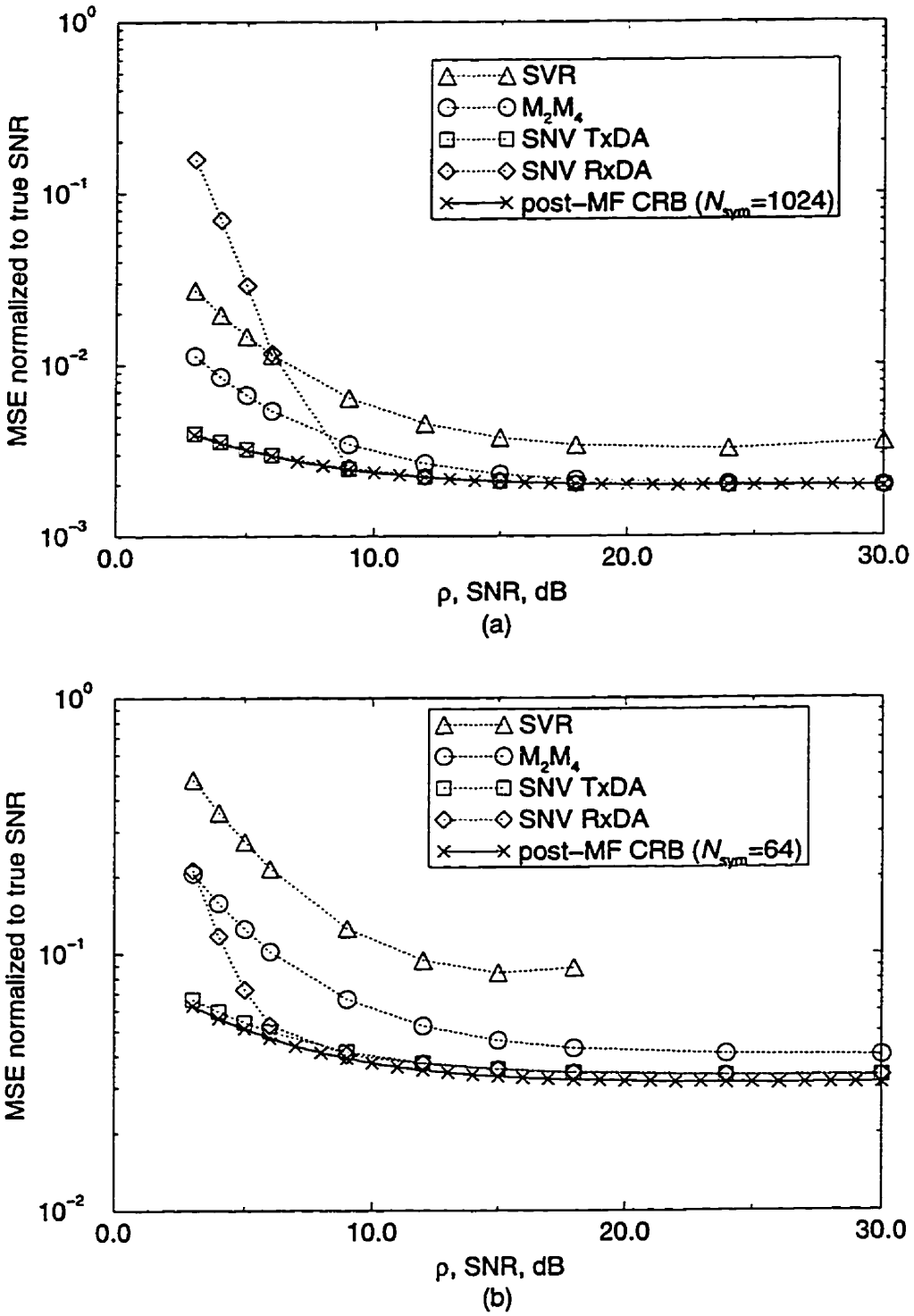

Figure 2.12: Normalized MSE with BPSK signals in real AWGN with $N_{ss} = 1$ and (a) $N_{sym} = 1024$, (b) $N_{sym} = 64$.

the SNV TxDA over the tested range of SNR. The MSE of the SNV RxDA estimator is identical to that of the SNV TxDA estimator at high SNR, but begins to diverge as the SNR is decreased to the point where receiver errors induce a bias in the SNV RxDA estimates. The gap between the MSE of the SNV TxDA estimator and that of the SNV RxDA estimator decreases with decreasing block length as the effect of the estimator variance (which grows as the block length is shortened) on the MSE becomes comparable to the effect of the error-induced bias (which is relatively insensitive to the block length).

The M_2M_4 estimator is asymptotically efficient at high SNR, as is evident by comparing Figure 2.11(a) with Figure 2.11(b). At low SNR ($\rho \leq 8$ dB) and for long block lengths ($N_{\text{sym}} \sim 1000$ symbols), the MSE of the M_2M_4 estimator is smaller than that of the SNV RxDA estimator, but the advantage becomes less significant with decreasing block length so that, at $N_{\text{sym}} = 64$, the MSE of the M_2M_4 estimator is greater than that of the SNV RxDA estimator over the entire range of tested SNR. Though it has been observed by simulation that some SNR estimators based on higher-order moments (such as the SSME algorithms) suffer degraded performance at high SNR, the M_2M_4 estimator is an example of an exception since its bias, variance, and MSE are "well-behaved"; that is, they asymptotically approach some constant value at high SNR.

The SVR estimator generally does not perform as well as the M_2M_4 and SNV estimators, though in Figures 2.10 and 2.12(a) it can be seen that, under some conditions, the bias and MSE curves of the SVR estimator are better than those of the SNV RxDA estimator at very low SNR. Along with the SSME algorithms, the SVR estimator is another example of an estimator whose bias, variance and MSE begin to rise at larger values of SNR. The mechanism behind this degradation is not fully understood, but the effect is more pronounced as the block length is shortened. In fact, for $N_{\text{sym}} = 64$ and for $\rho > 20$ dB, the degradation is so severe that some of the generated estimates are meaningless; that is, the expression given by (2.108)

generates estimates that are negative. For this reason, some high-SNR points have been omitted from the plots of the bias, variance, and MSE of the SVR estimator in Figures 2.10 to 2.12.

2.6.2 Performance Results for 8-PSK in Complex AWGN

In Section 2.6.1, the performances of the real SNR estimators of Section 2.3 operating in a narrowband, real AWGN channel were presented and the effect of different block lengths was illustrated. In this section, the block length is fixed at $N_{\text{sym}} = 64$ and the performances of the complex estimators of Section 2.3 are investigated with 8-PSK signals corrupted by complex AWGN. The performances of the complex forms of the ML TxDA, ML RxDA, SNV TxDA, SNV RxDA, SVR, and M_2M_4 estimators are illustrated in Figures 2.13 to 2.15. The ML estimator results are plotted separately from the results of the other estimators since the ML estimators are the only algorithms investigated in this 8-PSK environment that use multiple samples per symbol. The performances of the SSME algorithms in the 8-PSK case were not studied since these algorithms were derived specifically for BPSK in real AWGN and are not easily adapted to M-ary PSK modulation in the complex AWGN channel.

Comparing Figure 2.8(b) with Figure 2.14(a), and comparing Figure 2.11(b) with Figure 2.14(b), it can be seen that the pre-MF and post-MF CRB's are a factor of two smaller in the 8-PSK case relative to the respective CRB's in the BPSK case. Since the ML and SNV estimators are nearly efficient, the variances of these estimators are also a factor of two smaller in the complex channel. However, comparing Figure 2.9(b) with Figure 2.15(a), and comparing Figure 2.12(b) with Figure 2.15(b), it can be seen that the low-SNR performances of the ML RxDA and SNV RxDA estimators with 8-PSK signals are inferior to their respective low-SNR MSE performances with BPSK signals. The reason for these differences in performances is that receiver errors are more likely with 8-PSK signals given that the signal power and noise power in the 8-PSK and BPSK cases are the same.

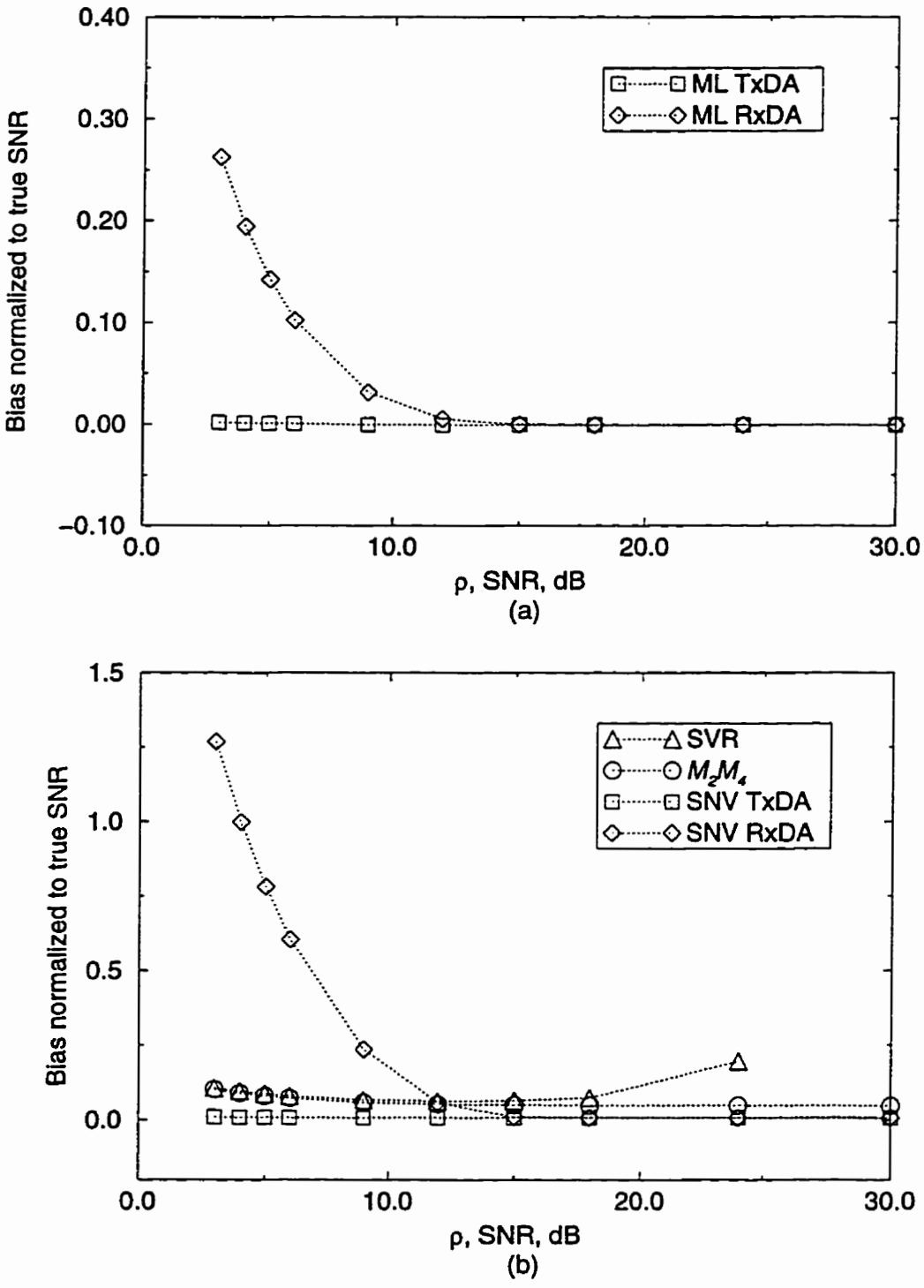

Figure 2.13: Normalized bias with 8-PSK signals in complex AWGN with $N_{\text{sym}} = 64$ and (a) $N_{\text{ss}} = 16$, (b) $N_{\text{ss}} = 1$.

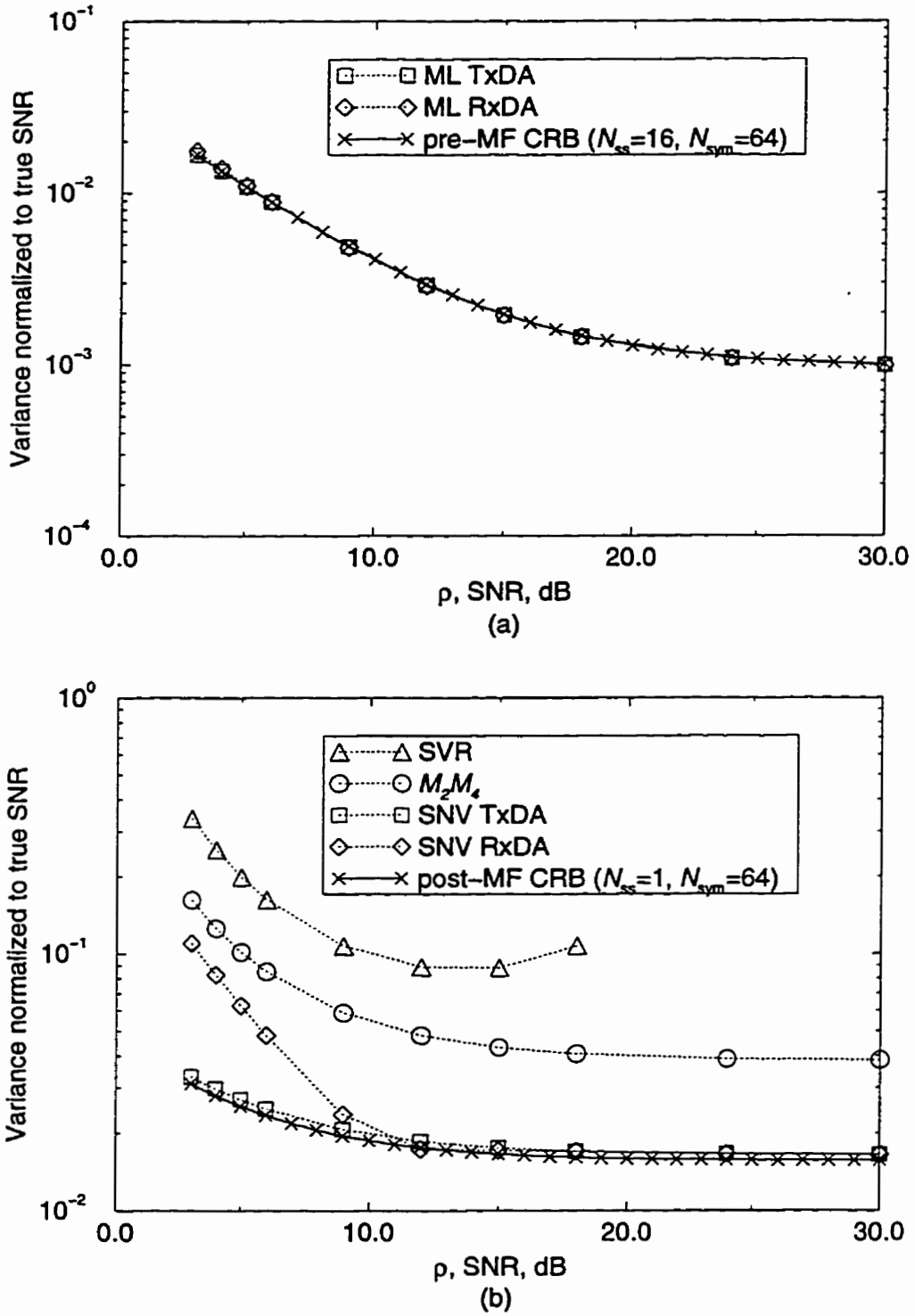

Figure 2.14: Normalized variance with 8-PSK signals in complex AWGN with $N_{\text{sym}} = 64$ and (a) $N_{\text{ss}} = 16$, (b) $N_{\text{ss}} = 1$.

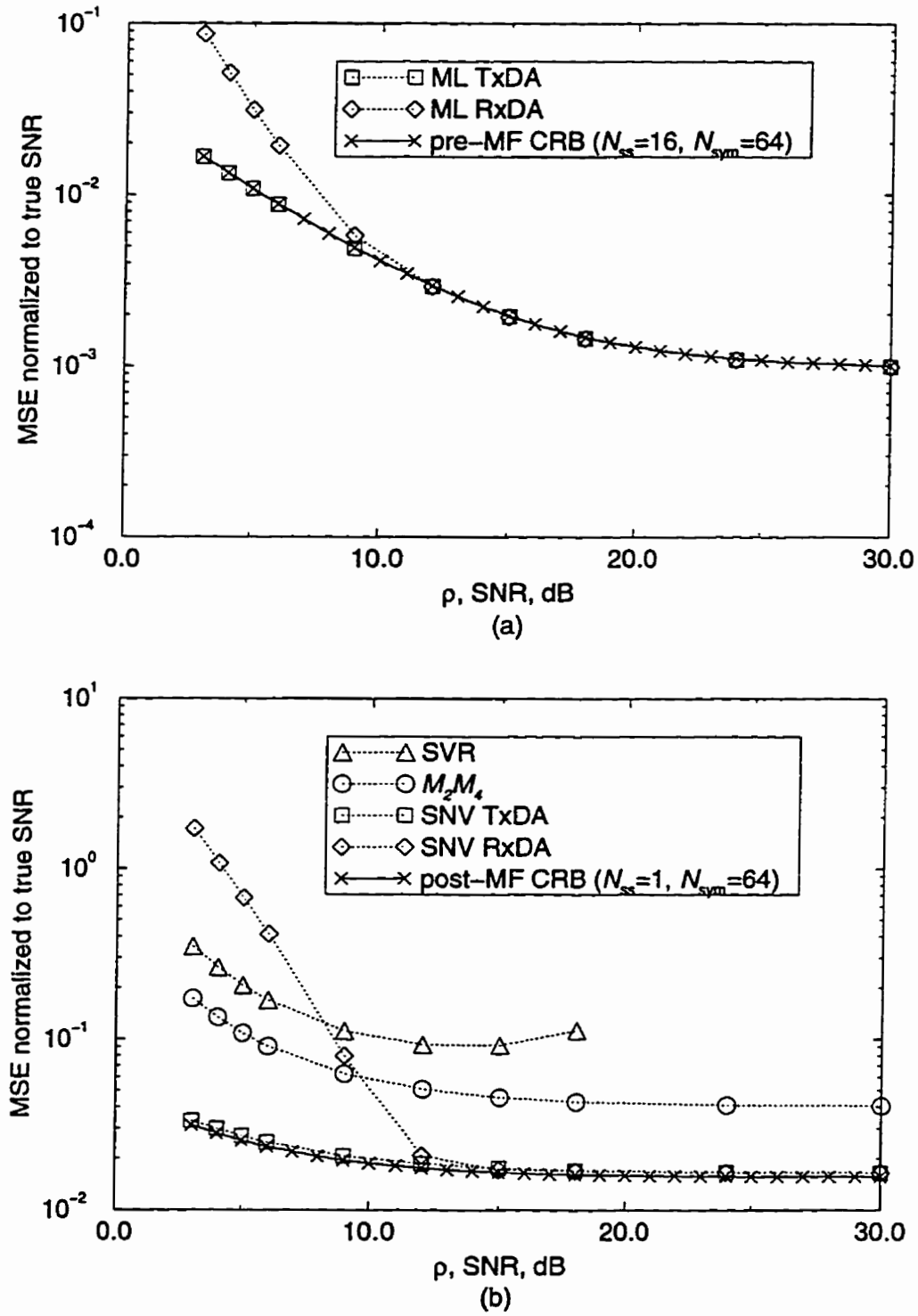

Figure 2.15: Normalized MSE with 8-PSK signals in complex AWGN with $N_{\text{sym}} = 64$ and (a) $N_{\text{ss}} = 16$, (b) $N_{\text{ss}} = 1$.

In contrast to the post-MF CRB and the variance of the SNV TxDA estimator, which are both a factor of two smaller in complex channels as compared to real channels, the variance of the M_2M_4 estimator is close to identical in real and complex channels, as is evident from a comparison of Figures 2.11(b) and 2.14(b). However, the fact that the M_2M_4 performance is almost unchanged in the BPSK and 8-PSK cases coupled with the fact that the performance of the SNV RxDA estimator is relatively worse at low SNR with 8-PSK as compared to BPSK means that the relative performance advantage of the M_2M_4 estimator over the SNV RxDA estimator at low SNR improves as the number of constellation points, M, increases. This behaviour may be observed by comparing Figure 2.12(b) with Figure 2.15(b).

Similarly to the M_2M_4 estimator, the bias, variance, and MSE of the SVR estimator with 8-PSK signals change very little as compared to the bias, variance, and MSE of the SVR estimator with BPSK signals as is seen by a comparison of Figures 2.10(b), 2.11(b), and 2.12(b) with Figures 2.13(b), 2.14(b), and 2.15(b), respectively. As a result, the SVR estimator, too, has a low-SNR performance that improves relative to the low-SNR performance of the SNV RxDA estimator as the order of the modulation increases. However, as in the BPSK case, the degradation in performance observed at high SNR is still present with 8-PSK signals. At high SNR, some of the SVR estimates are negative which is why some high-SNR points have been omitted from Figures 2.13(b), 2.14(b), and 2.15(b).

2.6.3 Additional Results

Some additional results of interest are reported here. In particular, the performance of the SSME_0 algorithm, designed specifically for operation with BPSK signals in wideband, real channels, is investigated and compared to the performance of the ML RxDA estimator. Also, a comparison is presented of one-sample per symbol ML SNR estimation before and after the MF.

Performance in the Wideband Channel

As discussed in Section 2.3.1, the $SSME_0$ SNR estimation algorithm is designed for BPSK signals in wideband, real AWGN channels. A simulation model based on Figure 2.2 was developed in which the $SSME_0$ was tested. For comparison, the ML RxDA estimator was also tested in this wideband channel after making the necessary adaptations to its algorithm. Simulations were run with block lengths of $N_{sym} = 64$ and $N_{sym} = 1024$, and the results are shown in Figures 2.16 and 2.17. Figures 2.16(a) and 2.16(b) show the bias plots for $N_{sym} = 64$ and $N_{sym} = 1024$, respectively (the plots are separated in order to avoid clutter). Figure 2.17(a) shows the variance for both $N_{sym} = 64$ and $N_{sym} = 1024$, and Figure 2.17(b) shows the MSE. It is interesting to note in Figure 2.17 that the variance and MSE of the $SSME_0$ using $N_{sym} = 1024$ approach the CRB for $N_{sym} = 64$ at high SNR. This is a curious result.

In order to investigate this result further, another set of simulations was run in the wideband channel with four samples per symbol rather than sixteen. Figure 2.18(a) shows the MSE curves for this case (the corresponding bias and variance plots are omitted since they do not reveal any additional information). Comparison of Figures 2.17(b) and 2.18(a) reveals that the MSE curves of the $SSME_0$ appear to be independent of the number of samples per symbol used to generate an SNR estimate. Further, Figure 2.18(b) shows that the $SSME_0$ MSE curves appear to be asymptotically efficient at high SNR with the CRB's for *one* sample per symbol; that is, the $SSME_0$ algorithm does not appear to take advantage of the available multiple samples per symbol. In contrast, the MSE performance of the ML RxDA estimator improves as the number of samples per symbol is increased since it takes advantage of the multiple samples per symbol by processing the samples in such a way as to effect a digital matched filter [72].

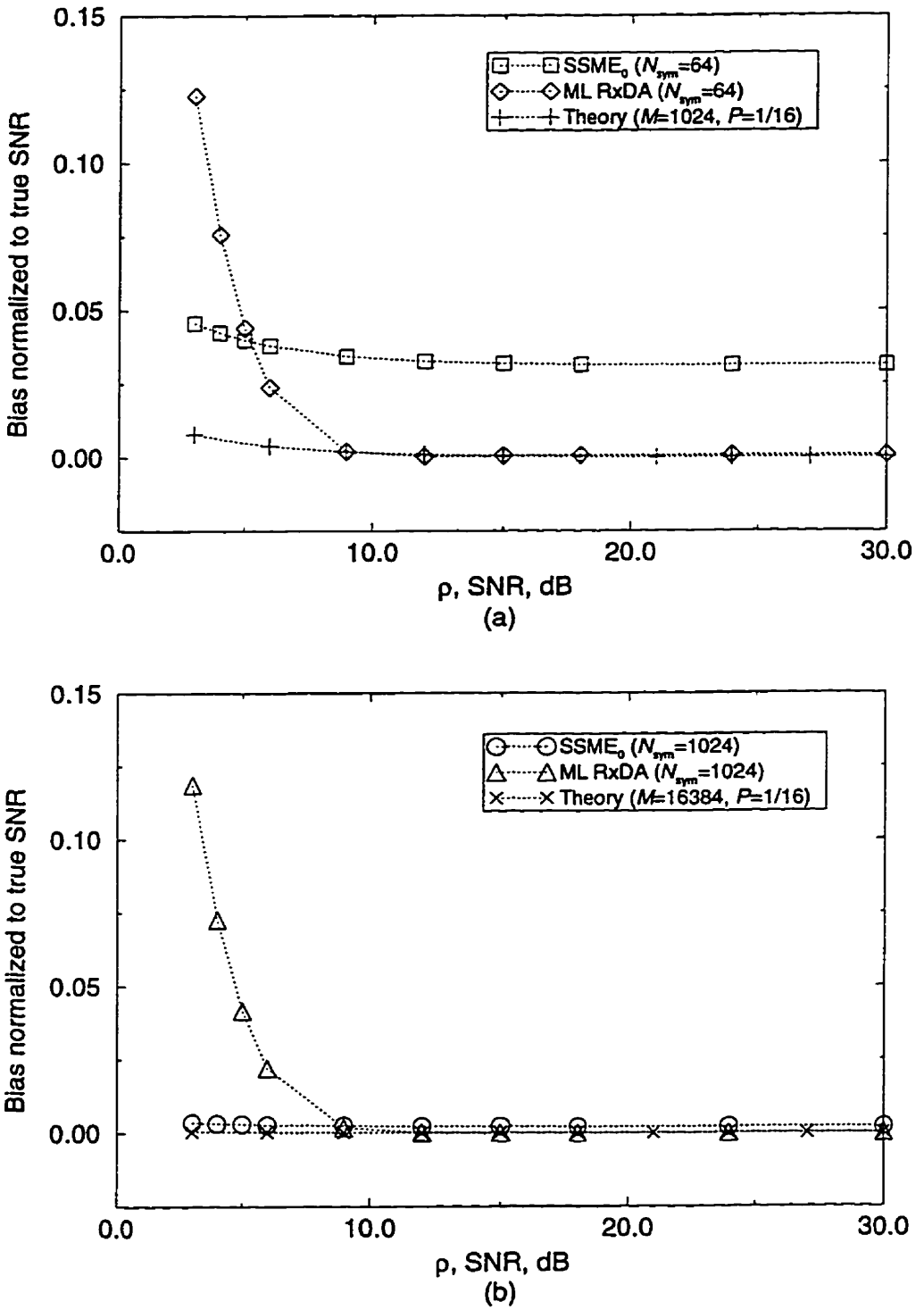

Figure 2.16: Normalized bias in wideband channel with $N_{ss} = 16$ and (a) $N_{sym} = 64$, (b) $N_{sym} = 1024$.

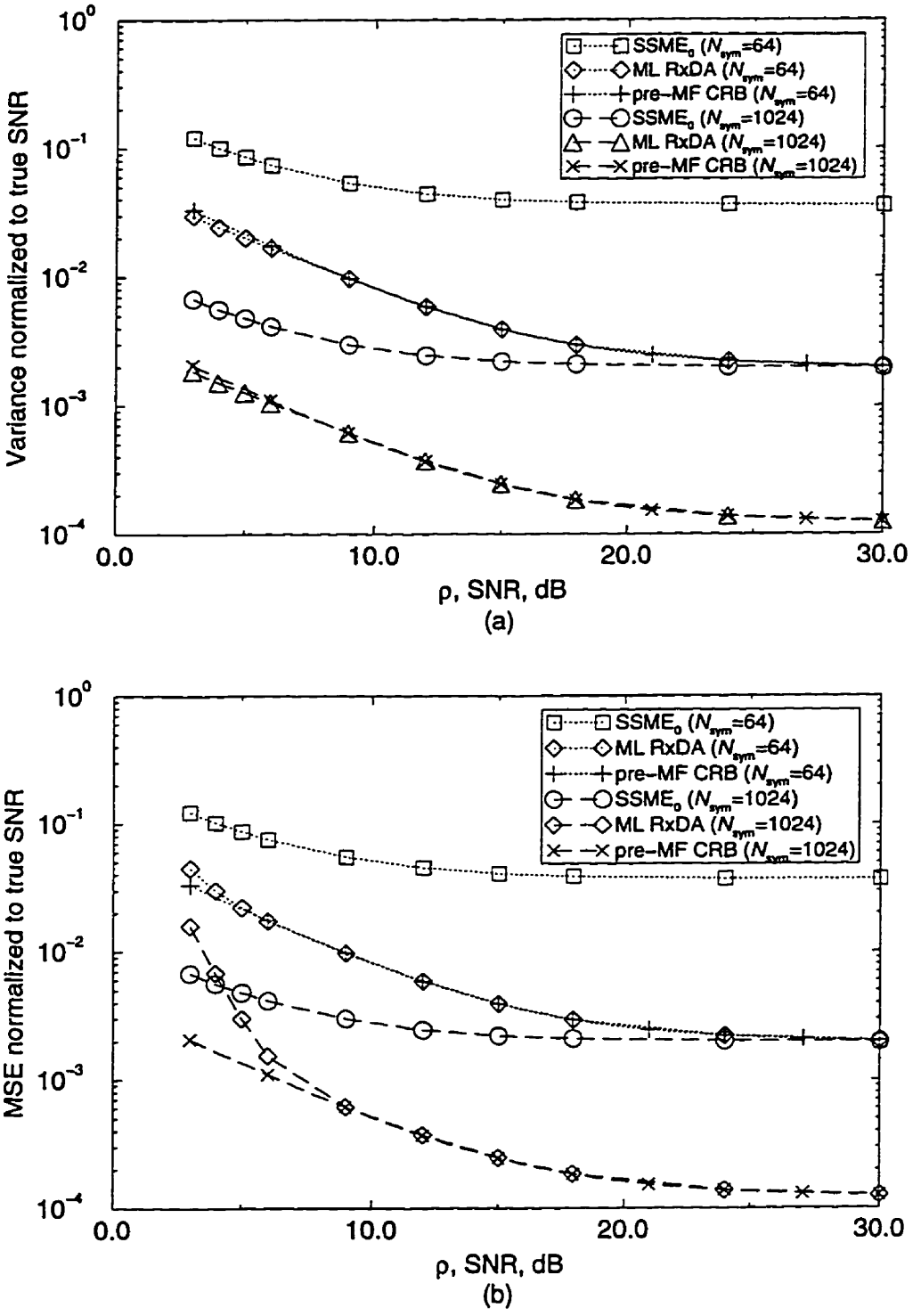

Figure 2.17: (a) Normalized variance in wideband channel with $N_{ss} = 16$, (b) Normalized MSE in wideband channel with $N_{ss} = 16$.

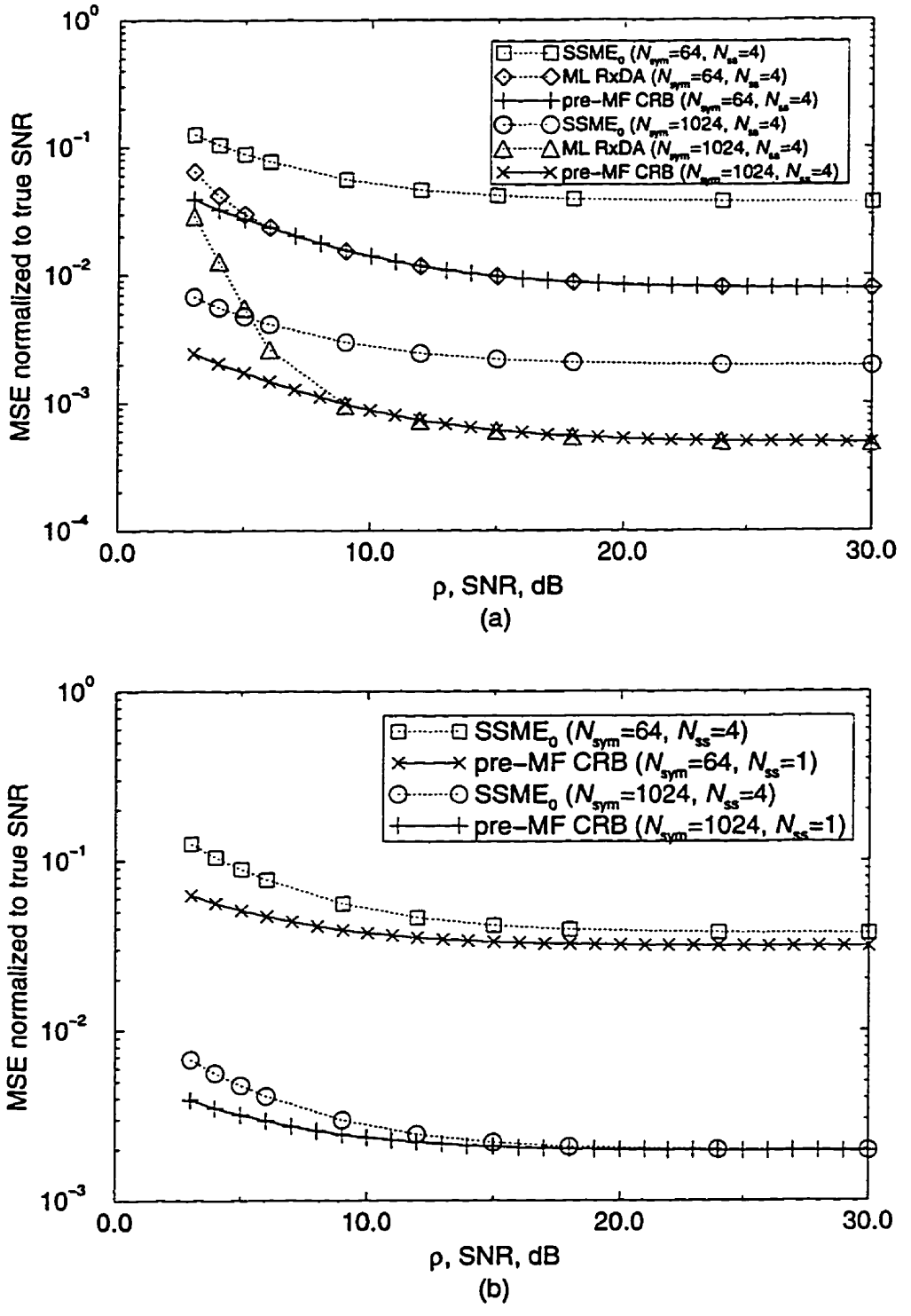

Figure 2.18: Normalized MSE of SSME_0 in wideband channel with $N_{ss} = 4$ compared with (a) MSE of ML RxDA and CRB ($N_{ss} = 4$), (b) CRB ($N_{ss} = 1$).

Comparison of Pre-MF Estimation to Post-MF Estimation in the Narrowband Channel

In Section 2.3.3, it was explained that the SNV estimator is essentially the ML estimator applied to the output samples of the MF taken at the optimal, baud-spaced sampling instants. It is interesting to consider the performance of a similar estimator using one sample per symbol *before* the MF. For simplicity, the system in which this comparison is conducted assumes BPSK signals in real AWGN.

The sequence of pre-MF samples is formed by sampling the received signal, given by (2.4) as

$$r_k = \sqrt{S}\, m_k + \sqrt{N}\, z_k,$$

at a sampling point corresponding to the symbol centres. Since the signal portion, m_k, of the received, pre-MF signal is corrupted by ISI (by design, due to the fact that m_k is the output of a *root*-raised cosine filter), the baud-spaced, pre-MF samples will not have constant magnitude; however, the average magnitude of the symbol centres will be equal to $\sqrt{S}h_0$ where h_l represents the taps of the transmit filter and h_0 is the peak of the impulse response. For the system model of Section 2.2,

$$h_0 \approx 0.0795781. \tag{2.132}$$

The performances of the one-sample per symbol pre-MF and post-MF ML estimators are compared here analytically by plotting the post-MF CRB and a modified, one-sample per symbol pre-MF CRB. The derivation of the CRB in Section 2.5.2 for real signals in real AWGN assumes the power of the sequence of information symbols is given by (2.55) as

$$\frac{1}{K}\sum_{k=0}^{K-1}(m_k^{(i)})^2 = R_m(0) = \frac{\sigma_a^2 g_0}{N_{ss}}.$$

In Section 2.2, σ_a^2 and g_0 are set to unity. For illustration purposes, maintain the assumption here that σ_a^2 is set to unity, but leave g_0 as a variable parameter. The

post-MF CRB may then be written using (2.124) with $N_{ss} = 1$ as

$$\mathrm{CRB}_{\mathrm{post-MF}} = 2\left(\frac{2}{\rho g_0 N_{\mathrm{sym}}} + \frac{1}{N_{\mathrm{sym}}}\right). \tag{2.133}$$

The pre-MF CRB for real signals in real AWGN cannot be obtained using (2.124) since the expression for the power of the sequence of information symbols given by (2.55) is not applicable. Instead, the power of the N_{sym}, baud-spaced pre-MF samples is given by

$$\frac{1}{N_{\mathrm{sym}}} \sum_{n=0}^{N_{\mathrm{sym}}-1} (m_k^{(i)}|_{k=nN_{ss}})^2 = h_0^2. \tag{2.134}$$

Using (2.134), the one-sample per symbol pre-MF CRB may be expressed as

$$\mathrm{CRB}_{\mathrm{pre-MF}}(N_{ss} = 1) = 2\left(\frac{2}{\rho h_0^2 N_{\mathrm{sym}}} + \frac{1}{N_{\mathrm{sym}}}\right). \tag{2.135}$$

Note that post-MF CRB given by (2.133) and the one-sample per symbol pre-MF CRB given by (2.135) have the same high-SNR asymptote, and differ only at low values of SNR. Figure 2.19 shows plots comparing these two expressions (substituting $g_0 = 1$, and substituting (2.132) for h_0) using block lengths of $N_{\mathrm{sym}} = 64$ and $N_{\mathrm{sym}} = 1024$. Figure 2.19 clearly shows that the performance of a post-MF estimator becomes increasingly superior to that of a one-sample per symbol pre-MF estimator with decreasing SNR. At high SNR, the CRB curves approach a common asymptote. The performance of the post-MF ML SNR estimator is better than that of the one-sample per symbol pre-MF ML SNR estimator because the effect of the MF is to maximize the SNR at the optimal sampling instant; in other words, the effect of the MF is to make $g_0 > h_0^2$.

These observations confirm that, if only one sample per symbol is available for SNR estimation in a particular implementation (for example, due to hardware processing speed limitations), then the estimation should be performed *after* the MF (at the optimal sampling instant) rather than before, as expected.

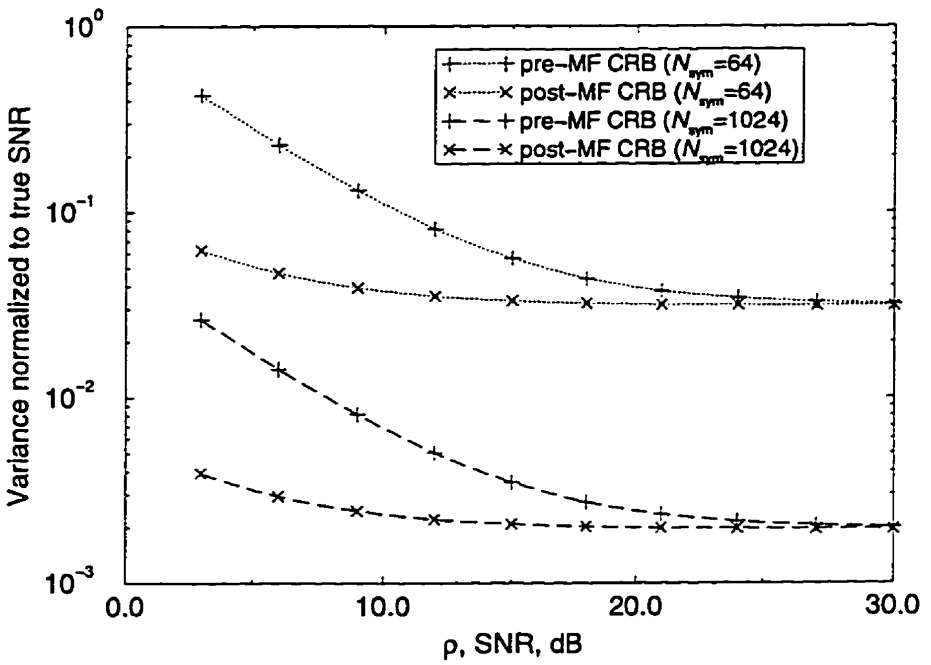

Figure 2.19: Comparison of one-sample per symbol pre-MF CRB to post-MF CRB.

2.7 Implementation Issues

An estimator is not chosen for a particular application based on its performance alone. If an estimator performs very well, but it is extremely difficult to implement, then it may not be the best choice. Factors that affect the ease of implementation are:

1. The computational power required to carry out the operations of a particular algorithm.

2. The peripheral circuitry required to present the received signal to the estimator in the required form.

The issue of computational power is considered first. The available computational power varies depending on the specific technology used to realize an algorithm. For example, it is much simpler to implement arithmetic multiplication and division in a DSP chip than it is in a programmable gate array device. All of the estimators tested here require the ability to compute the ratio of two numbers which makes a gate array realization difficult, unless there is a way to create a modified implementation that does not require division operations (an example of a gate array implementation of the M_2M_4 estimator is presented in [73]). The M_2M_4 and SVR expressions each have a square-root operation which is difficult to perform in DSP and even more difficult in a gate array implementation. The square-root operation would need to be approximated, perhaps by a series expansion or a lookup table. Probably the most complicated estimator to implement is the $SSME_2$ algorithm which requires the solution of a quadratic system of equations. An iterative technique, which demands much computing power, is often used to arrive at a solution in such a case.

Next consider the issue of "peripheral" circuitry which includes all of the pieces of the receiver between the receiver antenna and the estimator including:

- Low-noise amplifier (LNA).

- Frequency down-converters.

- Analog filters.
- Analog-to-digital converters.
- Matched filter.
- Symbol timing recovery (STR).
- Carrier phase recovery (CR).

For example, *all* of the estimators require symbol timing recovery in order to function properly, but not all estimators require carrier phase recovery. Those estimators not requiring CR are those that operate on even powers of the received signal which include the $SSME_2$, SVR, and M_2M_4 estimators.

Each estimator has its own particular constraints and advantages which need to be considered separately. For example, consider the following unique implementation advantage of the ML SNR estimators. In the description of the ML estimators in Section 2.3.2, it was pointed out that the ML RxDA estimator must choose the sequence, $m_k^{(i)}$, based on receiver decisions which it uses to correlate with the received signal. As pointed out in [45], this process is identical to the process of correlation detection which is equivalent to matched filtering [62]. This fact can be used to simplify the implementation of the ML RxDA estimator since many digital receivers employ a matched filter.

2.8 Summary

The performances of the various SNR estimators described in Section 2.3 were compared in Section 2.6 in narrowband, real and complex AWGN channels; and in a wideband, real AWGN channel. Based on the results presented, the "best" estimator to use depends on the given application.

If known data is available to the receiver, the ML TxDA and SNV TxDA estimators perform so well as to make it difficult to justify attempts to find better estimators

for the channel conditions considered here. These TxDA estimators can only generate estimates when known data is available to the receiver, but the particular application will dictate whether or not this limitation is objectionable.

If an application demands the continuous generation of SNR estimates, then the selection of one of the in-service SNR estimators would be more appropriate. It has been demonstrated in Section 2.6 that the choice of the "best" in-service SNR estimator depends on the block length (N_{sym}), the number of samples per symbol (N_{ss}) available, the type and order (M) of modulation used, and the SNR range of interest. The issue of complexity, discussed briefly in Section 2.7, is another factor to consider.

The simulated performances presented in Sections 2.6.1 and 2.6.2 of the SVR, M_2M_4, and SNV estimators are not specific to the system model of Section 2.2 but are identical among systems that employ any type of root-Nyquist filter in the transmitter and receiver. The simulated performances of the ML estimators are identical even among systems which employ filters other than those of the root-Nyquist type (Nyquist's criterion for ISI-free transmission need not be satisfied) in the transmitter and receiver—the only requirement for the results to remain identical is that the sum of the squares of the filter coefficients be unity. The simulated results of the SSME_0 algorithm are specific to the wideband, real AWGN channel, and the simulated results of the SSME_1 and SSME_2 are specific to the particular system configuration described in Section 2.2.

The SSME_1, SSME_2, and SVR algorithms exhibit biases, variances, and MSE's that rise with increasing SNR. Though the exact mechanism that causes the bias, variance, and MSE to rise at high SNR is not fully understood, it is not unreasonable that this effect should occur. Note that the CRB itself, normalized to the true SNR, does not fall indefinitely with increasing SNR but, instead, approaches a constant (see, for example, Figures 2.11 and 2.12). Since the performances of the *best* estimators cannot be better than the CRB, the normalized variances and normalized MSE's of

those estimators, at best, also approach a constant at high SNR. Consequently, it is reasonable to expect the variance and MSE of an inferior estimator either to approach some larger constant or to rise with increasing SNR.

2.8.1 Comments on Specific Estimators

The SSME Algorithms

From the point of view of overall performance, the $SSME_0$ algorithm is judged to be the best of the three SSME algorithms when used in wideband channels for which it was designed. The $SSME_0$ algorithm is the only one of the three that has stable performance at high SNR. It was pointed out in Section 2.6.3 that, though the $SSME_0$ operates on multiple samples per symbol, its performance is relatively independent of the number of samples per symbol available. In fact, from Figure 2.18(b) it is apparent that, regardless of the number of samples per symbol processed by the estimator, the performance of the $SSME_0$ is actually comparable to the CRB for one sample per symbol.

In the narrowband channel of Section 2.2, the SSME series of estimators have trouble with $N_{sym} = 1024$ and, with $N_{sym} = 64$, the generated SNR estimates are sometimes negative or non-existent. The exact mechanism behind the poor high-SNR performance is not fully understood, but a detailed examination of the simulation data reveals that the systems of equations used to generate estimates of the discrete signal power, S, and discrete noise power, N, are ill-conditioned and become increasingly so as N_{sym} is decreased and as the actual SNR is increased.

Of the two SSME algorithms designed for narrowband channels, $SSME_1$ and $SSME_2$, the $SSME_2$ algorithm performs better, thus showing that the approach taken in Section 2.3.1 for the development of an improved SSME algorithm for narrowband channels does, indeed, yield a better estimator.

The ML Algorithms

The variance and MSE of the ML TxDA estimator are indistinguishable from the CRB for block lengths of 64 and 1024 symbols using sixteen samples per symbol thus making the ML TxDA estimator the best SNR estimator in the category of estimators that use multiple samples per symbol. Likewise, its in-service, RxDA counterpart is also asymptotically efficient for high SNR where receiver errors are negligible, but the low-SNR performance of the ML RxDA estimator degrades relative to that of the ML TxDA estimator with increasing block length and order of modulation.

Note that the variance of the ML RxDA estimator is less than the CRB under certain conditions. For example, Figure 2.8(b) shows such a case at low SNR for $N_{\text{sym}} = 64$. Recall that the CRB plotted is for any *unbiased* estimator. The plotted unbiased CRB does not correspond perfectly to the variance of the ML RxDA estimator since this estimator exhibits a strong bias at low SNR. It is not a theoretical contradiction, therefore, that the variance of this estimator is less than the unbiased CRB at low SNR.

The SNV Algorithm

The SNV TxDA estimator performs best of all of the estimators studied in the category of estimators using only one sample per symbol in the sense that its variance and MSE come closest to the theoretical bound. The variance of the SNV TxDA estimator is indistinguishable from the post-MF CRB for a block length of 1024 symbols (see Figure 2.11(a)), but there is a noticeable gap between the variance of the SNV TxDA estimator and the CRB for a block length of 64 symbols (see Figure 2.11(b)) which demonstrates that the SNV TxDA estimator is not truely efficient, but is asymptotically efficient.

The variance and MSE of the in-service, ML RxDA estimator behave identically to those of its TxDA counterpart at high SNR where receiver errors are negligible, but its performance degrades at low SNR, and the performance gap between the

ML RxDA and ML TxDA estimators widens with increasing block length and order of modulation. The variance of the SNV RxDA estimator dips below the CRB for some values of low SNR but, as mentioned above in the discussion of the ML RxDA estimator, this observation poses no theoretical contradiction since the CRB is for unbiased estimators, and so it cannot be expected to correspond well to the variance of the SNV RxDA estimator at low SNR where the bias of the estimator is worst.

The M_2M_4 Algorithm

This method, based on higher-order moments, is well-behaved at both low and high values of SNR for block lengths of both $N_{\text{sym}} = 64$ and $N_{\text{sym}} = 1024$. The variance of the M_2M_4 estimator approaches the CRB at high SNR with increasing block length, and so is asymptotically efficient, but this estimator is not as efficient as the SNV TxDA estimator, which may be seen by comparison of Figures 2.11(a) and 2.11(b).

The M_2M_4 estimator is the best in-service estimator at low SNR given a long block length ($N_{\text{sym}} \sim 1000$) or a higher-order signal constellation. Receiver errors at low SNR do not deleteriously affect the M_2M_4 estimator as they do the SNV RxDA estimator.

Since this algorithm is based on even moments of the received signal, the M_2M_4 estimator has the advantage that carrier recovery is unnecessary.

The SVR Algorithm

This estimator performed least well of all of the estimators tested here that employ one sample per symbol to generate SNR estimates. With a block length of $N_{\text{sym}} = 64$ symbols, the SVR estimator exhibits similar poor high-SNR behaviour as the SSME_1 and SSME_2 estimators (with $N_{\text{sym}} = 1024$); that is, the bias, variance, and MSE all start to rise. The effect is noticeable, but less pronounced, when the SVR estimator uses a block length of $N_{\text{sym}} = 1024$ symbols.

As in the case of the SSME estimators, the mechanism behind the poor high-SNR performance of the SVR estimator is not fully understood. It is noteworthy, however, that the SVR estimator is based on the autocorrelation properties of a received signal. The estimators discussed in Sections 2.4.1 and 2.4.2 are also based on autocorrelation concepts and were judged to be unsuitable for this study due to the very large number of symbols required in order to obtain finite or non-negative estimates. The SVR estimator also appears to have trouble using a relatively small number of symbols to generate SNR estimates.

An advantage of this estimator is that, since the algorithm requires the computation of the square and the fourth power of the received signal, the phase of the received signal is unimportant, meaning that carrier recovery is not required for the operation of this estimator.

Chapter 3

SIR Estimation in Mobile Radio Channels

3.1 Problem Definition

The goal is to extend the concept of SNR estimation in the AWGN channel to signal-to-impairment ratio[1] (SIR) estimation in wireless channels such as those encountered in mobile radio. Impairments affecting mobile radio channels include AWGN; fast, frequency-selective Rayleigh fading[2]; and cochannel interference (CCI) [24]. The performances of many system functions may be improved through use of the SIR. In Section 1.2, several applications that require a quantitative measure of channel quality were described.

There are various ways to express the SIR in any given environment, depending on the measurement method and the way in which the SIR is to be used. The ultimate measure of channel quality is the raw error rate of the channel—the error rate without equalization or coding [77]. There are a few different definitions of error rate, but the most common, and the simplest, is the average bit error rate (BER) which is simply the ratio of the number of bit decision errors to the total number of transmitted bits over some observation interval. The raw, average BER does not

[1]Note that some references (see [74, 2]) use "SIR" to denote "signal-to-interference ratio" whereas, in this study, the "I" stands for the combined effect of all channel "impairments."

[2]A brief summary of some basic fading concepts and terms used in this study is offered in Appendix E. Tutorial papers on the subject of fading channels include [75, 76].

reflect the burstiness of errors, but it is an adequate descriptor of channel quality nonetheless. For example, the European second-generation digital cellular standard (GSM) uses the raw, average BER to define reception quality [77, 78].

Though the BER is the ultimate measure of channel quality, direct estimation of the BER may not necessarily be the easiest way to generate estimates of channel quality. Instead, some observable property of the received signal is sought that is easy to measure which, after a simple transformation, yields the SIR of the channel. It is proposed here that whatever metric is used to quantify the SIR of a channel, it ideally should have a monotonic, one-to-one correspondence with the BER. The SNR in the AWGN channel is a good SIR metric in this respect since the SNR is the *only* parameter that determines the BER in the AWGN channel. For example, the probability of error (or BER) for coherent BPSK is given by (2.13) of Section 2.2 as $P_b = Q(\sqrt{\rho})$ where ρ is the SNR. In fading channels, the SNR alone is not sufficient to determine the error rate. In fading channels, the BER is found to be dependent on the SNR, the overall multipath delay spread of the channel, and the overall Doppler spread [75, 25]. CCI must also be taken into account [24, 25].

SIR estimates are generated from the observation of a number of received symbols, just as in the case of SNR estimation. The SIR estimators under consideration are mainly of the in-service type (that is, the SIR is derived directly from the information-bearing received signal). The in-service estimators are either RxDA estimators or moment-based estimators. Some TxDA SIR estimators are also considered (see Section 2.1 for an explanation of the terms, *TxDA* and *RxDA*).

First, some sort of definition of SIR must be provided before searching for algorithms to estimate this ambiguous parameter. A few possible definitions of SIR are offered here.

1. The first proposed definition of SIR has the form

$$\mathrm{SIR}_1 = \frac{P_{\text{avg, desired signal}}}{P_{\text{avg, impairments}}}$$

where $P_{\text{avg, desired signal}}$ is the average power of the desired signal in the strongest path from the transmitter to the receiver (usually the shortest path). This quantity is measured over a duration long enough to average out short-term fading effects. The quantity, $P_{\text{avg, impairments}}$ is the total average impairment power due to variance in the average amplitude of the desired signal (caused by fast fading), multipath-induced ISI, AWGN, and CCI. A similar definition is offered by Hladik *et al* in [27], and Chennakeshu and Saulnier in [28]. SIR_1 does not have a good one-to-one correspondence with the BER of the channel since it does not distinguish between the various types of impairments. In general, different types of impairments have different statistics and so do not have the same deleterious effects on signal detection. SIR_1 is the SIR definition most closely related to the concept of SNR discussed in Chapter 2. Note that this definition is an adequate descriptor of SIR in Chapter 2 since the statistics of the impairment process (AWGN) are completely known.

2. The second possible definition of SIR has the form

$$\text{SIR}_2 = \frac{\varphi^2}{1-\varphi^2}$$

where φ is the correlation between the corrupted received signal and a "clean" copy of the transmitted signal. This type of estimator was introduced in Chapter 2 in the context of DA SNR estimation and is also discussed in [63, 64]. The SNV and ML SNR estimators from Chapter 2 are examples of methods that correlate the transmitted message sequence (or an estimate of the transmitted message sequence) with the noisy received signal to generate SNR estimates. SIR_2 suffers from the same drawback as SIR_1—SIR_2 does not have a one-to-one correspondence with the BER of the general mobile radio channel since it does not distinguish between different types of impairments.

3. The third variant of SIR has the form

$$\mathrm{SIR}_3 = f(\rho, \tau_{\mathrm{rms}}, \nu_{\mathrm{max}}, S/C).$$

SIR_3 is an explicit function of the individual impairments on the channel: ρ is the SNR, τ_{rms} is the rms delay spread, ν_{max} is the Doppler spread, and S/C is the ratio of the discrete signal power to discrete noise power at the decision device in the receiver. It is easily appreciated that the separate estimation of each of the impairment parameters is a difficult task; however, SIR_3 *does* have a one-to-one correspondence with the BER of the channel. The one-to-one correspondence only holds, however, if the four impairments considered by SIR_3 are the *only* impairments that exist. The existence of unaccounted impairments with different statistics (impulsive ignition noise, for example) or impairments whose statistics change with time (from Rayleigh to Rician fading, for example) spoils the one-to-one correspondence. In other words, SIR_3 is dependent on the assumed characteristics of the individual impairments.

4. The final proposed form of SIR is

$$\mathrm{SIR}_4 = f(\mathrm{BER}).$$

The ultimate SIR estimator is one that is an explicit, one-to-one function of the BER of the channel. In the most direct form, SIR_4 could be just the BER itself, but this study searches for methods that measure SIR using parameter estimation methods, rather than direct measurement of the BER (see [36, 37]).

Most of the SIR estimators investigated in this study have forms similar to SIR_1 and SIR_2.

As in the case of the search for the "best" SNR estimator, the search for the "best" SIR estimator is carried out by the comparison of the performances of various published algorithms in a common channel, using common performance metrics. A few

novel estimators are also investigated. The performance of each estimator is evaluated by software simulation in various mobile radio channels. The system environment is specified in the next section.

3.2 Environment Assumptions

In order to compare the performances of SIR estimators in mobile radio channels, a few assumptions must be made in order to put this study in a relevant context. It is logical to use the second-generation digital cellular standards[3] as a guide since these standards define the direction of the wireless communications industry [85, 88, 86]. In particular, the North American EIA/TIA IS-54 standard [79] is selected for this investigation of SIR estimation.

A few pertinent characteristics of the IS-54 standard are listed in Table 3.1 compiled using [85, 10, 87] as references. Note that the handoff and diversity functions listed in the table are two applications that can benefit from SIR estimation as discussed in Section 1.2.

Based on the IS-54 characteristics listed above, the channel simulator implemented for this study of SIR estimators models $\pi/4$ DQPSK-modulated signals transmitted at a bit rate of 48.6 kbit/s (baud rate of 24.3 kbaud/s), sent in blocks of 150 symbols[4]. Baseband differential detection is employed at the receiver assuming perfect carrier and symbol synchronization. A RRC filter ($\alpha = 0.35$) shapes the transmitted signal and provides matched filtering at the receiver (Appendix B describes how the RRC filter may be implemented by an FIR filter using the frequency sampling technique). CCI is modeled as a single cochannel user employing the same $\pi/4$ DQPSK

[3]The three main second-generation digital cellular standards are: North American EIA/TIA IS-54 [79], pan-European GSM [80, 81, 82], and Japanese PDC (or JDC) [83, 84]. Comparisons of the three systems are provided in [85, 86, 10, 87]

[4]A block length of 150 symbols is chosen instead of exactly 162 symbols because not all 162 symbols are "good". The IS-54 slot structure (see, for example, Figure 5 of [85]) allocates six leading bits (three symbols) as "guard bits", and the following six bits (another three symbols) as "ramp" bits. To be conservative, it was decided that the SIR estimators in this study would use 150 of the 162 available symbols.

Forward band	869–894 MHz
Reverse band	824–849 MHz
Modulation	$\pi/4$ DQPSK
Filtering	Raised-cosine ($\alpha = 0.35$)
Demodulation	not specified (coherent, differential, or discriminator techniques may be employed)
Access method	TDMA
Gross bit rate	48.6 kbit/s
Symbol duration	41.2 μs
Number of bits per time slot	324
Number of symbols per time slot	162
Time slot duration	6.7 ms
Adaptive equalization	Yes
Handoff method	Mobile Assisted Handoff (MAHO)
Diversity	Antenna diversity may be implemented at the mobiles
Channel coding	Convolutional with interleaving

Table 3.1: Some system characteristics specified by the IS-54 digital cellular standard.

modulation.

The desired signal and the cochannel interferer are subjected to fast, frequency-selective fading. The fading process experienced by the desired signal is assumed independent of the fading experience by the CCI. The fading channel is assumed to be a "wide-sense stationary with uncorrelated scattering" (WSSUS) channel (first introduced by Bello [89]) which may be completely specified by its amplitude distribution, Doppler power spectrum, and delay power spectrum [90].

Cox [91] concluded from measurements of propagation characteristics in New York City that the urban mobile radio channel is characterized well by a *Gaussian* WSSUS (GWSSUS) model over distances of about 30 metres. "Gaussian" implies that the quadrature components of the impulse response of the channel are Gaussian distributed which means the signal amplitude variations follow a Rayleigh distribution. In this special case, all that is required to specify the channel is the Doppler power spectrum and delay power spectrum [92]. The WSSUS assumption may be assumed to hold at least over short distances which is justified in this study since the time

slot duration, as specified in Table 3.1, is only 6.7 ms which corresponds to a spatial displacement of less than 25 cm at 120 km/hr.

There is some debate regarding the selection of an appropriate model for the amplitude variations of a received signal in a mobile radio environment. Some (see, for example, [93, 94]) have advocated distributions such as the Nakagami-m distribution as good models of the signal envelope variations. Others (see, for example, [75]) have objected that there is not enough evidence to support proposed distributions such as the Nakagami-m, and that the signal envelope variations in the mobile radio environment are best characterized by the simple Rayleigh distribution [95]. Since many recent papers related to mobile radio assume a Rayleigh-distributed fading envelope (for example, [96, 11, 97, 98, 99, 100, 2]), the Rayleigh assumption is adopted for this study. Long-term variations of the average value of the signal (log-normal fading) due to shadowing is not of concern in this study due to the WSSUS assumption.

The Doppler power spectrum for a vertical monopole antenna in isotropic scattering is characterized by the so-called "Jakes spectrum" [25, 95] (see Appendix E) given by

$$\Phi_\nu(\nu) = \frac{1}{\pi \nu_{\max} \sqrt{1 - (\nu/\nu_{\max})^2}}, \quad |\nu| \leq \nu_{\max} \tag{3.1}$$

where $\nu_{\max}$ is the maximum Doppler frequency known as the *Doppler spread.* The Doppler spread, $\nu_{\max}$ is related to the speed of travel, V, of the mobile receiver (or transmitter) and the wavelength, λ, of the RF signal by the expression[5]

$$\nu_{\max} = V/\lambda.$$

Some (see, for example, [101]) have suggested that true isotropic scattering is an idealization that is invalid in most cases. Doppler spectra often exhibit a "spiky" appearance which is a manifestation of Doppler shifts due to individual, dominant scatterers. For the sake of simplicity, isotropic scattering is assumed in this study of SIR estimation so that the Jakes spectrum of (3.1) applies.

[5]For example, for a mobile radio operating at 900 MHz in a vehicle moving at 120 km/hr, the Doppler spread is 100 Hz.

The delay power spectrum (or delay profile) of a "typical" mobile radio channel has traditionally been accepted to have an exponential decay [95]. This traditional exponential profile, along with the "double-spike" profile and Gaussian profile, are often used today to characterize the multipath characteristics of a channel [102, 74, 103, 2]. These three profiles may be expressed as

$$\Phi_\tau(\tau) = \begin{cases} \frac{1}{\tau_{\text{rms}}} e^{-\tau/\tau_{\text{rms}}} & \text{one-sided exponential} \\ \frac{1}{2}[\delta(\tau) + \delta(\tau - 2\tau_{\text{rms}})] & \text{double-spike} \\ \frac{1}{\sqrt{2\pi}\tau_{\text{rms}}} e^{-\tau^2/(2\tau_{\text{rms}}^2)} & \text{Gaussian} \end{cases} \tag{3.2}$$

where τ_{rms} is the rms delay spread (see [104] and Appendix E), $\Phi_\tau(\tau)$ represents the normalized delay power spectrum, and it is understood that $\tau \geq 0$ in all three cases. It has been reported by several authors (for example, [2, 103, 105]) that the rms delay spread has a strong influence on the BER performance, and for $\tau_{\text{rms}}/T > 0.1$ or 0.2 (τ_{rms} is normalized with respect to the symbol period, T), the shape of the power delay profile is important. Recalling that the symbol period specified by the IS-54 standard as given in Table 3.1 is 41.2 μs, this rough guideline implies that BER performance in an IS-54-compliant system is sensitive to the shape of the power delay profile for rms delay spreads in excess of 4–8 μs. Propagation tests of some areas have revealed that common rms delay spreads for urban areas range from 1–4 μs, and those of hilly areas range from 5–8 μs, but rms delay spreads of up to 12 μs or more may occur (see [24, 75, 95, 91, 102, 74, 104]). Since these numbers show that the BER in typical, real-world multipath channels can depend on the shape of the delay profile, the fading simulator for this study must have the capability of modeling a continuous delay profile.

Some suggestions for the modeling of the fading environment have been made by the digital cellular standards. The IS-54 standard, for example, specifies a "worst-case" scenario for the multipath environment which is the double-spike model with $\tau_{\text{rms}} = T/2$. That is, the interfering multipath echo is of equal strength to the shorter echo and delayed by exactly one symbol period. This model does not take

into account continuous shapes of the delay profile as the multipath phenomenon is modeled simply as two impulses. Equalizer simulations for the North American digital cellular radio system relied heavily on this model. It was found after field testing that the simulations yielded optimistic BER results as compared to operation in a real mobile radio environment [106] so it was necessary to refine the equalizers using more realistic channel simulation models.

Proposals for test channels were made for the pan-European GSM cellular mobile system as well. The so-called CEPT-COST 207 recommendations specify four different delay profiles for various terrains including: a rural area, a typical urban area, a bad urban area, and hilly terrain [92, 107, 108]. These models are all based on either the exponential profile, or the double-spike model with each "spike" having an exponential decay. Note that the CEPT-COST 207 report also makes recommendations for the Doppler power spectrum. Early echos are modeled by the Jakes spectrum of (3.1), while late echos are specified to have Gaussian shapes (see an example in [92]).

In order to balance the need for realistic channel models with ease of software implementation, the fading channel simulator chosen for this study models the Gaussian WSSUS (Rayleigh amplitude distribution) with a normalized Doppler power spectrum specified by (3.1), and a delay power spectrum specified by the exponential profile, given by (3.2), for values of rms delay spread ranging from 3–12 μs.

Now that the fading channel for this study has been specified, a fading simulator which can model this channel must be identified. The simulator should be flexible so that it may be easily modified to model more complicated channels for possible future study. A few simulators are considered in the next section.

3.3 A Brief Survey of Potential Fading Simulators

Various fading channel simulators have been discussed in the literature. A representative sample of some of these techniques is presented below.

Smith [109] presents an algorithm which generates a fixed length of Rayleigh-

faded samples based on the addition of two filtered WGN sequences in quadrature. This sequence is multiplied with the transmitted signal to produce a flat-faded signal with a given Doppler spread. Frequency-selective fading models may be simulated by incorporating two or more channels of this type, each delayed and scaled with respect to the channel representing the shortest path from transmitter to receiver (see, for example, [99, 100]). This simulator is relatively easy to implement for flat-fading channels (FORTRAN code is provided in [109]), but modeling continuous delay profiles with this method is awkward since a large number of sets of Rayleigh-faded samples have to be generated, each scaled and delayed appropriately with respect to the first-arriving echo. The generation of each set of fading samples requires two calls to an FFT routine. Note that FFT routines typically operate on sets of samples which have a length that is some power of two [71] so that this places restrictions on the number of fading samples generated and could be a cause of inefficiency. For example, if only 550 fading samples are required, this algorithm must generate 1024 samples because that is the smallest number greater than 550 that is a power of two.

Another technique that is easy to implement is the "sum-of-sines" method introduced by Jakes [25] and commonly used to simulate mobile radio channels (see, for example, [110, 111]). A number of sinusoids with distinct Doppler frequencies up to the maximum Doppler frequency are added to form coloured Gaussian noise which is then multiplied with the desired signal to form a faded signal. This is a multiplicative (flat) fading process. A frequency-selective fading simulator may be constructed in a similar manner as described earlier for the simulator presented by Smith; that is, a frequency-selective fading channel may be constructed by the generation of many flat-faded echos, each of which is scaled and delayed with respect to some reference echo.

A fading simulator which more accurately represents continuous or quasi-continuous delay profiles is introduced by Fechtel [112]. This simulator is based on channel orthogonalization techniques in the delay and correlation domain. This simulator is

able to model continuous delay profiles with significantly reduced complexity as compared to the construction of a continuous delay profile with a *large* number of discrete, flat-faded channels, as described earlier.

The operation of some fading simulators is based on measurement-based models which allow very accurate modeling of real fading channels (see, for example [113, 114, 115]). Highly accurate modeling of channel fading conditions comes at the cost of having to construct a database of actual channel measurements.

A method of channel modeling is discussed by D'Aria *et al* in [116] which employs a Fourier series approximation of the time-varying channel transfer function. Enough harmonics are included in the sum to model satisfactorily the time-varying channel which is represented either by actual channel measurements, or by proposed delay profiles (such as the CEPT-COST 207 recommendations [107]).

Dersch and Rüegg [101] describe a fading simulator that models the physical wave interference process and is based on concepts discussed by Braun and Dersch in [94]. The electric field at the receiver antenna is modeled as the superposition of many partial waves having characteristics defined statistically by the Doppler power spectrum, delay power spectrum, and amplitude distribution.

Other statistically-based simulators model the *channel impulse response* as a superposition of many received echos [90, 92, 108, 117]. Here, too, the Doppler power spectrum, delay power spectrum, and amplitude distribution dictate the characteristics of the individual echos that make up the composite received signal. Hoeher [92] describes an approximation to the GWSSUS model where the individual echos are assigned a random arrival phase, delay, and Doppler frequency according to specified probability distributions. The presumed amplitude distribution is Rayleigh and is completely determined by the random arrival phases. Müeller [90] extends the method to arbitrary amplitude distributions and coined this method the "Monte Carlo method." This term is used here to refer to the method described by Hoeher as well.

The model described by Hoeher [92] is chosen as the fading simulator for this study of SIR estimation since it models the GWSSUS channel, the simulator is easy to implement, it allows the modeling of arbitrary quasi-continuous delay power profiles and arbitrary power spectra, and an arbitrary number of fading samples may be generated. This simulator is described in the next section.

3.4 Implementation of the Monte Carlo Method

3.4.1 Monte Carlo Method for Continuous-Time Signals

Let the instantaneous impulse response of the channel be denoted by $f(\tau;t)$ which is the response of the channel at time, t, to an impulse transmitted on the channel at time $t-\tau$. The instantaneous channel impulse response may be written as (see [92])

$$f(\tau;t) = \lim_{M\to\infty} \frac{1}{\sqrt{M}} \sum_{m=1}^{M} e^{j(\theta_m + 2\pi\nu_m t)} \delta(\tau - \tau_m) \tag{3.3}$$

where M is the number of echos, θ_m is the angle of arrival, ν_m is the Doppler frequency, and τ_m is the delay (relative to the echo with the shortest delay) of the m^{th} echo. The parameters, θ_m, ν_m, and τ_m are continuous random numbers each having a specified probability distribution. The larger the number of echos, M, the better the approximation to the GWSSUS channel.

Note that each of the M echos is scaled by the constant factor, $1/\sqrt{M}$. The reason why random amplitudes are not required for this model is that the WSSUS model is *completely* statistically specified by the probability distributions of θ_m, ν_m, and τ_m [92]. The random phase samples, θ_m, are uniformly distributed on $[0, 2\pi)$. In Appendix 1 of [92], it is shown that the Doppler power spectrum and delay power spectrum are, respectively, proportional to the pdf's of ν_m and τ_m so that the pdf's of ν_m and τ_m are, in fact, the normalized Doppler power spectrum and the normalized delay power spectrum, respectively. The Monte Carlo method can model *any* arbitrary discrete, continuous, or piece-wise continuous Doppler or delay profile by the appropriate specification of the pdf's for ν_m and τ_m.

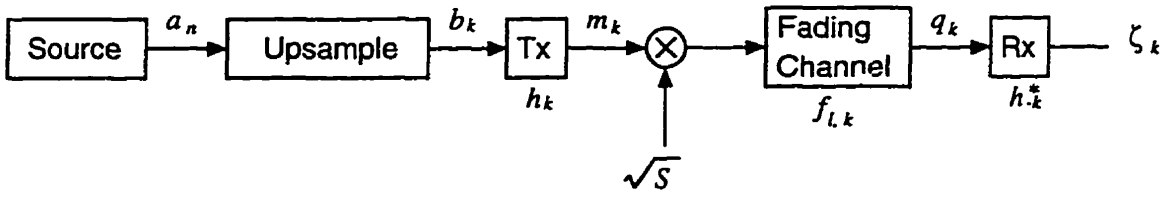

Figure 3.1: Block diagram of system illustrating generation of faded signal only.

The normalized Doppler and delay spectra chosen for this study were identified in Section 3.2 so that the pdf of ν_m may be written directly as

$$p_\nu(\nu) = \Phi_\nu(\nu) = \frac{1}{\pi\nu_{\max}\sqrt{1-(\nu/\nu_{\max})^2}}, \quad |\nu| \leq \nu_{\max} \tag{3.4}$$

and, likewise, the pdf of τ_m may be written directly as

$$p_\tau(\tau) = \Phi_\tau(\tau) = \frac{1}{\tau_{\text{rms}}} e^{-\tau/\tau_{\text{rms}}}, \quad \tau \geq 0. \tag{3.5}$$

3.4.2 Discrete-Time Representation of the Faded Signal

The block diagram of Figure 3.1 illustrates the generation of the discrete-time faded signal (impairments are incorporated into the model later). The source block generates $N_{\text{sym}} + 1$ complex, $\pi/4$-DQPSK-encoded symbols, a_n, $n \in \{0, 1, \ldots, N_{\text{sym}}\}$, from N_{sym} information symbols according to the mapping rule given in Table F.3 of Appendix F. Each encoded symbol corresponds to one of the eight possible points in the $\pi/4$-DQPSK constellation illustrated by Figure F.5 of Appendix F. The upsampled, complex message sequence is expressed as

$$b_k = \sum_n a_n \delta_{k,nN_{ss}} \tag{3.6}$$

where $N_{ss} = 16$ is the number of samples per symbol, as defined in Section 2.2. The pulse-shaped, information-carrying signal is given by

$$m_k = b_k \otimes h_k \tag{3.7}$$

where h_k represents the coefficients of a root-raised cosine filter with rolloff factor $\alpha = 0.35$, and length $L = 127$ taps. The faded signal, q_k, at the input to the receiver is the (discrete) convolution of the scaled message sequence, m_k, with the discrete-time channel impulse response, $f_{l,k}$, giving

$$q_k = \sqrt{S}\, m_k \otimes f_{l,k}. \tag{3.8}$$

The discrete-time impulse response of the channel, $f_{l,k}$ is obtained from the continuous-time representation by letting $t = kT_S$ and $\tau = lT_S$, where T_S is the sample period, and by setting a finite limit to M so that (3.3) becomes

$$f_{l,k} = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} e^{j(\theta_m + 2\pi\nu_m kT_S)} \delta_{lT_S,\tau_m}. \tag{3.9}$$

Finally, the output samples of the MF in the receiver may be represented as

$$\begin{aligned}
\zeta_k &= q_k \otimes h^*_{-k} \\
&= (\sqrt{S}\, b_k \otimes h_k \otimes f_{l,k}) \otimes h^*_{-k} \\
&= \sqrt{S}\, b_k \otimes (f_{l,k} \otimes g_k) \\
&= \sqrt{S}\, b_k \otimes \psi_{l,k} \\
&= \sqrt{S} \sum_{l=L_-}^{L_+} \psi_{l,k}\, b_{k-l} \\
&= \sqrt{S} \sum_{l=L_-}^{L_+} \psi_{l,k}\, a_{\frac{k-l}{N_{ss}}},\ k \in \{0, 1, \ldots, (N_{\text{sym}} + 1)N_{\text{ss}} - 1\}
\end{aligned} \tag{3.10}$$

where $g_k = h_k \otimes h^*_{-k} = h_k \otimes h_k$ (as in Section 2.2), the impulse response of the channel plus the transmit and receive filters is denoted by

$$\begin{aligned}
\psi_{l,k} &= f_{l,k} \otimes g_k \\
&= \sum_{\lambda=-\infty}^{\infty} f_{\lambda,k}\, g(lT_S - \lambda T_S) \\
&= \sum_{\lambda=-\infty}^{\infty} \left[\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} e^{j(\theta_m + 2\pi\nu_m kT_S)} \delta_{\lambda T_S,\tau_m} \right) g(lT_S - \lambda T_S) \right] \\
&= \frac{1}{\sqrt{M}} \sum_{m=1}^{M} e^{j(\theta_m + 2\pi\nu_m kT_S)} g(lT_S - \tau_m),
\end{aligned} \tag{3.11}$$

and the integers L_- and L_+ in (3.10) are chosen such that $\psi_{l,k} = 0$ (or is negligibly small) outside the range $L_- \leq l \leq L_+$.

3.4.3 Generation of the Fading Channel Parameters

In Section 3.4.1, it is noted that, for the continuous-time case, the random phase samples, θ_m, are uniformly distributed on $[0, 2\pi)$, the Doppler samples, ν_m, are distributed according to (3.4), and the delay samples, τ_m, are distributed according to (3.5). This section discusses how to generate random samples with a given distribution.

The phase samples, θ_m, are easily generated by any good uniform random number generator [71, 118, 119], and the same algorithm employed in Chapter 2 for the generation of WGN samples is used here. The random deviates, ν_m and τ_m, may be generated by the transformation of a uniform random variable using the inverse cumulative distribution function (cdf) of the desired parameter. This transformation is known as the *inverse probability integral transform* [118] and is applied below to the generation of random Doppler and delay samples.

Let u_m represent a continuous, uniformly-distributed random variable taking values from $[0, 1)$, and let w_m be related to u_m by

$$w_m = F_w^{-1}(u_m)$$

where F is the cdf of w_m. The random deviates, ν_m and τ_m, may be generated by the substitution of w_m with either ν_m or τ_m, and by the substitution of F_w with F_ν (the cdf of ν) or F_τ (the cdf of τ). (Note that a transformation would be required for θ_m as well for a model in which the arrival phase is not assumed to be uniformly distributed as it is here.)

To find the transformation for ν_m, $F_\nu(\nu)$ is required. The cdf of ν is evaluated from the pdf of ν (given by (3.4)) as

$$F_\nu(\nu) = \int_{-\nu_{max}}^{\nu} p_\nu(\alpha)\, d\alpha$$

$$= \int_{-\nu_{\max}}^{\nu} \frac{1}{\pi \nu_{\max} \sqrt{1-(\alpha/\nu_{\max})^2}}\, d\alpha$$
$$= \frac{1}{\pi}\left[\sin^{-1}\left(\frac{\nu}{\nu_{\max}}\right) + \pi/2\right]. \tag{3.12}$$

Let $u = F_\nu(\nu)$ and solve for ν to get

$$\nu = \nu_{\max} \cos \pi u,\ u \in [0,1). \tag{3.13}$$

This expression is only valid for $|\nu| < \nu_{\max}$. Note that (3.13) is slightly different from the expression for ν given in [92].

Similarly, to find the transformation for τ_m, $F_\tau(\tau)$ is required. First, the pdf of τ_m given by (3.5) must be truncated since a practical simulator cannot model a pdf that has an infinite tail. Let $\tau_{\max}$ denote the longest (finite) delay—delays greater than $\tau_{\max}$ can safely be assumed to have negligible probability. Let a be a normalization factor such that

$$a \int_0^{\tau_{\max}} e^{-\tau/\tau_{\rm rms}}\, d\tau = 1$$
$$a = \frac{1}{\tau_{\rm rms}(1-e^{-\tau_{\max}/\tau_{\rm rms}})}.$$

The pdf for τ_m over the restricted range $[0, \tau_{\max})$ may then be expressed as

$$p_\tau(\tau) = \frac{e^{-\tau/\tau_{\rm rms}}}{\tau_{\rm rms}(1-e^{-\tau_{\max}/\tau_{\rm rms}})},\ 0 \le \tau \le \tau_{\max}.$$

The cdf may then be evaluated as

$$F_\tau(\tau) = \int_0^\tau p_\tau(\alpha)\, d\alpha$$
$$= \frac{1-e^{-\tau/\tau_{\rm rms}}}{(1-e^{\tau_{\max}/\tau_{\rm rms}})},\ 0 \le \tau \le \tau_{\max}. \tag{3.14}$$

Finally, let $u = F_\tau(\tau)$ and solve for τ to get

$$\tau = -\tau_{\rm rms} \ln(1-u(1-e^{-\tau_{\max}/\tau_{\rm rms}})),\ u \in [0,1). \tag{3.15}$$

3.5 Verification of the Monte Carlo Method

The discrete formulation of the Monte Carlo method is used in the software implementation of a fading simulator (written in C) for this study. The fading parameter transformations of (3.13) and (3.15), and the expression for the discrete, fading channel impulse response given by (3.9) are verified in the following sections.

3.5.1 Verification of the Fading Parameters used in the Monte Carlo Method

This section verifies that the implemented C code correctly generates random variables with the appropriate distributions for θ_m, ν_m, and τ_m. The verification is performed in each case by generating 10000 random samples from which an estimate of the cdf is computed. The verification results are presented in graphical form.

Figure 3.2(a) compares the cdf of the simulated phase samples, θ_m, to the theoretical uniform cdf given by

$$F_\theta(\theta) = \frac{\theta}{2\pi}, \ \theta \in [0, 2\pi).$$

Figure 3.2(b) compares the cdf of the simulated Doppler samples, ν_m, generated by the transform given by (3.13), to the theoretical cdf given by (3.12). Figure 3.2(c) compares the cdf of the simulated delay samples, τ_m, generated by the transform given by (3.15), to the theoretical cdf given by (3.14). The cdf's of the simulated samples match theory well in all cases.

3.5.2 Verification of the Discrete, Monte Carlo-Based Fading Simulator Modeling Flat-Fading Channels

In this section, the Monte Carlo-based fading simulator represented by (3.9) is tested to see how closely it can approximate a fast, flat, Rayleigh fading channel.

As a preliminary check, the flat-fading channel is verified qualitatively by the observation of typical amplitude and phase fluctuations over time. Figure 3.3 shows a

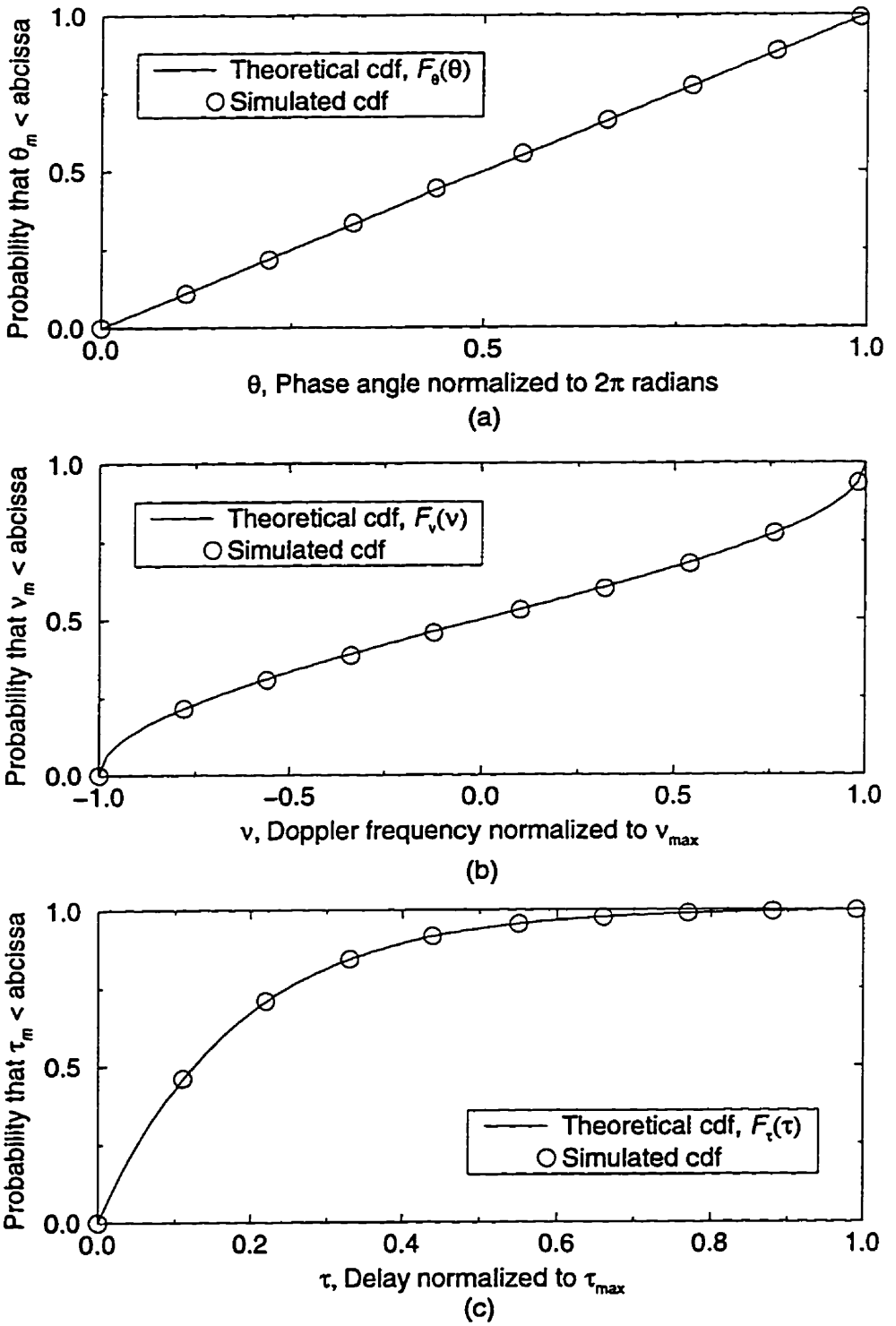

Figure 3.2: Comparison of plots of cdf's generated from simulated samples with theoretical plots of (a) $F_\theta(\theta)$, (b) $F_\nu(\nu)$, and (c) $F_\tau(\tau)$.

125 msec sample of envelope and phase fluctuations for a Rayleigh flat-fading channel having a Doppler spread of $\nu_{\max} = 100$ Hz approximated by a model with $M = 40$ echos. By visual inspection, the envelope fluctuations displayed in Figure 3.3(a) appear consistent with other published plots, such as Figure 3 of [100] and Figure 1(a) of [120]. Figure 1(b) of [120] indicates that rapid changes in phase coincide with deep amplitude fades; the correlation between rapid phase changes and deep amplitude fades observed in Figures 3.3(a) and 3.3(b) is consistent with the reported results.

Next, the flat-fading simulator is tested more rigorously by statistical tests of the complex, faded samples. Four different simulators are tested, each using a different number of echos, M, to model a Rayleigh flat-fading channel. The number of echos corresponding to the four different simulators is $M = 6$, 10, 20, and 40. The reason for considering various channels comprising different numbers of echos is to determine the minimum number of echos required to ensure Rayleigh statistics. Hoeher [92] cites a result published in [25] that the Rayleigh approximation should be good for $M \geq 6$. This claim is tested here. Each of the four flat-fading channels is evaluated by generating a number of faded samples and computing the cdf's of the envelope and phase, the LCR, average duration of fades, and the autocorrelation of the samples. These statistical tests are discussed in [24, 95, 121], and the theoretical expressions (which assume an omnidirectional receiver antenna) given below are from these references.

The LCR is the rate at which the envelope of the faded signal crosses (with positive slope) some constant level, Γ. Letting N_Γ denote the LCR, it may be expressed as

$$N_\Gamma = \sqrt{2\pi}\nu_{\max}\Gamma e^{-\Gamma^2}. \tag{3.16}$$

The average duration of fades, $\bar{\tau}_\Gamma$, is the average period over which the envelope stays below the level, Γ, and is given by

$$\bar{\tau}_\Gamma = \frac{e^{\Gamma^2} - 1}{\sqrt{2\pi}\nu_{\max}\Gamma}. \tag{3.17}$$

The normalized temporal autocorrelation, $\varphi(\tau)$, of the (complex) faded signal is given

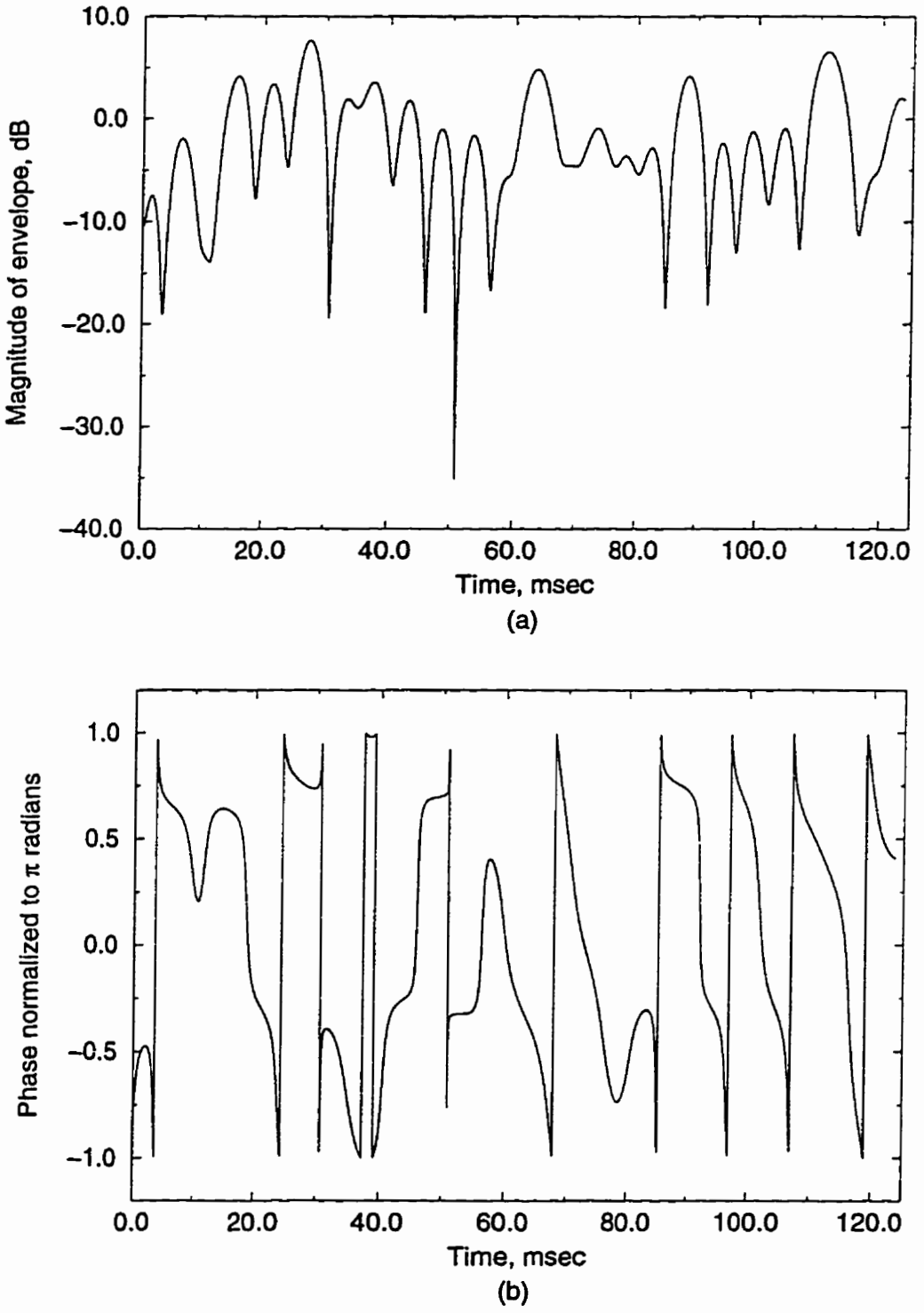

Figure 3.3: Representative samples of 125 msec of (a) envelope fluctuations and (b) phase fluctuations in a flat-fading channel with $\nu_{\max} = 100$ Hz and $M = 40$ echos.

by

$$\varphi(\tau) = J_0(2\pi\nu_{\max}\tau) \tag{3.18}$$

where $J_0(x)$ is the zero-order Bessel function of the first kind.

The results of the five quantitative tests are presented for each of the channels as:

1. Plots of cdf of amplitude variations (should be Rayleigh distributed).
2. Plots of cdf of phase variations (should be uniformly distributed).
3. Plots of LCR (should be consistent with (3.16)).
4. Plots of average duration of fades (should be consistent with (3.17)).
5. Plots of temporal autocorrelation (should be consistent with (3.18)).

The rms level of the envelope fluctuations is normalized to 0 dB, and the maximum Doppler frequency, $\nu_{\max}$, is 100 Hz. The number of faded samples used to generate the cdf's of the envelope and phase samples, the LCR, and the autocorrelation is 10^6, whereas 5×10^7 samples are used for the generation of the average duration of fades. Theory is compared to simulation results in Figures 3.4 to 3.8 corresponding to items 1 to 5, respectively, from the list above.

From Figures 3.4 to 3.7, it can be seen that $M = 6$ echos are sufficient to generate samples that have statistics consistent with a flat, Rayleigh fading channel, based on measurements of the cdf's of the amplitude and phase, LCR, and average duration of fades of the simulated, faded samples. However, the autocorrelation of the faded samples based on $M = 6$ echos is not very consistent with theory as revealed by Figure 3.8(a). In fact, twenty or forty echos appear necessary to generate samples that have the required autocorrelation. Although it appears that the claim made in [25] that the Rayleigh approximation is good for $M \geq 6$ is valid based on tests of the cdf of the envelope and phase, the LCR, and the average duration of fades of the faded samples, the autocorrelation of the faded samples reveals that this claim holds only

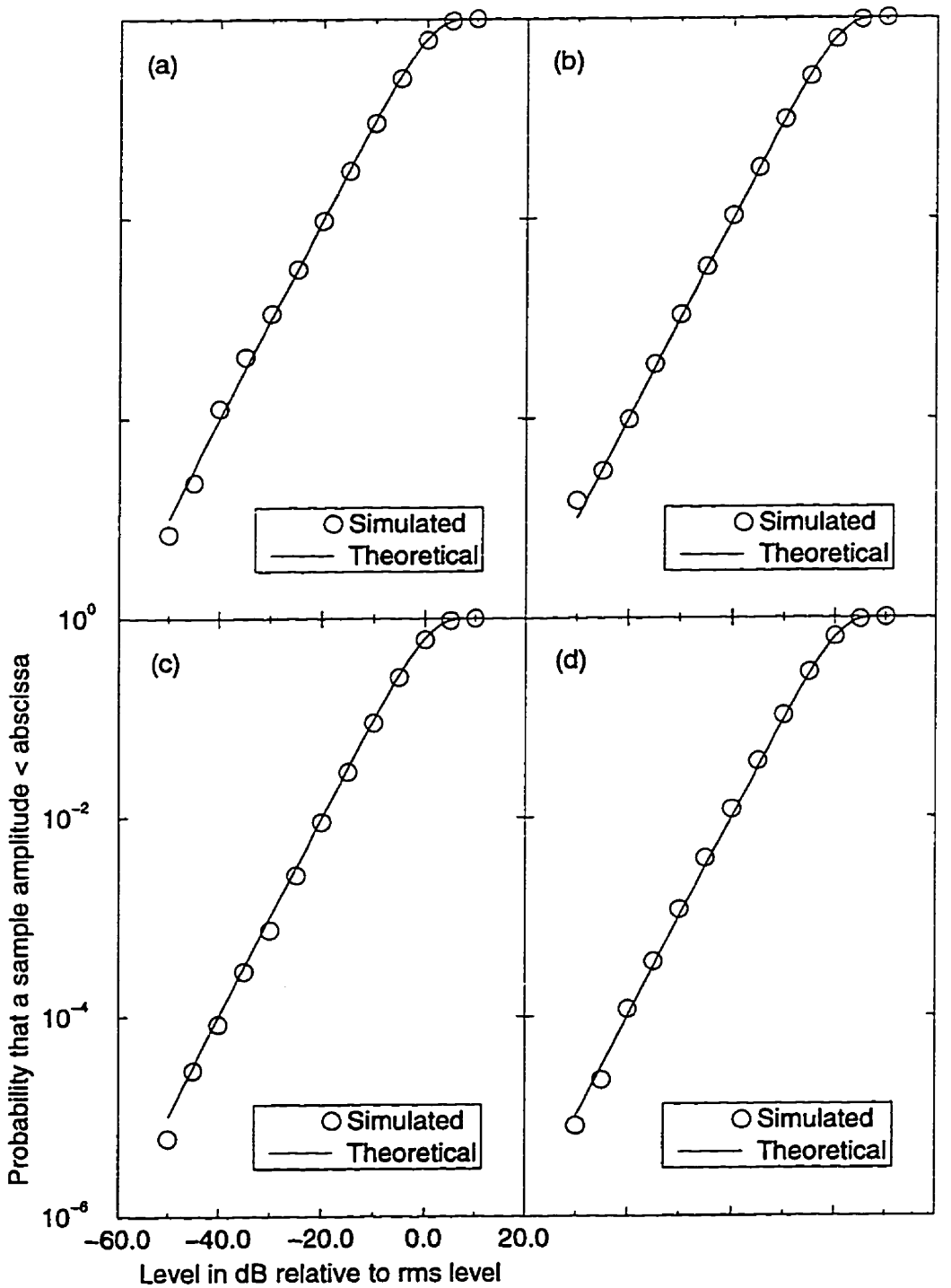

Figure 3.4: Comparison of the theoretical Rayleigh cdf with the cdf of simulated envelope samples for (a) $M = 6$, (b) $M = 10$, (c) $M = 20$, (d) $M = 40$ echos.

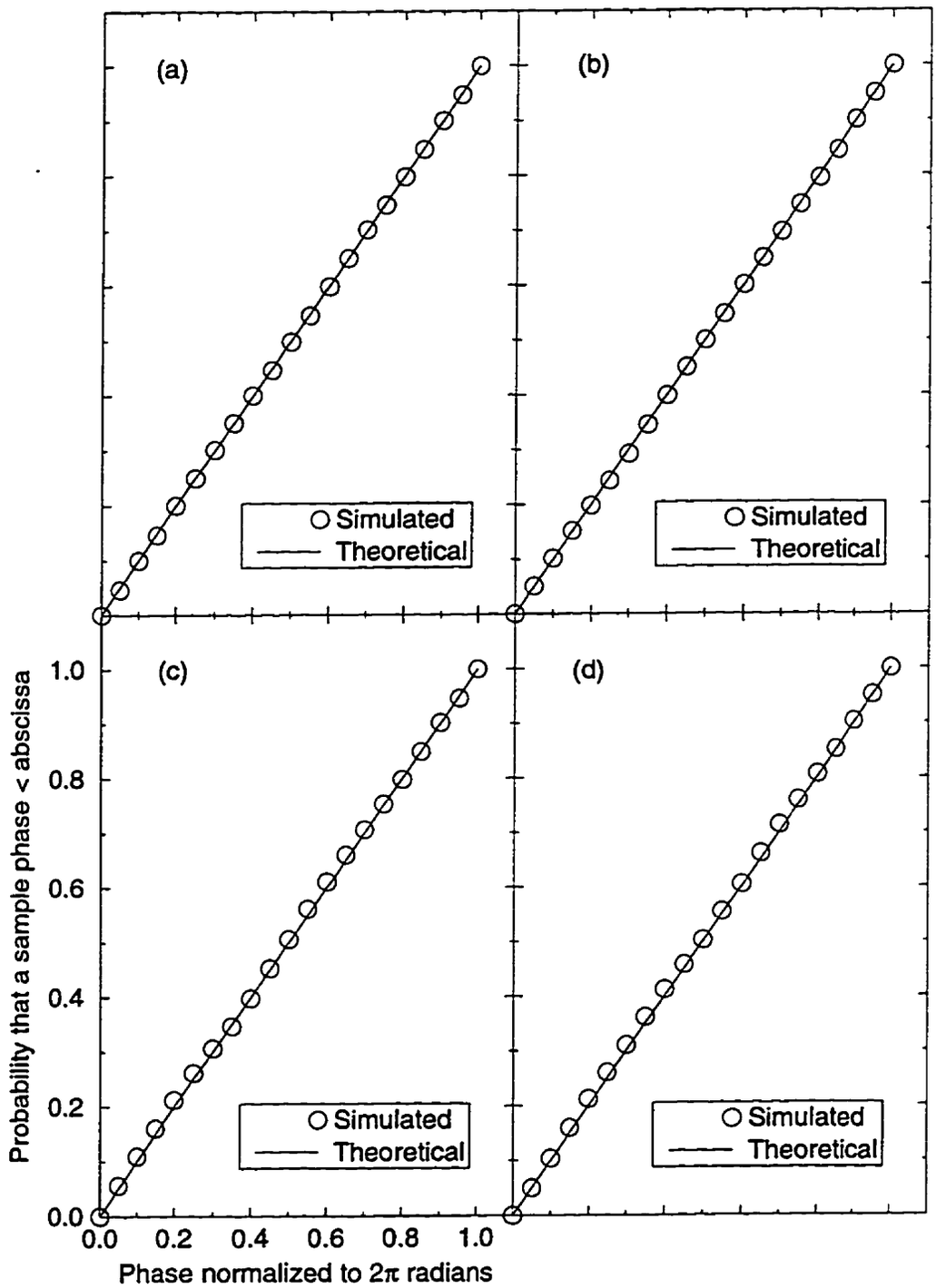

Figure 3.5: Comparison of the theoretical uniform cdf with the cdf of simulated phase samples for (a) $M = 6$, (b) $M = 10$, (c) $M = 20$, (d) $M = 40$ echos.

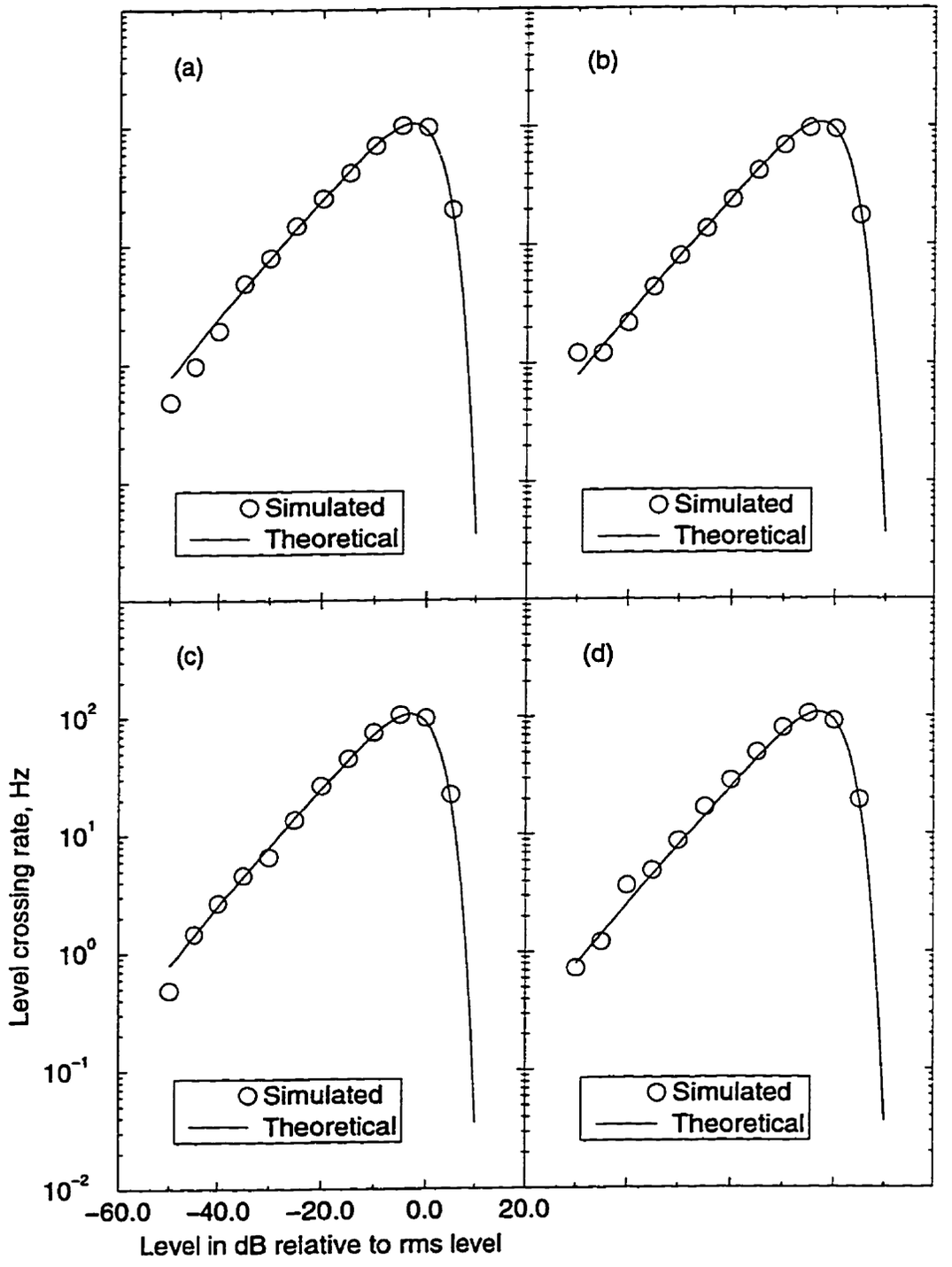

Figure 3.6: Comparison of the theoretical LCR with the LCR of simulated envelope samples for (a) $M = 6$, (b) $M = 10$, (c) $M = 20$, (d) $M = 40$ echos.

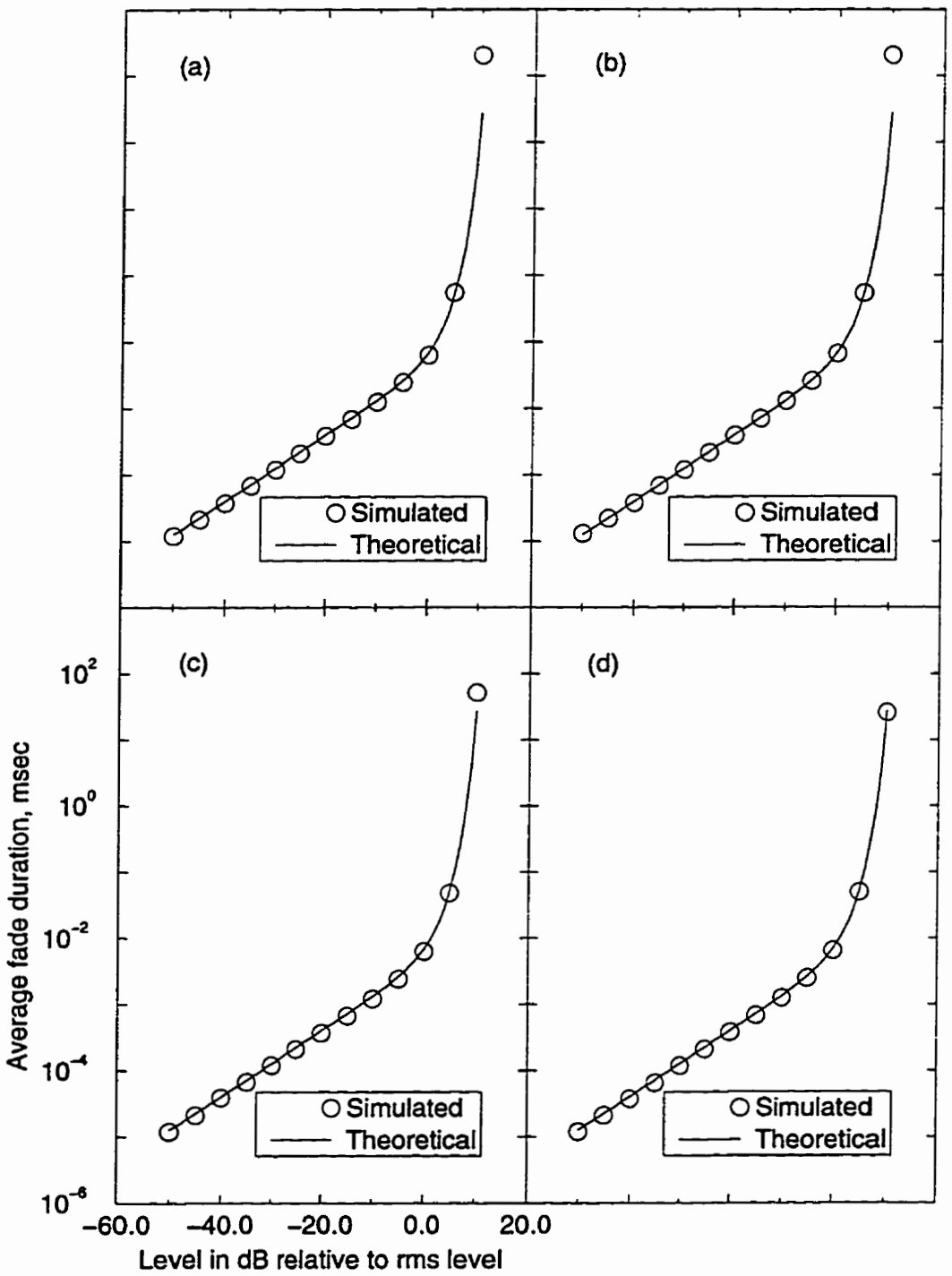

Figure 3.7: Comparison of the theoretical average duration of fades with the average duration of fades of simulated envelope samples for (a) $M = 6$, (b) $M = 10$, (c) $M = 20$, (d) $M = 40$ echos.

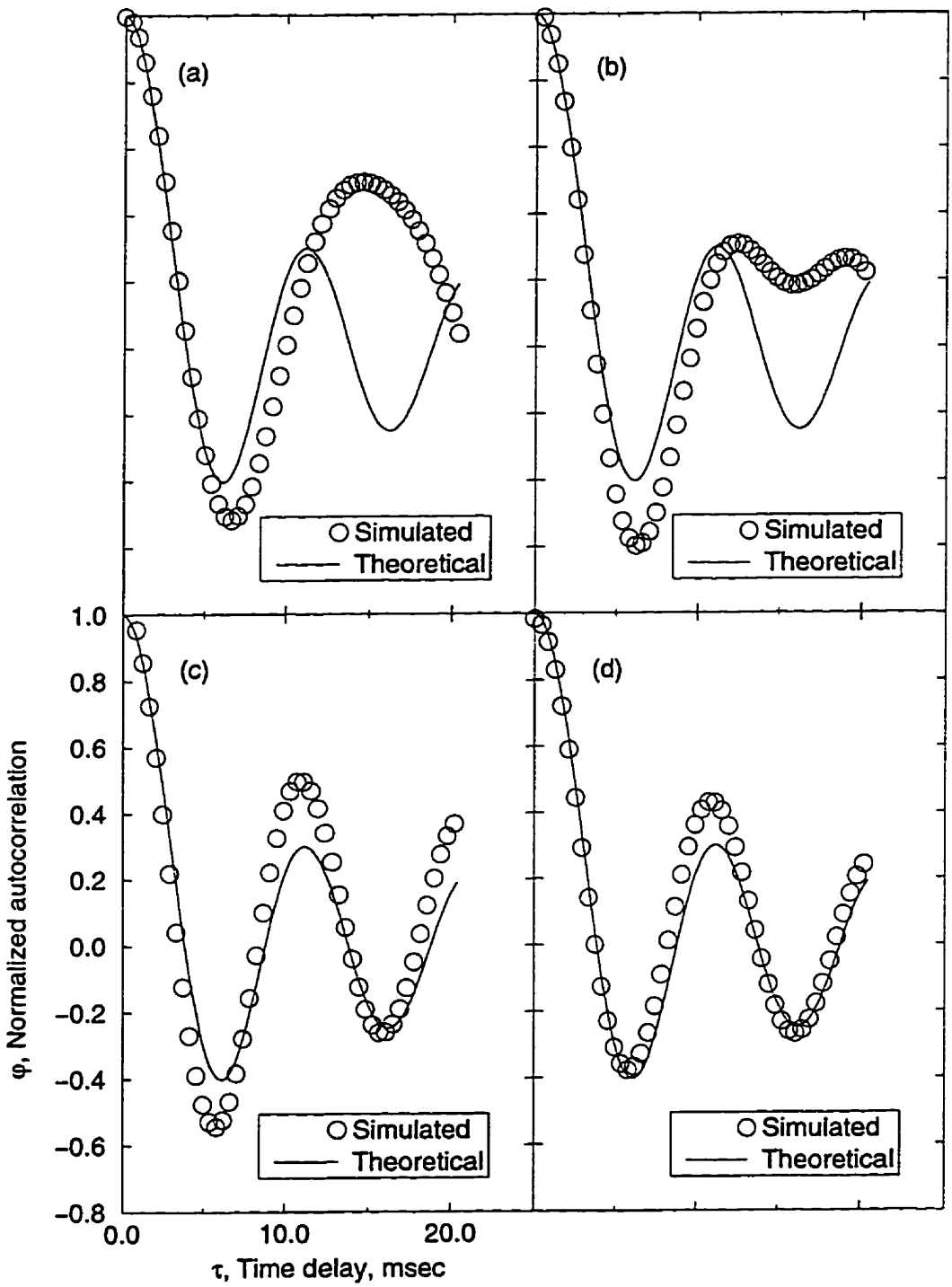

Figure 3.8: Comparison of the theoretical autocorrelation with the autocorrelation of the faded samples for (a) $M = 6$, (b) $M = 10$, (c) $M = 20$, (d) $M = 40$ echos.

for simulation times of 5 msec or less (for the channel configuration considered here). Since, in this study, the autocorrelation of the simulated samples must be good for simulation times of about 6.2 msec (corresponding to a block length of 150 symbols), $M = 40$ echos are chosen to model Rayleigh, flat-fading channels.

In a *frequency-selective* fading channel, each delayed echo should have Rayleigh statistics. For example, if a minimum of M echos are required to ensure Rayleigh statistics in a flat-fading channel, then a two-ray, double-spike model (see (3.2)) requires $2M$ echos (M echos for each ray). In a frequency-selective fading channel having a continuous delay profile, it is not as clear how many echos are required. (A new, more efficient, approach to the Monte Carlo-based fading simulator is presented in [117] in which the issue of ensuring Rayleigh statistics for all delayed echos is addressed.) Hoeher [92] presents examples of simulated, fast, frequency-selective fading channels where 500 echos are used "for illustration purposes," so this is the number of echos chosen to model fast, frequency-selective fading channels in this study.

3.5.3 The Correlation between Two Tones on the Frequency-Selective Fading Channel

One way to test the frequency selectivity of the simulated fading channel is to measure the correlation between two tones, separated in frequency by Δf Hz. Clarke [95] has shown that if the multipath delays have an exponential distribution (as is assumed in this study) with an rms delay spread of τ_{rms}, then the normalized correlation of the two tones as a function of frequency separation may be written as

$$|\varphi(\Delta f)| = \frac{1}{\sqrt{1 + (2\pi\Delta f\tau_{\text{rms}})^2}}. \tag{3.19}$$

This expression is compared to simulation results for four different values of τ_{rms} in a frequency-selective channel with an exponential delay profile, and the results are presented in Figure 3.9. It can be seen that the correlation of the two simulated tones is very close to theory for all delay spreads considered and beyond the frequency range

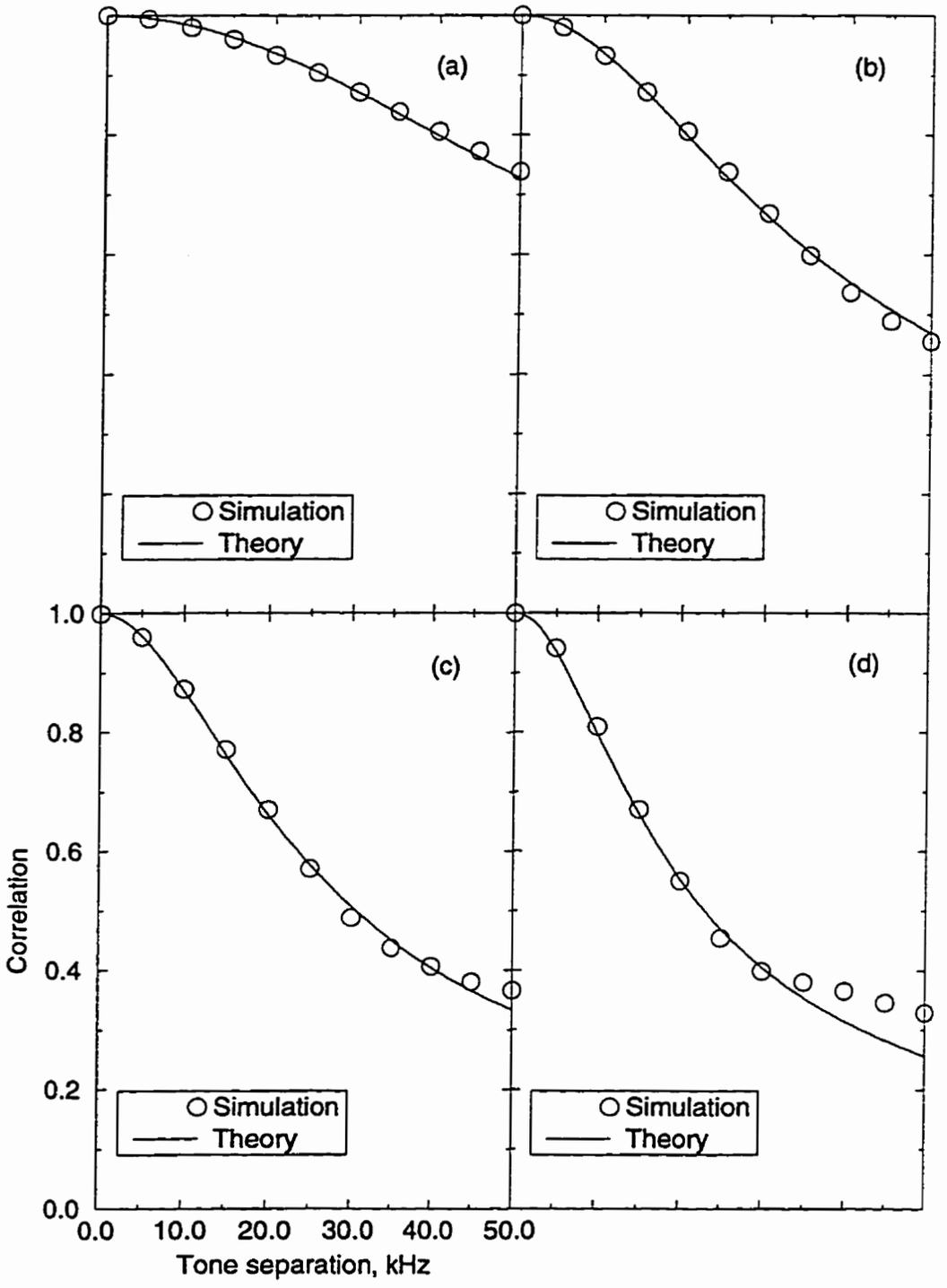

Figure 3.9: Comparison of the theoretical correlation of two tones with the simulated correlation of two tones for (a) $\tau_{\text{rms}} = 3$ μs, (b) $\tau_{\text{rms}} = 6$ μs, (c) $\tau_{\text{rms}} = 9$ μs, (d) $\tau_{\text{rms}} = 12$ μs.

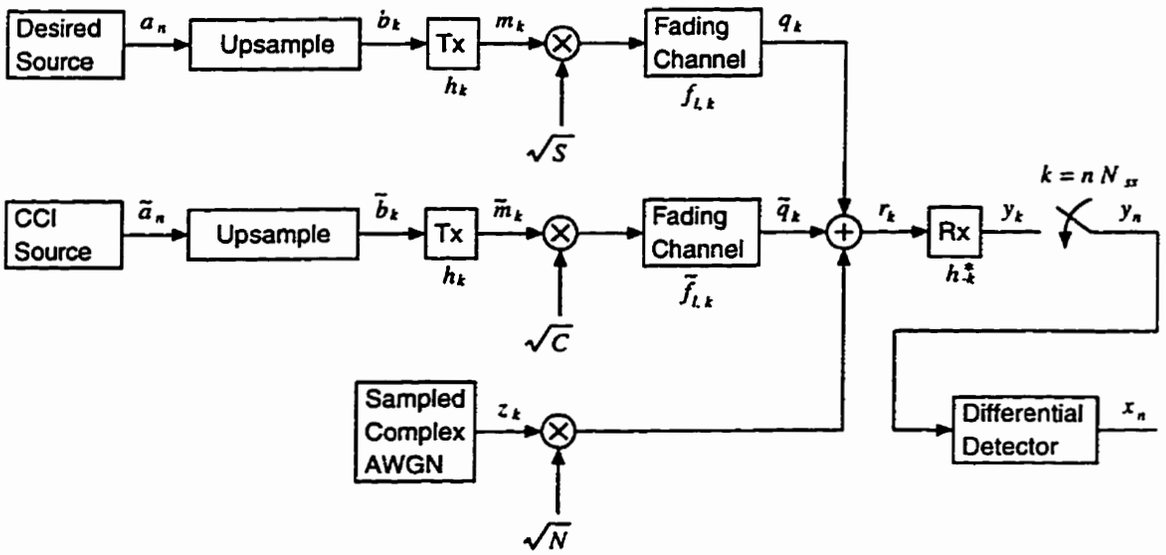

Figure 3.10: Block diagram of end-to-end system showing desired faded signal, faded CCI, and AWGN.

of interest[6].

3.6 System Models

In this section, the general system model is developed and the specific test channels used to evaluate the SIR estimators are described.

3.6.1 General System Model

A block diagram of the general system is provided in Figure 3.10 from which it may be seen that the system models a desired signal corrupted by a time-varying fading channel, a single cochannel interferer, and complex AWGN. The single interferer is

[6]Since root-raised cosine filtering with $\alpha = 0.35$ is employed in the transmitter and receiver, and since the baud rate is $48.6/2 = 24.3$ kbaud/s (see Table 3.1), the signal energy, ideally, only extends out (see pp. 535–536 of [65]) to

$$\frac{1+\alpha}{2T} = \frac{1+0.35}{2} \times 24.3 = 16.4 \text{ kHz}.$$

The "frequency range of interest" thus spans dc to 16.4 kHz.

subjected to a fading process similar to that experienced by the desired signal, but the two fading processes are assumed uncorrelated.

The discrete signals, a_n, b_k, m_k, and q_k are described in Section 3.4.2. The corresponding signals for the CCI, $\tilde{a}_n$, $\tilde{b}_k$, $\tilde{m}_k$, $\tilde{q}_k$, are entirely analogous. Further, $f_{l,k}$ represents the time-varying impulse response of the channel for the desired signal as given by (3.9); $\tilde{f}_{l,k}$ represents the corresponding fading channel for the CCI (a different set of fading parameters, θ_m, ν_m, and τ_m, is chosen to make the fading processes for the desired signal and CCI uncorrelated); h_k represents the tap coefficients of a root-raised cosine filter described in Section 3.4.2; and S, C, and N are real, positive scale factors for the desired signal, CCI, and AWGN, respectively.

Some assumptions regarding the first- and second-order statistics of the desired $\pi/4$-DQPSK-encoded symbols, a_n, the $\pi/4$-DQPSK-encoded CCI symbols, $\tilde{a}_n$, and the complex noise samples, z_k, must be made for the subsequent analysis. All three of the sequences are assumed zero-mean so that

$$E\{a_n\} = E\{\tilde{a}_n\} = E\{z_k\} = 0.$$

The autocorrelations of the desired symbol sequence and the CCI symbol sequence are assumed identical so that

$$E\{a_m a_n^*\} = E\{\tilde{a}_m \tilde{a}_n^*\} = \sigma_a^2 \delta_{mn}.$$

The power of the desired source symbols and the CCI symbols are also assumed identical such that

$$E\{|a_n|^2\} = E\{|\tilde{a}_n|^2\} = \sigma_a^2 = 1.$$

The sequence of desired symbols and the CCI sequence are assumed uncorrelated giving

$$E\{a_m \tilde{a}_n^*\} = 0.$$

The noise is assumed white with autocorrelation

$$E\{z_k z_l^*\} = \sigma_z^2 \delta_{kl},$$

and the total power in the complex noise is

$$E\{|z_k^2|\} = \sigma_z^2 = 1.$$

The noise samples are uncorrelated with both the information sequence and the CCI sequence so that

$$E\{a_m z_k^*\} = E\{\tilde{a}_m z_k^*\} = 0.$$

The received signal, r_k, at the input to the MF is

$$r_k = q_k + \tilde{q}_k + \sqrt{N} z_k \tag{3.20}$$

where q_k is given by (3.8), and $\tilde{q}_k$ is written similarly as

$$\tilde{q}_k = \sqrt{C}\, \tilde{m}_k \otimes \tilde{f}_{l,k}.$$

The received signal after the MF may be written as

$$\begin{aligned} y_k &= r_k \otimes h_{-k}^* \\ &= (q_k \otimes h_k) + (\tilde{q}_k \otimes h_k) + \sqrt{N}(z_k \otimes h_k) \\ &= \zeta_k + \tilde{\zeta}_k + \sqrt{N}\, w_k \end{aligned}$$

where $h_{-k}^* = h_k$ (since the RRC coefficients are assumed real and even), ζ_k is given by (3.10), $\tilde{\zeta}_k$ is given by

$$\tilde{\zeta}_k = \tilde{q}_k \otimes h_{-k}^* = \tilde{q}_k \otimes h_k,$$

and w_k is coloured noise expressed as

$$w_k = z_k \otimes h_{-k}^* = z_k \otimes h_k.$$

The statistics of the complex coloured noise, w_k, are entirely analogous to the statistics of the real coloured noise treated in Chapter 2 so that its mean is given by

$$E\{w_k\} = 0,$$

and its autocorrelation can be written as

$$\begin{aligned} E\{w_{k+l}w_k^*\} &= \sum_\alpha \sum_\beta h_\alpha h_\beta E\{z_{k+l-\alpha} z_{k-\beta}^*\} \\ &= \sum_\alpha h_\alpha h_{l-\alpha} \\ &= g_l \end{aligned}$$

so that the samples of w_k taken at the baud rate are uncorrelated (white). As in Chapter 2, it is assumed that $g_0 = 1$.

The optimally-sampled output of the MF can be expressed as

$$y_n = y_k|_{k=nN_{ss}} = (\zeta_k + \tilde{\zeta}_k + \sqrt{N}w_k)\Big|_{k=nN_{ss}}, \; n \in \{0, 1, \ldots, N_{\text{sym}}\}. \tag{3.21}$$

It is difficult to separate (3.21) into two terms whose ratio represents the SIR. More specifically, it is difficult to separate ζ_k into a "good" term and an "impairment" term, unlike the CCI and noise terms in (3.21) which are clearly impairments. It is not valid simply to declare that the strongest received, desired multipath ray is the true desired signal and all other multipath components are pure impairments that adversely affect BER performance. On the contrary, though all delayed multipath rays contribute to ISI, the overall effect of some multipath rays (especially the earliest-arriving delayed rays) is to add to the signal power which can counter their contribution to ISI. Also, the strongest received multipath component itself cannot be identified as a pure signal term since it is deleteriously affected, in general, by the random phase and amplitude fluctuations of fast, Rayleigh fading.

Instead of debating which part of ζ_k is pure signal and which parts are impairments, it is simply assumed that it is possible to re-write (3.21) as[7]

$$y_n = \zeta_{d_n} + \xi_n, \; n \in \{0, 1, \ldots, N_{\text{sym}}\} \tag{3.22}$$

where ζ_{d_n} is the pure, desired signal term and a new variable, ξ_n, is defined to represent the sum of all of the impairments. Note that if the channel has no fading and if there

[7] In the sequel, it is understood that $N_{ss} = 1$; that is, only symbol-spaced samples are considered so that "N_{ss}" may be omitted from the subscripts of ζ_{d_n} and ξ_n.

is no CCI present, then

$$y_n = \zeta_{d_n} + \xi_n = \sqrt{S}a_n + \sqrt{N}w_n,$$

where the signal term results from a simplification of (3.10) assuming no fading (i.e. $\psi_{k,l} = g_0 = 1$).

The assumed first- and second-order statistics of ζ_{d_n} are given by

$$E\{\zeta_{d_n}\} = 0 \tag{3.23a}$$

$$E\{\zeta_{d_m}\zeta_{d_n}^*\} = S\delta_{mn} \tag{3.23b}$$

$$E\{|\zeta_{d_n}|^2\} = S \tag{3.23c}$$

$$E\{\xi_m\zeta_{d_n}^*\} = 0. \tag{3.23d}$$

For simplicity, the very gross assumption is made that ξ is Gaussian, and that its first- and second-order statistics are given by

$$E\{\xi_n\} = 0 \tag{3.24a}$$

$$E\{\xi_m\xi_n^*\} = \Xi\delta_{mn} \tag{3.24b}$$

$$E\{|\xi_n|^2\} = \Xi \tag{3.24c}$$

$$E\{\zeta_{d_m}\xi_n^*\} = 0 \tag{3.24d}$$

where Ξ represents the total impairment power. Of course, the Gaussian assumption is invalid, in general, for all of the impairment processes other than AWGN; however, for the sake of simplicity and unless otherwise mentioned, the estimators are derived assuming Gaussian statistics, and then applied *ad hoc* to channels with non-Gaussian impairments.

Most of the SIR estimators are derived in terms of the symbols at the output of the differential detector (DD) given by

$$x_n = y_n y_{n-1}^*, \; n \in \{1, 2, \ldots, N_{\text{sym}}\}. \tag{3.25}$$

3.6.2 Five Mobile Radio Test Channels

The aim of this study is to investigate statistically SIR estimator performance given a particular set of channel conditions. Each SIR estimate is generated from a block of $\pi/4$-DQPSK-encoded symbols corrupted by fading, CCI, and AWGN. For each successive SIR estimate, a new random sequence of data, AWGN, and/or CCI is generated. Five test channels are chosen in which to evaluate the performances of the SIR estimators, and their characteristics are described in this section. For convenience, the five channels are referred to here as Cases 1–5.

The parameters used to define the characteristics of each channel are the SNR, S/C (the ratio of desired signal power to CCI power), $\nu_{\max}$, and τ_{rms}. The number of information symbols used to generate an SIR estimate is fixed at $N_{\text{sym}} = 150$ (which implies $N_{\text{sym}} + 1 = 151$ differentially-encoded symbols) for all cases. The parameters chosen for each of the five test cases were selected by referring to "typical" values for SNR, S/C, $\nu_{\max}$, τ_{rms}, and BER reported in the literature [74, 2, 11, 97, 100]. In each case, a particular channel parameter, such as the SNR, is swept (in discrete steps), and the SIR is measured using a particular SIR estimator. The sweep ranges of the channel parameters are chosen such that the BER of each test channel is some "reasonable" value in a range from 10^{-1} to 10^{-5}. The simulated BER curves for each test channel are presented in Section 3.9.

It is illustrative to see examples of the effect of each channel on the constellation of the received, corrupted, complex baseband signal. For each of the five test channels, four different types of signal constellations are shown constructed from samples of simulation data. The different constellations portray:

- Typical received samples (16 samples per symbol) at the output of the MF.
- Symbol centres before the DD.
- Symbol centres after the DD.
- Symbol centres de-rotated to the real, positive axis (modulation removed).

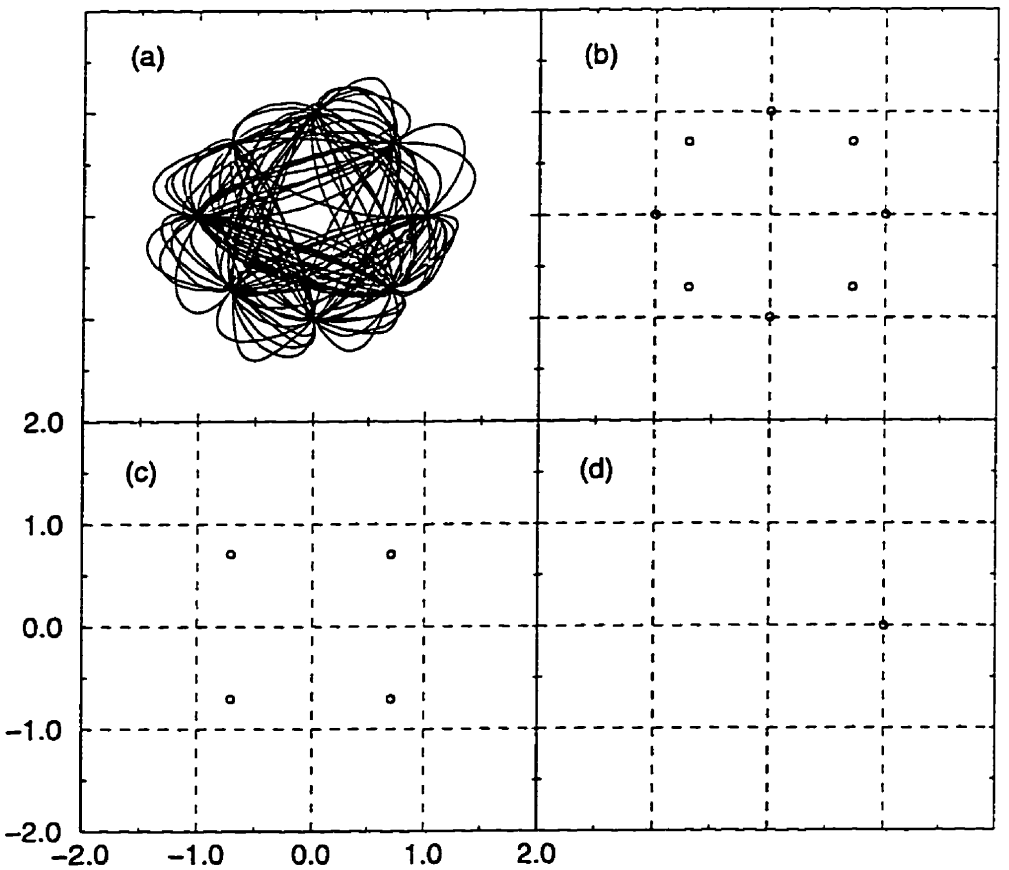

Figure 3.11: "Clear sky" reference signal (a) samples after MF, (b) symbol centres before DD, (c) symbol centres after DD, (d) post-DD symbols with modulation removed.

The "de-rotation" referred to above is an effective removal of the modulation from the baseband signal after the DD. This de-rotation concept arises in the derivation of some SIR estimators in Section 3.7. For reference, Figure 3.11 shows each of the four constellations as they would appear in a "clear-sky" channel with no noise, no CCI, and no fading.

Case 1: Complex AWGN Channel

The only impairment in this channel is complex AWGN. This is the same complex AWGN channel used in Chapter 2. The sample signal constellations for Case 1 are

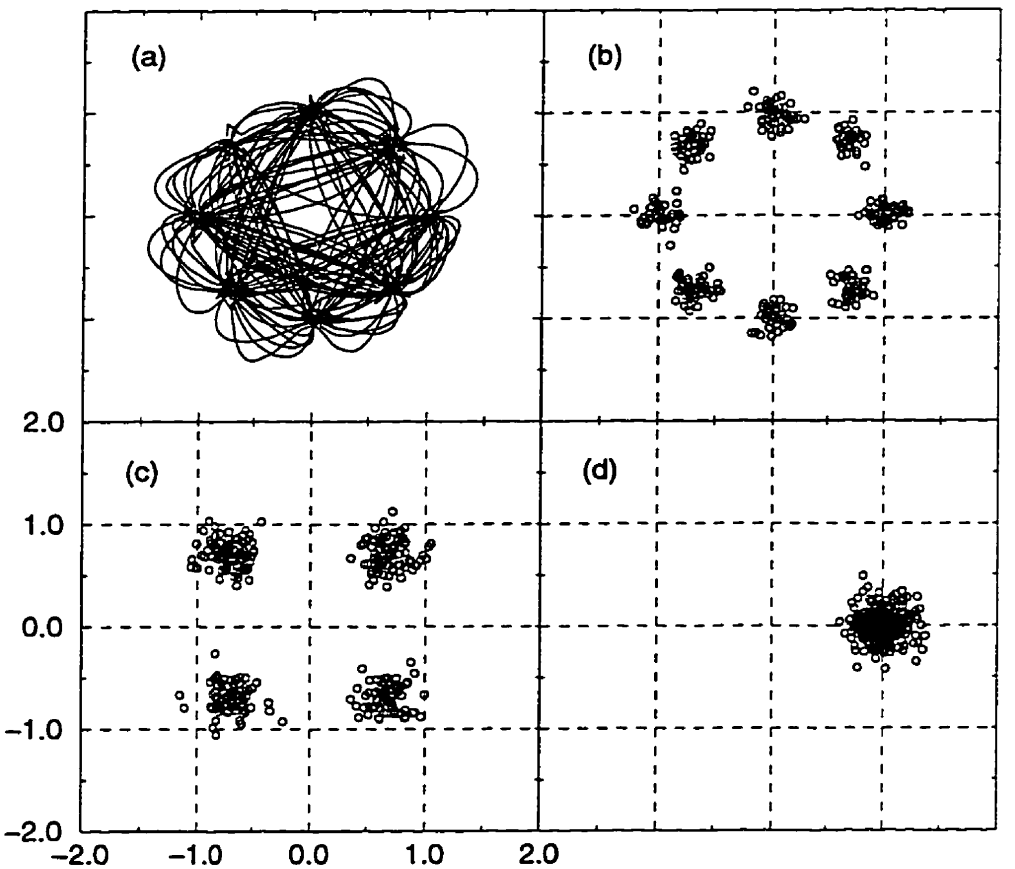

Figure 3.12: Received signal for Case 1 with $S/N = 18$ dB (a) samples after MF, (b) symbol centres before DD, (c) symbol centres after DD, (d) post-DD symbols with modulation removed.

presented in Figure 3.12 with $S/N = 18$ dB.

Case 2: Fast, Flat Fading Channel

This channel models fast Rayleigh fading in a flat channel with AWGN but no CCI. The Doppler power spectrum is the Jakes spectrum given earlier in (3.1). The spacing between successive amplitude fades due to the fast Rayleigh fading is about the same or greater than the block length of $N_{\text{sym}} = 150$ symbols, depending on the specific value of ν_{max}, which creates a modeling problem peculiar to this test channel (and Case 5, which is discussed later). Specifically, the amplitude and phase fluctuations

of the Rayleigh model can either be left to evolve over time for each successive SIR estimate, or the channel conditions can be "replayed" for each successive SIR estimate. If the latter model is selected, a specific start time must be chosen.

For the Case 2 channel, the decision was made to simulate both models. The channel that is allowed to evolve over time for each successive SIR estimate is referred to as the *Case 2A* channel, and the *Case 2B* channel is the name used to refer to the model in which the channel conditions are replayed identically for every SIR estimate. The Case 2A channel is more realistic since it models a real-world scenario, but a severe drawback is the fact that the "true" SIR varies considerably about some long-term mean over time (since each SIR estimate is allowed to see only an N_{sym}-wide window) which creates a false, misleadingly high variance in the SIR estimates. In contrast, the variance of the SIR estimates measured in the Case 2B model is attributable solely to the specific SIR estimator tested and not to time-dependent channel conditions; however, the model does not reflect well the dependence of the SIR on the Doppler spread, ν_{max}, due to the relatively narrow observation window.

For the Case 2A model, the range of ν_{max} and the other channel variables must be specified. In Figure 7 of [11], it can be seen that for Rayleigh-faded, $\pi/4$-DQPSK modulated signals, the effect of ν_{max} is negligible for values of SNR less than 30 dB, but the effect of ν_{max} for values of SNR greater than about 60 dB is very distinct. For this reason, for the Case 2A model, S/N is held constant at 70 dB[8]. The Doppler spread, ν_{max}, is swept from 40 Hz to 100 Hz in 30 Hz steps. An example of typical amplitude and phase fluctuations observed in this Case 2A channel for $\nu_{\text{max}} = 100$ Hz was given in Figure 3.3.

For the Case 2B model, ν_{max} is held constant at 100 Hz while S/N is swept from 20 dB to 40 dB in increments of 5 dB. This lower range of SNR was chosen since the error probability is relatively unchanged for $S/N > 50$ dB with ν_{max} fixed. The ampli-

[8]It is understood that it may be difficult to find a real-world Rayleigh fading channel with an associated SNR of 70 dB, but this somewhat artificial value of SNR is assumed here in order to observe the effect of the Doppler spread on the SNR.

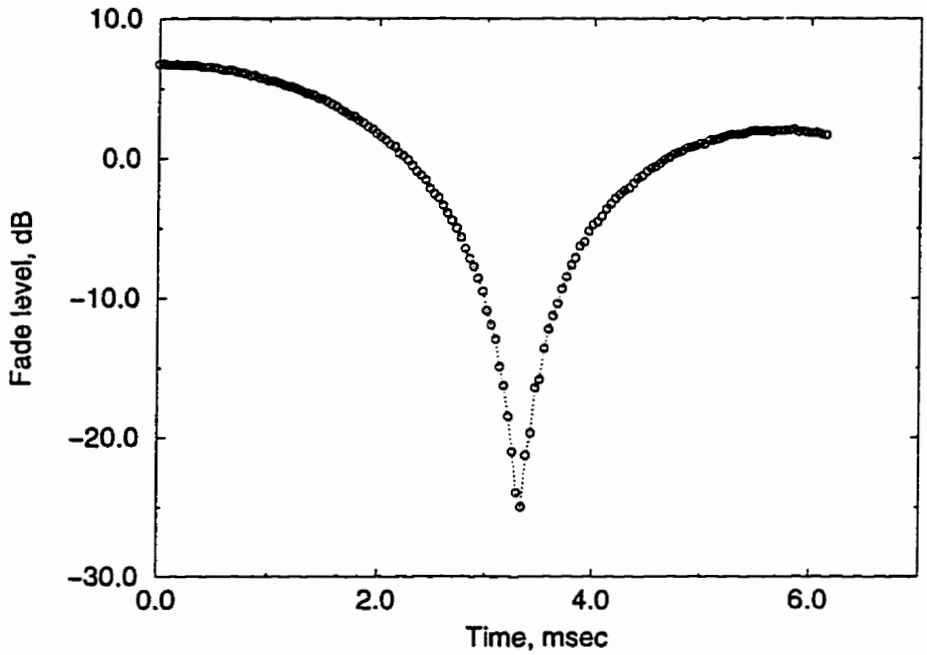

Figure 3.13: Envelope of received signal due to Rayleigh fading in the Case 2B channel over an interval of 150 symbols (6.18 μs). This amplitude fluctuation is replayed for the generation of each SIR estimate.

tude and phase fluctuations due to the fast fading are replayed every $N_{\text{sym}} + 1 = 151$ differentially-encoded symbols so that each SIR estimate is generated under identical channel conditions. The start time is chosen such that a deep null occurs near the center of each received block. The amplitude fading profile is given in Figure 3.13. The spacing between successive fades is about equal to $N_{\text{sym}} = 150$ symbols.

The sample signal constellations for Case 2A are presented in Figure 3.14 with $\nu_{\text{max}} = 100$ Hz, and $S/N = 40$ dB. The signal constellations for Case 2B are omitted since the effect of Rayleigh fading on the signal constellations is best appreciated by observing the effect of the time-evolving channel of Case 2A.

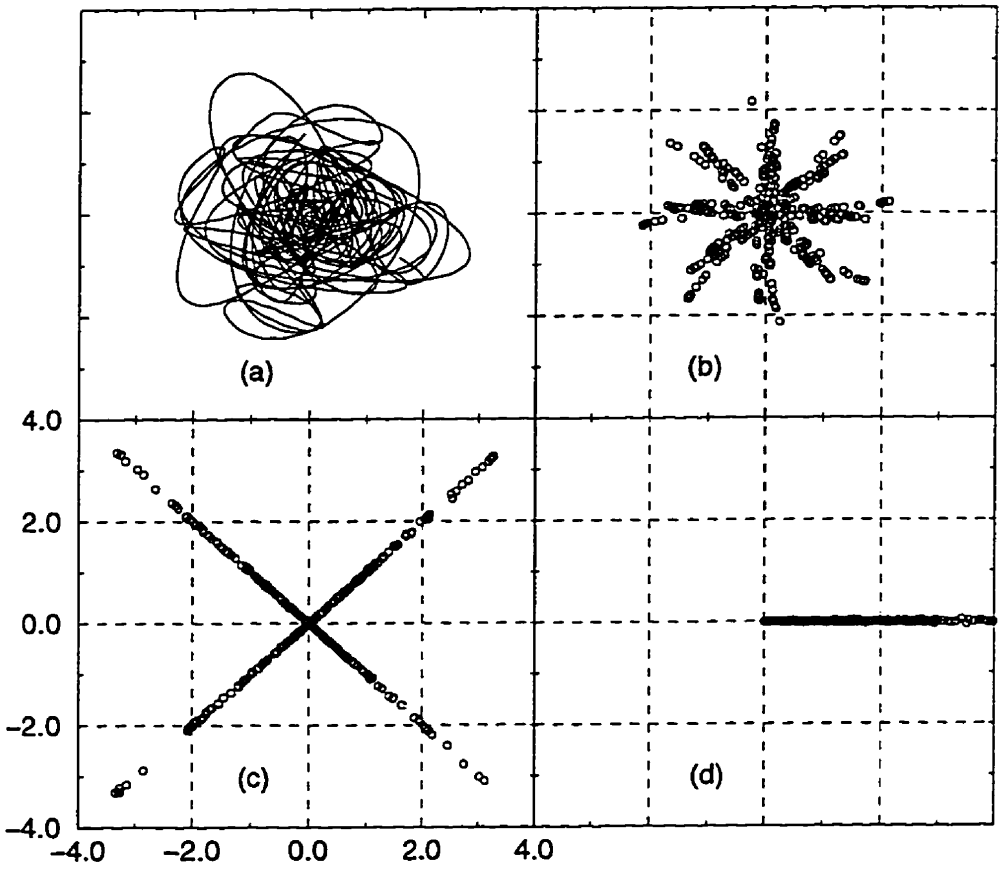

Figure 3.14: Received signal for Case 2 with $\nu_{\max} = 100$ Hz and $S/N = 40$ dB (a) samples after MF, (b) symbol centres before DD, (c) symbol centres after DD, (d) post-DD symbols with modulation removed.

Case 3: Slow, Frequency-Selective Fading Channel

In this case, it is assumed that the transmitter and receiver are not moving with respect to one another so that $\nu_{\max} = 0$, but there is channel distortion due to multipath-induced ISI. In other words, the channel is frequency-selective. The channel is assumed to have an exponential delay power spectrum given by the first line of (3.2). The delay spread of the channel is swept from $\tau_{\text{rms}} = 3$ μs to 12 μs in 3 μs steps, while S/N is held constant at 15 dB. This small amount of AWGN is included since the multipath-induced ISI alone produces a very small BER, especially at small values of τ_{rms}, which takes an excessive amount of simulation time.

The sample signal constellations for Case 3 are presented in Figure 3.15 with $\tau_{\text{rms}} = 9$ μs and $S/N = \infty$. No noise has been added to the signal used to generate these plots so that the effect of the multipath-induced ISI is isolated. There is a downward shift in the post-DD constellation (see Figure 3.15(c)) which brings the upper two post-DD symbol centres closer to a decision boundary. Note, however, that even with $\tau_{\text{rms}} = 9$ μs, though the post-DD symbol-centres are smeared and come close to the decision-boundaries, not one of the symbols pictured would be detected incorrectly—there is only a very small probability of a decision error. It is for this reason that a small amount of noise ($S/N = 15$ dB) is added for the SIR simulations.

Case 4: Single Cochannel Interferer

This case assumes a single, $\pi/4$ DQPSK-modulated cochannel interferer, which implies the interfering signal has the same statistics as the desired signal. The worst-case timing scenario is assumed where the symbols of the desired signal and the symbols of the CCI are precisely time-synchronous. The desired transmitter, the cochannel interferer, and the receiver are assumed fixed in position so that the channel transfer function is time-invariant. Some AWGN ($S/N = 15$ dB) is included, however, as the BER curve due to CCI alone has a "brick-wall" shape. That is, in the absence of any

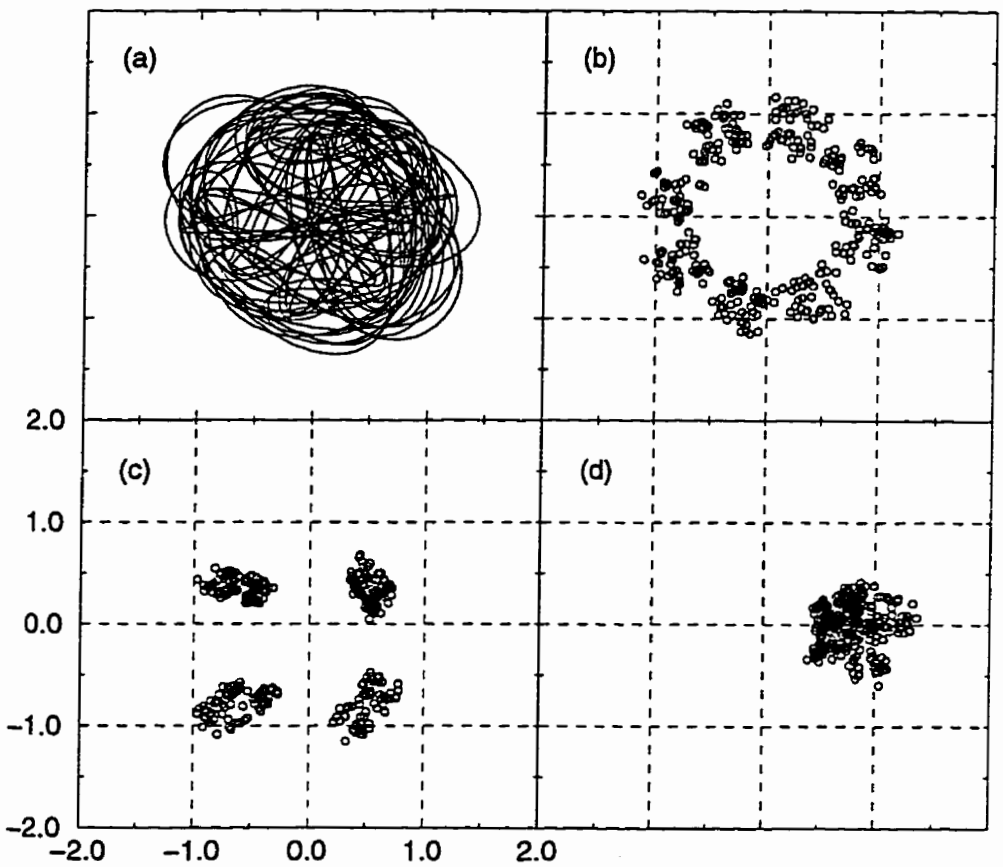

Figure 3.15: Received signal for Case 3 with $\tau_{\text{rms}} = 9$ μs and $S/N = \infty$ (a) samples after MF, (b) symbol centres before DD, (c) symbol centres after DD, (d) post-DD symbols with modulation removed.

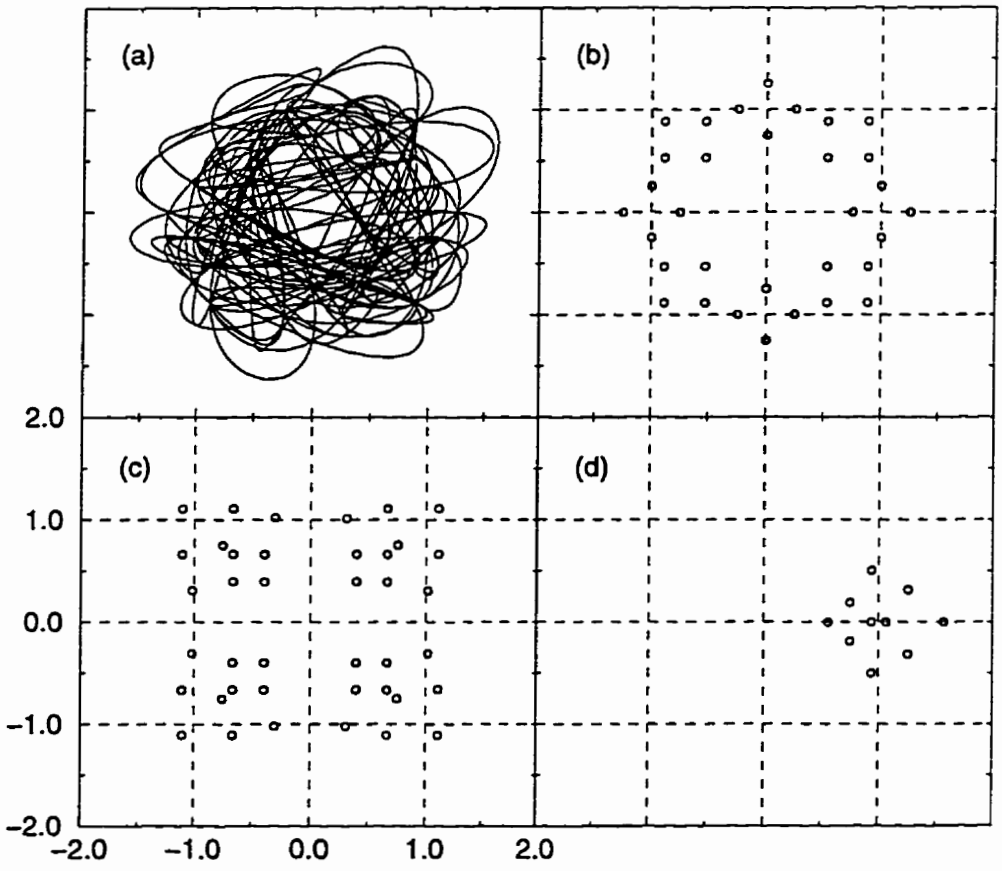

Figure 3.16: Received signal for Case 4 with $S/C = 12$ dB and $S/N = \infty$ (a) samples after MF, (b) symbol centres before DD, (c) symbol centres after DD, (d) post-DD symbols with modulation removed.

other impairment, as S/C is decreased from infinity, the BER remains zero until S/C reaches a threshold at which point the BER increases dramatically. Injecting some noise causes a more gradual degradation in BER with decreasing S/C.

The sample signal constellations for Case 4 are presented in Figure 3.16 with $S/C = 12$ dB, and $S/N = \infty$. No noise has been added to the signal used to generate these plots so that the effect of the CCI is isolated. Some interesting patterns arise from the perfectly time-synchronized CCI. From these plots it is easy to see that, in the absence of noise, the BER has a threshold. No errors are produced in the scenario shown since the patterns of points after the DD (see Figure 3.16(c)) are contained

entirely within their respective quadrants. It is only when the CCI power becomes large enough such that the four clusters encroach into other quadrants that errors are produced. Once that threshold is crossed, the BER degrades dramatically.

Case 5: A "Typical" Mobile Radio Channel

The last case considered is a combination of the various channel degradations. Both the signal and CCI are subjected to similar, but uncorrelated fading conditions. That is, the shapes of the Doppler and delay power spectra are identical, and the parameters $\nu_{\max}$ and τ_{rms} are equal in both channels, but the sets of random variables, $(\theta_m, \nu_m, \tau_m)$, $m \in \{1, 2, \ldots, M\}$, in (3.9) are chosen independently.

Since fast Rayleigh fading is present in this test channel, the same issue arises as discussed earlier in the context of the Case 2 channel regarding whether to allow the channel conditions to evolve over time, or to replay the exact same channel conditions for each successive SIR estimate. For the sake of conciseness, only one type of model need be implemented for the Case 5 channel. A model analogous to the Case 2B channel is use here; that is, the channel conditions are replayed identically for each SIR estimate so that the inherent effect of an estimator algorithm on the variance of the SIR estimates is isolated from the effect of the time-varying channel conditions.

The start time for the simulation is chosen so that a deep null, such as is present in Case 2B, is avoided since the effect of the null swamps out the effect of the other channel impairments. Instead, a segment of time is chosen where the amplitude fluctuations are less severe. The amplitude profile chosen for Case 5 is shown in Figure 3.17. In order to see the amplitude fluctuation more clearly, all other impairments were omitted for the generation of Figure 3.17.

The sample signal constellations for Case 5 are presented in Figure 3.18 with $\nu_{\max} = 100$ Hz, $\tau_{\text{rms}} = 3$ μs, $S/C = 18$ dB, and $S/N = 15$ dB. For the Case 5 simulations, the SNR is the independent variable which assumes values from the set $\{9, 12, 15, 30\}$ dB.

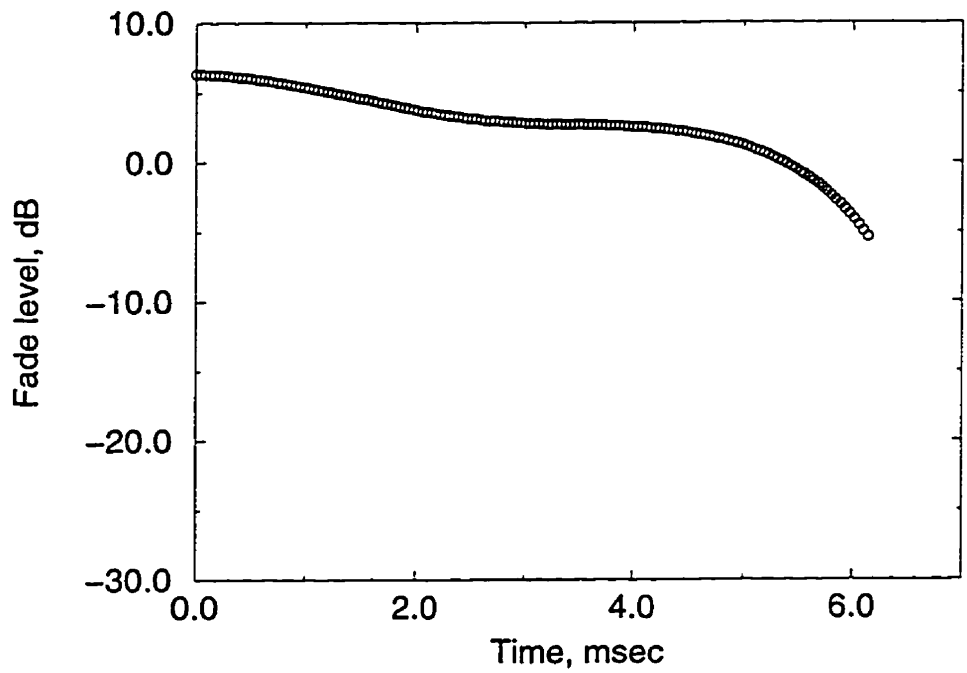

Figure 3.17: Envelope of received signal due to Rayleigh fading in the Case 5 channel over an interval of 150 symbols (6.18 μs). This amplitude fluctuation is replayed for the generation of each SIR estimate.

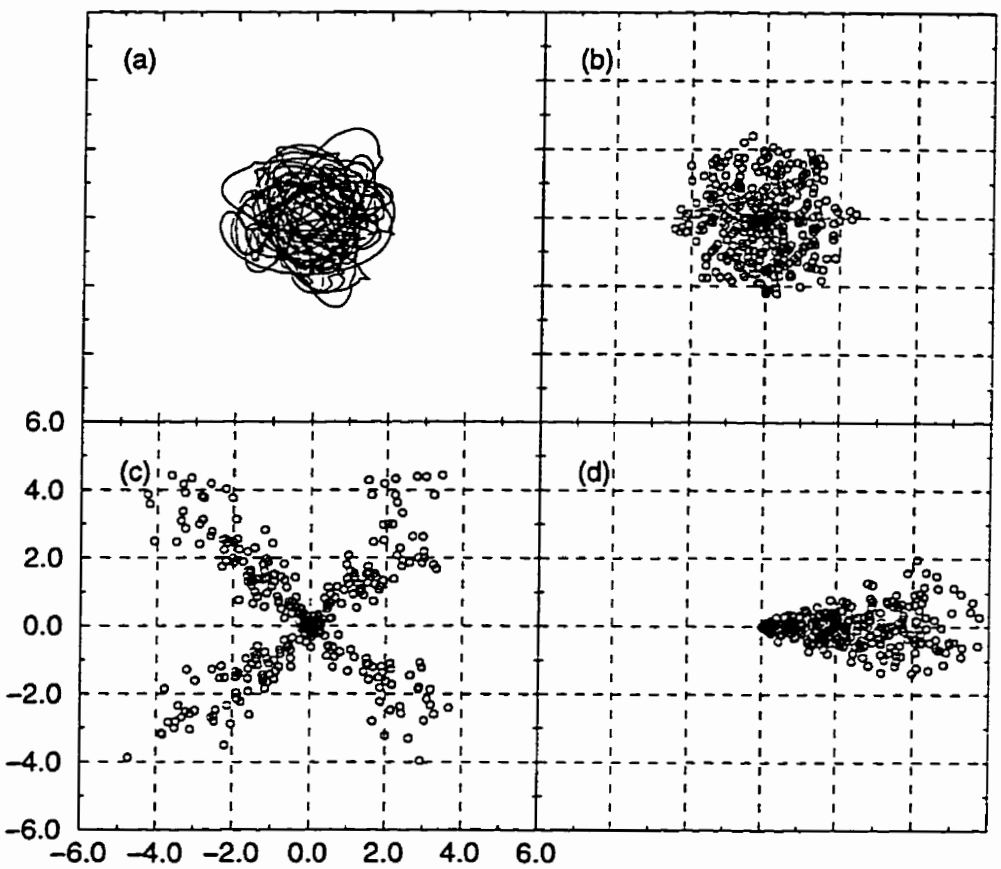

Figure 3.18: Received signal for Case 5 with $\nu_{\max} = 100$ Hz, $\tau_{\text{rms}} = 3$ μs, $S/C = 18$ dB and $S/N = 15$ dB (a) samples after MF, (b) symbol centres before DD, (c) symbol centres after DD, (d) post-DD symbols with modulation removed.

Channel	$\nu_{\max}$, Hz	τ_{rms}, μs	S/C, dB	S/N, dB
Case 1	0	0	∞	$\{6, 9, 12, 15, 18, 24, 30\}$
Case 2A	$\{30, 70, 100\}$	0	∞	70
Case 2B	100	0	∞	$\{20, 25, 30, 35, 40\}$
Case 3	0	$\{3, 6, 9, 12\}$	∞	15
Case 4	0	0	$\{6, 9, 12, 18\}$	15
Case 5	100	3	18	$\{9, 12, 15, 30\}$

Table 3.2: Summary of channel characteristics for the five types of test channels.

The channel parameters for the five cases are summarized in Table 3.2.

3.7 SIR Estimators under Study

The SIR estimators under study are described in this section. Some of the estimators are adapted from Section 2.3, others make their first appearance here. Where possible, estimator names and acronyms are those used by the original authors; otherwise, names and acronyms are chosen for ease of reference.

3.7.1 The Second- and Fourth-Order Moments (M_2M_4) Estimator

The M_2M_4 estimator was described in Section 2.3.4. The form of the M_2M_4 estimator applicable to complex signals in complex AWGN is adapted here to the model described in Section 3.6.1. The M_2M_4 estimator is applied to the baud-spaced samples at the output of the MF (but before the DD) given by (3.22) as

$$y_n = \zeta_{d_n} + \xi_n, \ n \in \{0, 1, \ldots, N_{\text{sym}}\}.$$

The general, complex M_2M_4 estimator for the desired signal power is given by (2.98) of Section 2.3.4 as

$$\hat{S}_{M_2M_4,\text{complex}} = \frac{M_2(k_w - 2) \pm \sqrt{(4 - k_a k_w){M_2}^2 + M_4(k_a + k_w - 4)}}{k_a + k_w - 4}, \qquad (3.26)$$

and the general, complex M_2M_4 estimator for the noise power is given by (2.99) of Section 2.3.4 as

$$\hat{N}_{M_2M_4,\text{complex}} = M_2 - \hat{S}_{M_2M_4,\text{complex}}, \tag{3.27}$$

where the constants, k_a and k_w, represent the kurtosis of the signal samples and the kurtosis of the noise samples, respectively.

In the context of the model of Section 3.6.1, the signal power estimator given by (3.26) is re-written as

$$\hat{S}_{M_2M_4} = \frac{M_2(k_\xi - 2) \pm \sqrt{(4 - k_a k_\xi)M_2{}^2 + M_4(k_a + k_\xi - 4)}}{k_a + k_\xi - 4}, \tag{3.28}$$

where k_ξ is the kurtosis of the impairment process, and M_2 and M_4 are given in their time-average forms as

$$M_2 = \frac{1}{N_{\text{sym}} + 1} \sum_{n=0}^{N_{\text{sym}}} |y_n|^2,$$

and

$$M_4 = \frac{1}{N_{\text{sym}} + 1} \sum_{n=0}^{N_{\text{sym}}} |y_n|^4$$

which are similar to (2.101) and (2.102), respectively, of Section 2.3.4. The noise power estimator given by (3.27) becomes the *impairment* power estimator expressed as

$$\hat{\Xi}_{M_2M_4} = M_2 - \hat{S}_{M_2M_4}. \tag{3.29}$$

The kurtosis of the $\pi/4$-DQPSK-modulated symbols is $k_a = 1$ and, if it is assumed that ξ_n is a complex, Gaussian process, it's kurtosis is $k_\xi = 2$. However, if ξ_n is dominated by CCI so that all other impairments (including AWGN) are negligible, the kurtosis of ξ_n is $k_\xi = 1$ since the CCI is also a $\pi/4$-DQPSK-modulated signal. Clearly, the optimal form of the M_2M_4 estimator depends on the statistics of the impairment process so that either the M_2M_4 estimator must be able to identify different types of impairment processes and adapt accordingly, or some compromise must be made so that a static M_2M_4 estimator is applicable to a range of channel environments. The latter approach is taken here. Specifically, the compromise made is to let $k_\xi = 1.5$

so that, using (3.28) (taking the negative root) and (3.29), an M_2M_4 SIR estimator may be expressed as

$$\hat{\rho}_{\mathrm{M_2M_4}} = \frac{\hat{S}_{M_2M_4}}{\hat{\Xi}_{M_2M_4}} = \frac{1}{3}\left[\frac{M_2 + \sqrt{10{M_2}^2 - 6M_4}}{2M_2 - \sqrt{10{M_2}^2 - 6M_4}}\right]. \tag{3.30}$$

3.7.2 The Signal-to-Variation Ratio (SVR) Estimator

The SVR estimator was described in Section 2.3.5. The form of the SVR estimator applicable to complex signals in complex AWGN is adapted here to the model described in Section 3.6.1. Just as the M_2M_4 estimator, the SVR estimator is applied to the $N_{\mathrm{sym}} + 1$ baud-spaced samples at the output of the MF (but before the DD) given by (3.22).

The general, complex SVR estimator is adapted from (2.112) of Section 2.3.5 by replacing k_w with k_ξ to yield

$$\hat{\rho}_{\mathrm{SVR,complex}} = \frac{(\beta - 1) \pm \sqrt{(\beta-1)^2 - [1 - \beta(k_a - 1)][1 - \beta(k_\xi - 1)]}}{1 - \beta(k_a - 1)} \tag{3.31}$$

where k_ξ is defined in Section 3.7.1, k_a is the kurtosis of the signal, and β is given in time-average form as

$$\beta = \frac{\frac{1}{N_{sym}}\sum_{n=1}^{N_{sym}} |y_n|^2 |y_{n-1}|^2}{\frac{1}{N_{sym}}\sum_{n=1}^{N_{sym}} |y_n|^4 - \frac{1}{N_{sym}}\sum_{n=1}^{N_{sym}} |y_n|^2 |y_{n-1}|^2}$$

which is very similar to (2.114) of Section 2.3.5. The kurtosis of the signal is $k_a = 1$, and the kurtosis of the noise is chosen to be $k_\xi = 1.5$ as a compromise between the kurtosis of complex AWGN (for which $k_\xi = 2$) and the kurtosis of a CCI signal (for which $k_\xi = 1$), as done for the M_2M_4 SIR estimator in Section 3.7.1. The SVR SIR estimator may then be expressed using (3.31) (taking the positive root) as

$$\hat{\rho}_{\mathrm{SVR}} = \beta - 1 + \sqrt{\beta(\beta - 1.5)}. \tag{3.32}$$

This is precisely the result given by Brandão *et al* in [56]. The idea to use $k_\xi = 1.5$ as the kurtosis of the impairment process was originally suggested in [56].

3.7.3 The Absolute Difference of Absolute I and Absolute Q (ADIQ) Estimator

This method was first introduced by Yoshida *et al* in [57] as a means to measure link quality degradation as a result of multipath-induced ISI in systems that employ $\pi/4$-DQPSK modulation. In [58] it was shown that the same method described in [57] could be used to provide a measure of both multipath delay-spread and CCI power in systems which employ any type of "QPSK-like" signals. The method is quite simple in that the estimator only need be able to compute the absolute value and difference of samples. No name was attributed to this method by the original authors, so the descriptive name, the *Absolute Difference of Absolute I and Absolute Q (ADIQ) Estimator* is assigned here for ease of reference.

The ADIQ method may be applied to any "QPSK-like" signal; that is, it may be applied to any signal whose complex, baseband signal constellation comprises four symbols forming a square centred on the origin. The differentially-detected $\pi/4$-DQPSK constellation complies with this requirement.

Unfortunately, the ADIQ expression given in [58, 57] yields estimates that don't have any particular meaning in an absolute sense, in contrast to all of the other SIR estimators considered in this study. For example, in the complex AWGN channel, all of the SIR estimators considered in this study except for the ADIQ estimator as given in [58, 57] yield estimates that correspond to the SNR. As a consequence, the ADIQ estimator as given in [58, 57] can only be used in applications which require a *relative* measure of channel quality. However, it is shown below how the original form of the ADIQ estimator may be transformed into a modified estimator which yields the SNR in an AWGN channel. The original ADIQ estimator of [58, 57] is denoted by $\hat{\varrho}_{\text{ADIQ}}$, whereas the modified ADIQ estimator presented here is denoted by $\hat{\rho}_{\text{ADIQ}}$.

The ADIQ estimator operates on the post-DD samples, x_n, given by (3.25) which may be denoted explicitly in terms of real and imaginary parts as $x_n = x_{I_n} + jx_{Q_n}$ so that the in-service ADIQ SIR estimator as given by [58, 57] may be written in

time-average form as

$$\hat{\varrho}_{\text{ADIQ}} = \frac{1}{N_{\text{sym}}} \sum_{n=1}^{N_{sym}} \frac{||x_{I_n}| - |x_{Q_n}||}{\sqrt{x_{I_n}^2 + x_{Q_n}^2}} \tag{3.33}$$

where decreasing $\hat{\varrho}_{\text{ADIQ}}$ corresponds to increasing signal quality. By inspection of (3.33), it may be seen that the magnitude of all symbols is normalized. This "instantaneous automatic gain control" (AGC) was incorporated by the original authors (see [58]) to eliminate fluctuations in the received signal. No detailed explanation is offered in [58], but it is easy to see that if the AGC were not employed, the amplitude of the received signal would impair the function of the ADIQ estimator. For example, for a desired signal with a given nominal amplitude corrupted by AWGN with a certain variance, the numerator of (3.33) varies depending on the amplitude of the noisy received signal (which, in turn, yields correspondingly different estimates of $\hat{\varrho}_{\text{ADIQ}}$); however, any SNR (or SIR) estimate should be independent of the amplitude of the sum of the desired signal and noise, thus the inclusion of the instantaneous AGC.

It was discovered empirically by inspection of the simulated SNR estimates produced by (3.33) in the AWGN channel of Case 1 that the true SNR is proportional to the inverse square of $\hat{\varrho}_{\text{ADIQ}}$. Specifically, the modified ADIQ SIR estimator which yields correct values of the SNR may be expressed as

$$\hat{\rho}_{\text{ADIQ}} = \frac{1}{\sqrt{\frac{2}{\pi}}\hat{\varrho}_{\text{ADIQ}}^2} \tag{3.34}$$

where $\hat{\varrho}_{\text{ADIQ}}^2$ is given by (3.33). The factor, $\sqrt{2/\pi}$, is the same factor that arises in the context of the relationship between the sample standard deviation and the more robust mean absolute deviation [122]. Specifically, consider a set of samples represented by x_k, $k \in \{1, 2, \ldots, K\}$. It is shown in [122] that if the unbiased sample mean of the samples is given by $\bar{x}$ as

$$\bar{x} = \frac{1}{K} \sum_{k=1}^{K} x_k,$$

and if the unbiased sample standard deviation of the samples is given by

$$\sigma = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(x_k - \bar{x})^2},$$

then the *mean absolute deviation* (MAD), given by

$$d = \frac{1}{K}\sum_{k=1}^{K}|x_k - \bar{x}|,$$

is related to the sample standard deviation, σ, by

$$d \to \sqrt{\frac{2}{\pi}}\sigma$$

in the limit that $K \to \infty$. It seems reasonable, but not obvious, that this $\sqrt{2/\pi}$ factor should arise in the context of ADIQ estimation. Proving the empirical observation represented by (3.34) is difficult due to the nested absolute value operations in (3.33) so a proof is not provided here.

As discussed in Section 3.7.7, the absolute value operations on x_{I_n} and x_{Q_n} in (3.34) and (3.33) imply an RxDA type of estimator. It is possible, therefore, to create a TxDA form of the ADIQ estimator by replacing $|x_{I_n}|$ and $|x_{Q_n}|$ with (3.55) and (3.56) (see Section 3.7.7) and by using the true transmitted differential phases, Δ_n, instead of the estimated differential phases, $\hat{\Delta}_n$, as discussed in Section 3.7.6. Though this modification would yield an estimator with better performance at low SIR, the TxDA form of the ADIQ estimator is not studied here since the originators [58, 57] consider only the RxDA form. Any mention of the ADIQ estimator here implies the RxDA form of (3.34) and (3.33) even though the "RxDA" label is not explicitly applied.

3.7.4 The Maximum-Likelihood Estimator for Post-MF (Pre-DD) Samples

In Section 2.6 the simulated performances of the ML TxDA and SNV TxDA estimators (both of which are ML estimators) were found to come closest to the theoretical

CRB. It would seem reasonable to try a similar ML approach to derive an ML SIR estimator for the channel described in Section 3.6.1. Unfortunately, it is a difficult problem to apply the formal ML method to the signal-plus-*impairment* samples at the input to the MF since the overall impairment process consists of several separate impairments, each having different statistics. Also, it is difficult to build the fast, frequency-selective fading channel into the ML derivation. Instead, the *ad hoc* application of the complex ML SNR estimator, presented in Section 2.3.2, to the general system model of Section 3.6.1 is considered here.

The reduced-bias, complex ML TxDA SNR expression is given by (2.72) of Section 2.3.2 as

$$\hat{\rho}'_{\text{ML TxDA,complex}} = \frac{N_{\text{ss}}^2 \left[\frac{1}{K}\sum_{k=0}^{K-1} \text{Re}\{r_k^* m_k\}\right]^2}{\frac{1}{K-1}\sum_{k=0}^{K-1}|r_k|^2 - \frac{N_{\text{ss}}}{K(K-1)}\left[\sum_{k=0}^{K-1}\text{Re}\{r_k^* m_k\}\right]^2}, \tag{3.35}$$

which takes as input the sequence of $K = N_{\text{sym}}N_{\text{ss}}$ corrupted samples at the input to the MF in the model of Section 2.2. In terms of the model of Section 3.6.1, the pre-MF sequence is given by (3.20) as

$$r_k = q_k + \tilde{q}_k + \sqrt{N}z_k,$$

where $k \in \{0, 1, \ldots, K-1\}$ and $K = (N_{\text{sym}}+1)N_{\text{ss}}$ due to differential encoding. The ML TxDA SNR estimator must form a re-modulated version of the transmitted signal from known transmitted symbols. The "clean" version of the transmitted signal to use here is given by (3.7) as

$$m_k = b_k \otimes h_k,$$

where b_k is given by (3.6) as

$$b_k = \sum_n a_n \delta_{k,nN_{\text{ss}}},$$

and a_n represents the sequence of $\pi/4$-DQPSK-encoded source symbols. As described in Section 2.3.2, the RxDA version of the ML estimator is obtained by replacing the known pulse-shaped samples, m_k, with estimated pulse-shaped samples, $m_k^{(i)}$, formed

from receiver decisions. The reduced-bias, complex ML RxDA estimator is given by (2.73) of Section 2.3.2 as

$$\hat{\rho}'_{\text{ML RxDA,complex}} = \frac{N_{ss}^2 \left[\frac{1}{K}\sum_{k=0}^{K-1} \text{Re}\left\{r_k^* m_k^{(\hat{i})}\right\}\right]^2}{\frac{1}{K-1}\sum_{k=0}^{K-1}|r_k|^2 - \frac{N_{ss}}{K(K-1)}\left[\sum_{k=0}^{K-1} \text{Re}\left\{r_k^* m_k^{(\hat{i})}\right\}\right]^2}. \tag{3.36}$$

Recall that the parameter, $\hat{i}$, denotes the $\hat{i}^{\text{th}}$ sequence selected by the receiver out of $M^{N_{sym}}$ possible transmitted sequences (for $\pi/4$-DQPSK modulation, $M = 4$).

Apart from the fact that the complex ML SNR estimators of Section 2.3.2 are derived specifically for the AWGN channel, there are two other issues of concern that make the ML SNR estimators unsuitable for the system model of Section 3.6.1. The first issue is the effect of receiver errors in the case of ML RxDA estimation, and the second issue is the effect of random phase fluctuations caused by Rayleigh fading.

First consider the effect of receiver errors on the ML RxDA estimator. The ML TxDA SNR estimator given by (3.35) was simulated in the Case 1 channel described in Section 3.6.2 and its performance was found to approach the CRB, as expected (see Figure 3.26). The ML RxDA SNR estimator given by (3.36) was also simulated in the Case 1 channel and was found to have a performance similar to that of its TxDA counterpart at high SNR, but its performance at low SNR was remarkably poor due to the catastrophic effect of receiver errors (see Figure 3.26 of Section 3.9.1).

To illustrate the mechanism behind the catastrophic failure of this estimator at low SNR, consider Table 3.3 (constructed using Table F.3 of Appendix F) which shows $\pi/4$-DQPSK-encoded source symbols formed from known bit-pairs at the transmitter (Tx), and re-modulated symbols formed from estimated bit-pairs at the receiver (Rx). It can be seen that the symbols generated by the Rx are different from the symbols generated by the Tx after the single bit error at index $n = 2$, so that the correlation between r_k and the estimated samples, $m_k^{(\hat{i})}$, is destroyed. The *differential* phases after the symbol error are not affected, but the *absolute* phases are. The only way the encoded symbol sequences of the Tx and Rx could become equal again is if a

Symbol Index, n	0	1	2	3	4	$\cdots$
Information bit-pairs (Tx)	—	00	10	01	11	$\cdots$
$\pi/4$-DQPSK-encoded source symbols (Tx)	$\pi/4$	$\pi/2$	$-\pi/4$	$\pi/2$	$\pi/4$	$\cdots$
Information bit-pairs (Rx)	—	00	00	01	11	$\cdots$
$\pi/4$-DQPSK-encoded source symbols (Rx)	$\pi/4$	$\pi/2$	$3\pi/4$	$-\pi/2$	$-3\pi/4$	$\cdots$

Table 3.3: The effect of a single bit error (at $n = 2$) on differentially-encoded symbols.

future, fortuitous receiver error were to occur such that the erroneous differential phase makes the absolute phases of the Tx and Rx symbols the same. Of course, the next error to occur after that would spoil the correlation between the absolute phases of the Tx and Rx symbols once again.

Next, consider the random phase fluctuations caused by Rayleigh fading. The phase fluctuations (see Figure 3.3(b)) of the pre-DD received samples deleteriously affect the correlation of the received signal with the clean, local copy of the transmitted waveform. This is a problem that affects both the ML RxDA and ML TxDA SNR estimators. To overcome this problem, the receiver must be able to track the random phase fluctuations which implies additional receiver complexity[9].

Both of these issues indicate that it would seem advantageous to try the ML method on the *post-DD* samples rather than on the *pre-DD* samples. For the first issue, post-DD estimation alleviates the catastrophic effect of errors on RxDA estimation since the post-DD symbols contain only *differential* phase information, not *absolute* phase information. For the second issue, since the Rayleigh fading-induced random phase changes are relatively small over the duration of a symbol period, the differentially-detected symbols are relatively unaffected by the random phase fluctuations. The next four sections discuss possible post-DD implementations of SIR

[9]Note that the M_2M_4 and the SVR SIR estimators operate on the same pre-DD samples, r_k, but since these estimators are based on even moments (the second and fourth moments) of the received signal, the phase information disappears so that the Rayleigh fading-induced random phase fluctuations do not affect these estimators the same way that the affect correlation-based, data-aided estimators (such as the ML estimator).

estimators which are "ML-like" in structure.

3.7.5 The Maximum-Likelihood Estimator for Post-DD Samples

An ML SIR estimator was derived by Chennakeshu and Saulnier in [28] based on the pdf of the phase of differentially-detected samples assuming an AWGN channel. The resulting ML estimator for AWGN is applied *ad hoc* to wireless channels in which exist the additional impairments of fading and CCI. This SIR estimator was developed for use in a timing and frequency offset estimation algorithm described in [28]. This estimator is explicitly referred to as a signal-to-impairment ratio estimator in [28], but in order to distinguish it from the other SIR estimators under study, it will be referred to here as the *DDML* SIR estimator (the "DD" in DDML indicates that this is an ML expression for post-DD samples). The derivation of the DDML SIR estimator is sketched below based on the development of [28].

Since this ML method is based on the phase of the post-DD samples, (3.25) must be re-written in polar form as

$$x_n = y_n y_{n-1}^* = A_n e^{j\phi_n},\ n \in \{1, 2, \ldots, N_{\text{sym}}\} \tag{3.37}$$

where A_n represents the envelope and ϕ_n represents the phase of the impaired samples at the output of the DD at time index n. Let Δ_n represent the true transmitted phase. The pdf of the phases of the post-DD samples conditioned on Δ_n and the "instantaneous" SNR, ρ_n, may then be written using the notation adopted here to obtain an expression similar to equation (A2.3) of [28] (based on equation (4.2.103) of [65]) as

$$\begin{aligned} f(\phi_n|\Delta_n, \rho_n) &= \frac{1}{2\pi} e^{-\frac{\rho_n}{2}} \left[1 + \sqrt{2\pi\rho_n}\cos(\Delta_n - \phi_n) e^{\frac{\rho_n}{2}\cos^2(\Delta_n - \phi_n)}\right. \\ &\qquad \left. \times \left(1 - \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\rho_n}{2}}\cos(\Delta_n - \phi_n)\right)\right)\right]. \end{aligned} \tag{3.38}$$

Following [28], the pdf of ϕ_n is approximated by making the assumptions that $\rho_n \gg 1$

and $|\Delta_n - \phi_n| < \pi/2$, so that (3.38) simplifies to

$$f(\phi_n|\Delta_n,\rho_n) = \sqrt{\frac{\rho_n}{2\pi}}\cos(\Delta_n-\phi_n)e^{-\frac{\rho_n}{2}\sin^2(\Delta_n-\phi_n)}. \tag{3.39}$$

Assuming the N_{sym} differential phases are independent, the pdf of the sequence of received differential phases may be written simply as

$$f(\boldsymbol{\phi}|\Delta_n,\rho_n) = \prod_{n=1}^{N_{\text{sym}}} f(\phi_n|\Delta_n,\rho_n), \tag{3.40}$$

where $\boldsymbol{\phi} \in \{\phi_1,\phi_2,\ldots,\phi_{N_{\text{sym}}}\}$. Taking the logarithm of (3.40) and differentiating with respect to ρ_n, one obtains

$$\begin{aligned}\frac{\partial}{\partial\rho_n}\ln f(\boldsymbol{\phi}|\Delta_n,\rho_n) &= \frac{\partial}{\partial\rho_n}\sum_{n=1}^{N_{\text{sym}}}\left[\frac{1}{2}\ln\rho_n - \frac{1}{2}\ln(2\pi) + \ln\cos(\Delta_n-\phi_n)\right.\\ &\qquad\left. - \frac{\rho_n}{2}\sin^2(\Delta_n-\phi_n)\right]\\ &= \sum_{n=1}^{N_{\text{sym}}}\left[\frac{1}{2\rho_n} - \frac{1}{2}\sin^2(\Delta_n-\phi_n)\right].\end{aligned} \tag{3.41}$$

Setting (3.41) to zero and solving for ρ_n, one obtains an expression for the DDML TxDA SIR estimator as

$$\hat{\rho}_{\text{DDML,TxDA}} = \frac{1}{\frac{1}{N_{\text{sym}}}\sum_{n=1}^{N_{\text{sym}}}\sin^2(\Delta_n-\phi_n)}. \tag{3.42}$$

This is a TxDA SIR estimator since the transmitted differential phase, Δ_n, is assumed known to the receiver. In the case that the received differential phase is estimated by the receiver, the DDML RxDA estimator may be expressed similarly as

$$\hat{\rho}_{\text{DDML,RxDA}} = \frac{1}{\frac{1}{N_{\text{sym}}}\sum_{n=1}^{N_{\text{sym}}}\sin^2(\hat{\Delta}_n-\phi_n)}, \tag{3.43}$$

where $\hat{\Delta}_n$ represents the estimated differential phase.

3.7.6 The Pseudo-Maximum-Likelihood (PML) Estimator

In Section 3.7.5, the derivation of the DDML SIR estimator was sketched based on the phases of the post-DD samples—the amplitude information of the samples was

ignored. It is worthwhile considering an ML structure for post-DD samples based on a Cartesian approach similar to that discussed in Section 3.7.4 for pre-DD samples. It is a difficult problem to apply the formal ML method to the signal-plus-noise samples at the output of the DD due to the non-linear differential-detection operation (though Chennakeshu and Saulnier applied the ML method to the differential phases, the resulting pdf had to be approximated by making simplifying assumptions in order to arrive at an ML expression). Instead, consider a "pseudo-ML" estimator formulated here based on the general structure of the SNV SNR estimator (a type of ML estimator) for complex AWGN channels. The SNV TxDA SNR estimator is given by (2.84) of Section 2.3.3 as

$$\hat{\rho}'_{\text{SNV TxDA,complex}} = \frac{\left[\frac{1}{N_{sym}}\sum_{n=0}^{N_{sym}-1}\text{Re}\{y_n^* a_n\}\right]^2}{\frac{1}{N_{sym}-1}\sum_{n=0}^{N_{sym}-1}|y_n|^2 - \frac{1}{N_{sym}(N_{sym}-1)}\left[\sum_{n=0}^{N_{sym}-1}\text{Re}\{y_n^* a_n\}\right]^2}. \quad (3.44)$$

It can be seen that the general approach of the complex SNV TxDA SNR estimator is to correlate the complex-conjugate of the symbol-spaced samples at the output of the MF, y_n^*, with the the known transmitted symbols, a_n, to estimate the signal power[10] (the numerator of $\hat{\rho}'_{\text{SNV TxDA,complex}}$). The noise power is estimated by subtracting the estimated signal power from an estimate of the total signal-plus-noise power. The SNR is then found trivially as the ratio of the estimated signal and noise powers. This general approach of the SNV SNR estimator is applied below to the output of the DD as the basis of the PML SIR estimator.

In order to obtain a PML TxDA estimate of the signal power, the known transmitted sequence of source symbols (or simply the sequence of phases) is required. The symbol-spaced samples of the DD output may be expressed in terms of (3.22) and (3.25) as

$$x_n = y_n y_{n-1}^* = \zeta_{d_n}\zeta_{d_{n-1}}^* + \zeta_{d_n}\xi_{n-1}^* + \xi_n\zeta_{d_{n-1}}^* + \xi_n\xi_{n-1}^*. \quad (3.45)$$

[10]Equivalently, the signal power may be found as the correlation of y_n with a_n^* since $\text{Re}\{y_n^* a_n\} = \text{Re}\{a_n^* y_n\}$.

The phase difference between ζ_{d_n} and $\zeta_{d_{n-1}}$ represents the transmitted differential phase at time index n (note that there are four possible differential phases—see Table F.3 of Appendix F). Let Δ_n represent the transmitted differential phase so that

$$\zeta_{d_n}\zeta^*_{d_{n-1}} = S\,e^{j\Delta_n}. \tag{3.46}$$

A PML TxDA estimate of the signal power is formed by the correlation of the received post-DD symbols with the complex conjugate of the transmitted differential symbols. Taking advantage of the independence of the signal and noise samples, the PML TxDA estimate of the signal may be written as

$$\begin{aligned} E\{x_n e^{-j\Delta_n}\} &= S\,E\{e^{j(\Delta_n-\Delta_n)}\} + E\{\zeta_{d_n}e^{-j\Delta_n}\}E\{\xi^*_{n-1}\} \\ &\quad + E\{\xi_n\}E\{\zeta^*_{d_{n-1}}e^{-j\Delta_n}\} + E\{\xi_n\xi_{n-1}\}E\{e^{-j\Delta_n}\} \\ &= S \end{aligned} \tag{3.47}$$

where $E\{\xi^*_{n-1}\} = E\{\xi_n\} = 0$ using (3.24a), and $E\{\xi_n\xi_{n-1}\} = \Xi\,\delta_{n,n-1} = 0$ using (3.24b).

It is illustrative to consider a graphical interpretation of the samples represented by $x_n e^{-j\Delta_n}$. In Section 3.6.2, plots are provided showing examples of the appearance of various signals in the reference "clear-sky" channel and in the five mobile radio channels. In particular, Figures 3.11(d), 3.12(d), 3.14(d), 3.15(d), 3.16(d), and 3.18(d) show examples of the "de-rotated" signal, $x_n e^{-j\Delta_n}$, under various conditions. The signal is "de-rotated" in the sense that the post-DD symbols are phase-rotated to the real, positive axis, essentially stripping the modulation from the symbols. Figures 3.12(d), 3.15(d), and 3.16(d) show that the de-rotated post-DD symbols in Cases 1, 3, and 4 are reasonably symmetric, whereas the amplitude variations caused by the fast, Rayleigh fading present in Cases 2 and 5 spreads the de-rotated symbols along the real axis, as is evident in Figures 3.14(d) and 3.18(d).

In a similar manner as for the complex SNV SNR estimator, the total power is

obtained as

$$\begin{aligned} E\{y_n y_n^*\} &= S + E\{\zeta_{d_n}\}E\{\xi_n^*\} + E\{\xi_n\}E\{\zeta_{d_n}^*\} + E\{\xi_n\xi_n^*\} \\ &= S + \Xi, \end{aligned}$$

again making use of (3.23) and (3.24), so that an expression for the impairment power is

$$\Xi = E\{y_n y_n^*\} - E\{x_n e^{-j\Delta_n}\}. \tag{3.48}$$

In practice, expected values are estimated by time averages so that the PML TxDA SIR estimator may be expressed as

$$\hat{\rho}_{\text{PML,TxDA}} = \frac{\frac{1}{N_{sym}}\sum_{n=1}^{N_{sym}} \text{Re}\{x_n e^{-j\Delta_n}\}}{\frac{1}{N_{sym}}\sum_{n=1}^{N_{sym}} \left(y_n y_n^* - \text{Re}\{x_n e^{-j\Delta_n}\}\right)}, \tag{3.49}$$

where the real part of $x_n e^{-j\Delta_n}$ is taken to ensure that the SIR estimate is real. Note that the summation starts at $n = 1$ because of the differential detection operation.

The PML RxDA SIR estimator is identical in form to (3.49), except that Δ_n is replaced by $\hat{\Delta}_n$, which represents the *estimated* transmitted differential phases generated by the receiver. The in-service, PML RxDA SIR estimator is expressed explicitly as

$$\hat{\rho}_{\text{PML,RxDA}} = \frac{\frac{1}{N_{sym}}\sum_{n=1}^{N_{sym}} \text{Re}\{x_n e^{-j\hat{\Delta}_n}\}}{\frac{1}{N_{sym}}\sum_{n=1}^{N_{sym}} \left(y_n y_n^* - \text{Re}\{x_n e^{-j\hat{\Delta}_n}\}\right)}. \tag{3.50}$$

3.7.7 The Signal-to-Impairment Variance (SIV) Estimator

This method is essentially an *ad hoc* application of the *real* form of the SNV SNR estimator (see Section 2.3.3) to each of the in-phase and quadrature components of the complex, baseband signal. The de-rotated post-DD symbols are not required for this estimation method.

Using (3.45) and (3.46), the DD output can be expressed as

$$x_n = y_n y_{n-1}^* = S\, e^{j\Delta_n} + \zeta_{d_n}\xi_{n-1}^* + \xi_n\zeta_{d_{n-1}}^* + \xi_n\xi_{n-1}^*. \tag{3.51}$$

For notational simplicity, let

$$\zeta_{d_n} = a_I + ja_Q \tag{3.52a}$$

$$\zeta_{d_{n-1}} = a'_I + ja'_Q \tag{3.52b}$$

$$\xi_n = b_I + jb_Q \tag{3.52c}$$

$$\xi_{n-1} = b'_I + jb'_Q. \tag{3.52d}$$

Equation (3.51) may then be expanded to show real and imaginary parts explicitly as

$$\begin{aligned} x_n &= x_{I_n} + jx_{Q_n} \\ &= S\cos\Delta_n + a_I b'_I + a_Q b'_Q + a'_I b_I + a'_Q b_Q + b_I b'_I + b_Q b'_Q \\ &\quad + j\left(S\sin\Delta_n + a_Q b'_I - a_I b'_Q + a'_I b_Q - a'_Q b_I + b_Q b'_I - b_I b'_Q\right). \end{aligned} \tag{3.53}$$

The real form of the SNV RxDA estimator is given by (2.78) of Section 2.3.3 in terms of expected values as

$$\hat{\rho}_{\text{SNV RxDA,real}} = \frac{E\{|y_n|\}^2}{\text{Var}\{|y_n|\}}. \tag{3.54}$$

Following this general form, the SIR estimator is to be written as some function of $E\{|x_{I_n}|\}$ and $E\{|x_{Q_n}|\}$, and $\text{Var}\{|x_{I_n}|\}$ and $\text{Var}\{|x_{Q_n}|\}$. To facilitate the derivations that follow, consider the replacement of $|x_{I_n}|$ and $|x_{Q_n}|$ with the equivalent expressions

$$|x_{I_n}| = x_{I_n}\sqrt{2}\,\text{Re}\{e^{j\hat{\Delta}_n}\} = \sqrt{2}\,x_{I_n}\cos\hat{\Delta}_n \tag{3.55}$$

and

$$|x_{Q_n}| = x_{Q_n}\sqrt{2}\,\text{Im}\{e^{j\hat{\Delta}_n}\} = \sqrt{2}\,x_{Q_n}\sin\hat{\Delta}_n, \tag{3.56}$$

respectively, where $\hat{\Delta}_n$ represents an *estimate* of the phase difference between symbols transmitted at time indices n and $n-1$, as discussed in Section 3.7.6, and $\text{Im}\{\cdot\}$ denotes the imaginary part of a complex quantity. Using (3.55) and (3.53), the expected value of $|x_{I_n}|$ may be evaluated as

$$
\begin{aligned}
E\{|x_{I_n}|\}/\sqrt{2} &= E\{x_{I_n}\cos\hat{\Delta}_n\} \\
&= S\,E\{\cos\Delta_n\cos\hat{\Delta}_n\} + E\{a_I\cos\hat{\Delta}_n\}E\{b'_I\} + E\{a_Q\cos\hat{\Delta}_n\}E\{b'_Q\} \\
&\quad + E\{a'_I\cos\hat{\Delta}_n\}E\{b_I\} + E\{a'_Q\cos\hat{\Delta}_n\}E\{b_Q\} \\
&\quad + E\{b_Ib'_I\}E\{\cos\hat{\Delta}_n\} + E\{b_Qb'_Q\}E\{\cos\hat{\Delta}_n\}. \qquad (3.57)
\end{aligned}
$$

Using (3.24a) and (3.24b), one has

$$
E\{b_I\} = E\{b'_I\} = E\{b_Q\} = E\{b'_Q\} = 0 \qquad (3.58a)
$$

$$
E\{b_Ib'_I\} = E\{b_Qb'_Q\} = \Xi\delta_{n,n-1} = 0. \qquad (3.58b)
$$

Using the fact that $\Delta_n \in \{\pm\frac{\pi}{4}, \pm\frac{3\pi}{4}\}$, with each of the four possible phases occurring with equal probability, one obtains

$$
\begin{aligned}
E\{\cos\Delta_n\cos\hat{\Delta}_n\} &= E\{\cos(\Delta_n-\hat{\Delta}_n)+\cos(\Delta_n+\hat{\Delta}_n)\}/2 \\
&\approx E\{1+\cos 2\Delta_n\}/2 \\
&= \tfrac{1}{2}\left(1+\tfrac{1}{4}\cos\left(\tfrac{\pi}{2}\right)+\tfrac{1}{4}\cos\left(-\tfrac{\pi}{2}\right)+\tfrac{1}{4}\cos\left(\tfrac{3\pi}{2}\right)+\tfrac{1}{4}\cos\left(-\tfrac{3\pi}{2}\right)\right) \\
&= 1/2, \qquad (3.59)
\end{aligned}
$$

where the approximation is due to the assumption that $\hat{\Delta}_n = \Delta_n$ (meaning the receiver makes perfect decisions). Substituting the expressions given by (3.58) and (3.59) into (3.57) yields

$$
E\{|x_{I_n}|\} = S/\sqrt{2}. \qquad (3.60)
$$

Similarly, using (3.56) and (3.53) the expected value of $|x_{Q_n}|$ may be evaluated as

$$
\begin{aligned}
E\{|x_{Q_n}|\}/\sqrt{2} &= E\{x_{Q_n}\sin\hat{\Delta}_n\} \\
&= S\,E\{\sin\Delta_n\sin\hat{\Delta}_n\} + E\{a_Q\sin\hat{\Delta}_n\}E\{b'_I\} - E\{a_I\sin\hat{\Delta}_n\}E\{b'_Q\} \\
&\quad + E\{a'_I\sin\hat{\Delta}_n\}E\{b_Q\} - E\{a'_Q\sin\hat{\Delta}_n\}E\{b_I\} \\
&\quad + E\{b_Q\}E\{b'_I\}E\{\sin\hat{\Delta}_n\} - E\{b_I\}E\{b'_Q\}E\{\sin\hat{\Delta}_n\}. \qquad (3.61)
\end{aligned}
$$

Note that

$$
E\{b_Qb'_I\} = E\{b_Q\}E\{b'_I\} \qquad (3.62a)
$$

$$E\{b_I b'_Q\} = E\{b_I\}E\{b'_Q\} \tag{3.62b}$$

because the in-phase and quadrature impairment processes are assumed independent. By similar arguments used to evaluate $E\{\cos\Delta_n \cos\hat{\Delta}_n\}$ in (3.59), one has

$$\begin{aligned} E\{\sin\Delta_n \sin\hat{\Delta}_n\} &= E\{\cos(\Delta_n - \hat{\Delta}_n) - \cos(\Delta_n + \hat{\Delta}_n)\}/2 \\ &\approx E\{1 - \cos 2\Delta_n\}/2 \\ &= \tfrac{1}{2}\left(1 - \tfrac{1}{4}\cos\left(\tfrac{\pi}{2}\right) - \tfrac{1}{4}\cos\left(-\tfrac{\pi}{2}\right) - \tfrac{1}{4}\cos\left(\tfrac{3\pi}{2}\right) - \tfrac{1}{4}\cos\left(-\tfrac{3\pi}{2}\right)\right) \\ &= 1/2. \end{aligned} \tag{3.63}$$

Substitution of the expressions given by (3.58a), (3.62) and (3.63) into (3.61) yields

$$E\{|x_{Q_n}|\} = S/\sqrt{2}. \tag{3.64}$$

Finally, using (3.60) and (3.64), an SIV RxDA estimate of the signal power can be expressed as

$$\hat{S}_{\text{SIV,RxDA}} = \frac{1}{\sqrt{2}}\left[E\{|x_{I_n}|\} + E\{|x_{Q_n}|\}\right]. \tag{3.65}$$

The next step is to derive the SIV RxDA estimate of the impairment power. The variances of $|x_{I_n}|$ and $|x_{Q_n}|$ are

$$\text{Var}\{x_{I_n}\} = E\{x_{I_n}^2\} - E\{|x_{I_n}|\}^2 \tag{3.66a}$$

$$\text{Var}\{x_{Q_n}\} = E\{x_{Q_n}^2\} - E\{|x_{Q_n}|\}^2, \tag{3.66b}$$

so that $E\{x_{I_n}^2\}$ and $E\{x_{Q_n}^2\}$ need to be computed. Let $\alpha = S\cos\hat{\Delta}_n$ so that $x_{I_n}^2$ may be written as

$$\begin{aligned} x_{I_n}^2 &= (\alpha + a_I b'_I + a_Q b'_Q + a'_I b_I + a'_Q b_Q + b_I b'_I + b_Q b'_Q)^2 \\ &= \alpha^2 + (a_I b'_I)^2 + (a_Q b'_Q)^2 + (a'_I b_I)^2 + (a'_Q b_Q)^2 + (b_I b'_I)^2 + (b_Q b'_Q)^2 \\ &\quad + 2\left(\alpha a_I b'_I + \alpha a_Q b'_Q + \alpha a'_I b_I + \alpha a'_Q b_Q + \alpha b_I b'_I + \alpha b_Q b'_Q\right. \\ &\quad + a_I a_Q b'_I b'_Q + a_I a'_I b_I b'_I + a_I a'_Q b_Q b'_I + a_I b_I (b'_I)^2 + a_I b'_I b_Q b'_Q \\ &\quad \left. + a'_I a_Q b_I b'_Q + a_Q a'_Q b_Q b'_Q + a_Q b_I b'_I b'_Q + a_Q b_Q (b'_Q)^2\right) \end{aligned}$$

$$
\begin{aligned}
&+ a'_I a'_Q b_I b_Q + a'_I (b_I)^2 b'_I + a'_I b_I b_Q b'_Q \\
&+ a'_Q b_I b'_I b_Q + a'_Q (b_Q)^2 b'_Q \\
&+ b_I b'_I b_Q b'_Q \big) . \qquad (3.67)
\end{aligned}
$$

The expected values of many of the terms on the right-hand side of (3.67) vanish. Since the signal and noise samples have zero mean and are uncorrelated, then one has

$$
\begin{aligned}
E\{\alpha a_I\}E\{b'_I\} &= E\{\alpha a_Q\}E\{b'_Q\} = E\{\alpha a'_I\}E\{b_I\} = E\{\alpha a'_Q\}E\{b_Q\} = 0 \\
E\{a_I\}E\{b_I(b'_I)^2\} &= E\{a_I\}E\{b'_I b_Q b'_Q\} = E\{a_Q\}E\{b_I b'_I b'_Q\} = E\{a_Q\}E\{b_Q(b'_Q)^2\} = 0 \\
E\{a'_I\}E\{(b_I)^2 b'_I\} &= E\{a'_I\}E\{b_I b_Q b'_Q\} = E\{a'_Q\}E\{b_I b'_I b_Q\} = E\{a'_Q\}E\{(b_Q)^2 b'_Q\} = 0.
\end{aligned}
$$

Also, the in-phase and quadrature signal samples are assumed uncorrelated, as are the in-phase and quadrature noise samples, so that

$$
\begin{aligned}
E\{a_I\}E\{a_Q\}E\{b'_I\}E\{b'_Q\} &= E\{a_I\}E\{a'_Q\}E\{b_Q\}E\{b'_I\} = 0 \\
E\{a'_I\}E\{a_Q\}E\{b_I\}E\{b'_Q\} &= E\{a'_I\}E\{a'_Q\}E\{b_I\}E\{b_Q\} = 0.
\end{aligned}
$$

Since the cross-correlation of the signal and noise sequences is zero, and since the autocorrelations of the signal and noise sequences are also zero for any nonzero time delay (see (3.23b) and (3.24b)), then

$$
\begin{aligned}
E\{\alpha\}E\{b_I b'_I\} &= E\{\alpha\}E\{b_Q b'_Q\} = 0 \\
E\{a_I a'_I\}E\{b_I b'_I\} &= E\{a_Q a'_Q\}E\{b_Q b'_Q\} = 0 \\
E\{b_I b'_I\}E\{b_Q b'_Q\} &= 0.
\end{aligned}
$$

Assume the statistics of Δ_n and $\hat{\Delta}_n$ are the same so that

$$
\begin{aligned}
E\{\alpha^2\} &= E\{S^2 \cos^2 \hat{\Delta}_n\} \\
&= \frac{S^2}{2}\left(1 + E\{\cos 2\hat{\Delta}_n\}\right) \\
&= S^2/2,
\end{aligned}
$$

using the fact that $E\{\cos 2\Delta_n\} = 0$, which is apparent from either (3.59) or (3.63). Again using the independence of the signal and noise samples, one has

$$E\{a_I^2\}E\{(b_I')^2\} = E\{a_Q^2\}E\{(b_Q')^2\} = E\{(a_I')^2\}E\{b_I^2\} = E\{(a_Q')^2\}E\{b_Q^2\} = \frac{S\,\Xi}{4}.$$

Finally, Price's Theorem [66] may be used to evaluate the squared autocorrelation of Gaussian noise giving

$$\begin{aligned} E\{b_I^2(b_I')^2\} &= 2\,E^2\{b_I b_I'\} + E^2\{b_I^2\} \\ &= 2\,\frac{\Xi^2}{4}\delta_{n,n-1} + \frac{\Xi^2}{4} \\ &= \frac{\Xi^2}{4}. \end{aligned}$$

Similarly, for the quadrature channel, one obtains

$$E\{b_Q^2(b_Q')^2\} = \frac{\Xi^2}{4}.$$

The expected value of (3.67) thus simplifies to

$$E\{x_{I_n}^2\} = \frac{S^2}{2} + 4\frac{S\,\Xi}{4} + 2\,\frac{\Xi^2}{4} = \frac{S^2}{2} + S\,\Xi + \frac{\Xi^2}{2}, \tag{3.68}$$

and using (3.60) and (3.68), (3.66a) becomes

$$\mathrm{Var}\{|x_{I_n}|\} = \frac{S^2}{2} + S\,\Xi + \frac{\Xi^2}{2} - \frac{S^2}{2} = S\,\Xi + \frac{\Xi^2}{2} = \sigma_{I,\mathrm{SIV}}^2. \tag{3.69}$$

Now let $\beta = S\,\sin\hat{\Delta}_n$, so that $x_{Q_n}^2$ may be written as

$$\begin{aligned} x_{Q_n}^2 &= (\beta + a_Q b_I' - a_I b_Q' + a_I' b_Q - a_Q' b_I + b_Q b_I' - b_I b_Q')^2 \\ &= \beta^2 + (a_Q b_I')^2 + (a_I b_Q')^2 + (a_I' b_Q)^2 + (a_Q' b_I)^2 + (b_Q b_I')^2 + (b_I b_Q')^2 \\ &\quad + 2\,\big(\beta a_Q b_I' - \beta a_I b_Q' + \beta a_I' b_Q - \beta a_Q' b_I + \beta b_Q b_I' - \beta b_I b_Q' \\ &\quad - a_I a_Q b_I' b_Q' + a_Q a_I' b_Q b_I' - a_Q a_Q' b_I b_I' + a_Q b_Q (b_I')^2 - a_Q b_I b_I' b_Q' \\ &\quad - a_I a_I' b_Q b_Q' + a_I a_Q' b_I b_Q' - a_I b_I' b_Q b_Q' + a_I b_I (b_Q')^2 \\ &\quad - a_I' a_Q' b_I b_Q + a_I' (b_Q)^2 b_I' - a_I' b_I b_Q b_Q' \\ &\quad - a_Q' b_I b_I' b_Q + a_Q' (b_I)^2 b_Q' \\ &\quad - b_I b_I' b_Q b_Q'\big). \end{aligned} \tag{3.70}$$

Using arguments similar to those made for the derivation of the expected value of $x_{I_n}^2$, the expected values of all but the first seven of the expanded terms in (3.70) vanish. The first term, whose expected value is non-zero, evaluates to

$$\begin{aligned} E\{\beta^2\} &= E\{S^2 \sin^2 \hat{\Delta}_n\} \\ &= \frac{S^2}{2}\left(1 - E\{\cos 2\hat{\Delta}_n\}\right) \\ &= S^2/2. \end{aligned}$$

Again using the independence of the signal and noise samples, one has

$$E\{a_Q^2\}E\{(b_I')^2\} = E\{a_I^2\}E\{(b_Q')^2\} = E\{(a_I')^2\}E\{b_Q^2\} = E\{(a_Q')^2\}E\{b_I^2\} = \frac{S\,\Xi}{4}.$$

The expected values of the two remaining terms are evaluated as

$$E\{b_Q^2(b_I')^2\} = E\{b_I^2\}E\{(b_Q')^2\} = \frac{\Xi^2}{4},$$

and, very similarly,

$$E\{b_I^2(b_Q')^2\} = \frac{\Xi^2}{4}.$$

Note that it is not necessary to use Price's Theorem to evaluate the expected values of $E\{b_I^2(b_Q')^2\}$ and $E\{b_Q^2(b_I')^2\}$ since the in-phase and quadrature impairment samples are independent. The expected value of (3.70) thus simplifies to

$$E\{x_{Q_n}^2\} = \frac{S^2}{2} + 4\frac{S\,\Xi}{4} + 2\frac{\Xi^2}{4} = \frac{S^2}{2} + S\,\Xi + \frac{\Xi^2}{2}, \tag{3.71}$$

and using (3.64) and (3.71), (3.66b) becomes

$$\mathrm{Var}\{|x_{Q_n}|\} = \frac{S^2}{2} + S\,\Xi + \frac{\Xi^2}{2} - \frac{S^2}{2} = S\,\Xi + \frac{\Xi^2}{2} = \sigma_{Q,\mathrm{SIV}}^2. \tag{3.72}$$

Summing (3.72) and (3.69), one obtains

$$\sigma_{I,\mathrm{SIV}}^2 + \sigma_{Q,\mathrm{SIV}}^2 = \Xi^2 + 2\,S\,\Xi. \tag{3.73}$$

Solving (3.73) in terms of Ξ and taking the positive root, an expression for the SIV RxDA estimate of the impairment power is

$$\hat{\Xi}_{\mathrm{SIV,RxDA}} = \sqrt{\hat{S}_{\mathrm{SIV,RxDA}}^2 + \hat{\sigma}_{I,\mathrm{SIV}}^2 + \hat{\sigma}_{Q,\mathrm{SIV}}^2} - \hat{S}_{\mathrm{SIV,RxDA}}. \tag{3.74}$$

The in-service, SIV RxDA SIR estimator is expressed as the ratio of $\hat{S}_{\text{SIV,RxDA}}$ to $\hat{\Xi}_{\text{SIV,RxDA}}$ as

$$\begin{aligned}\hat{\rho}_{\text{SIV,RxDA}} &= \frac{\hat{S}_{\text{SIV,RxDA}}}{\sqrt{\hat{S}^2_{\text{SIV,RxDA}} + \hat{\sigma}^2_{I,\text{SIV}} + \hat{\sigma}^2_{Q,\text{SIV}}} - \hat{S}_{\text{SIV,RxDA}}} \\ &= \frac{1}{\sqrt{1 + \frac{\hat{\sigma}^2_{I,\text{SIV}} + \hat{\sigma}^2_{Q,\text{SIV}}}{\hat{S}^2_{\text{SIV,RxDA}}}} - 1}. \end{aligned} \tag{3.75}$$

In practice, expected values are replaced by time averages, so that (3.65) becomes

$$\hat{S}_{\text{SIV,RxDA}} = \frac{1}{\sqrt{2}N_{\text{sym}}} \sum_{n=1}^{N_{sym}} (|x_{I_n}| + |x_{Q_n}|), \tag{3.76}$$

and the sum of (3.72) and (3.69) is evaluated as

$$\begin{aligned}\hat{\sigma}^2_{I,\text{SIV}} + \hat{\sigma}^2_{Q,\text{SIV}} &= \tfrac{1}{N_{sym}-1} \textstyle\sum_{n=1}^{N_{sym}} \left(x^2_{I_n} + x^2_{Q_n}\right) \\ &\quad - \tfrac{1}{N_{sym}(N_{sym}-1)} \left[\left(\textstyle\sum_{n=1}^{N_{sym}} |x_{I_n}|\right)^2 + \left(\textstyle\sum_{n=1}^{N_{sym}} |x_{Q_n}|\right)^2\right]. \end{aligned} \tag{3.77}$$

The SIV TxDA SIR estimator has exactly the same form as (3.75), except that $\hat{S}_{\text{SIV,RxDA}}$ and $\hat{\sigma}^2_{I,\text{SIV}} + \hat{\sigma}^2_{Q,\text{SIV}}$, given by (3.76) and (3.77), respectively, are replaced by

$$\hat{S}_{\text{SIV,TxDA}} = \frac{1}{N_{\text{sym}}} \sum_{n=1}^{N_{sym}} (x_{I_n} \cos\Delta_n + x_{Q_n} \sin\Delta_n) \tag{3.78}$$

and

$$\begin{aligned}\hat{\sigma}^2_{I,\text{SIV}} + \hat{\sigma}^2_{Q,\text{SIV}} &= \tfrac{1}{N_{sym}-1} \textstyle\sum_{n=1}^{N_{sym}} \left(x^2_{I_n} + x^2_{Q_n}\right) \\ &\quad - \frac{\left(\sum_{n=1}^{N_{sym}} \sqrt{2}\, x_{I_n} \cos\Delta_n\right)^2 + \left(\sum_{n=1}^{N_{sym}} \sqrt{2}\, x_{Q_n} \cos\Delta_n\right)^2}{N_{\text{sym}}(N_{\text{sym}}-1)}, \end{aligned} \tag{3.79}$$

respectively, where (3.55) and (3.56) are used with the estimated differential phase at time index n, $\hat{\Delta}_n$, replaced by the actual transmitted differential phase, Δ_n.

3.7.8 The Modified Signal-to-Impairment Variance (MSIV) Estimator

The MSIV estimator is a hybrid of the PML and SIV estimators. The signal power is computed by averaging the de-rotated differentially-detected symbols in the same

manner as the PML estimator. However, the impairment power is computed using the variances of the *radial* and *tangential* components of the post-DD symbols rather than the variances of the real and imaginary components as for the SIV estimator. The distinction between the two methods is clarified below.

The MSIV TxDA expression for the signal power is identical to (3.47) which is the expression used by the PML TxDA estimator of Section 3.7.6. The MSIV TxDA estimate of the signal power may thus be simply stated as

$$\hat{S}_{\text{MSIV,TxDA}} = E\{x_n e^{-j\Delta_n}\}. \tag{3.80}$$

The impairment power is computed as a function of the variances of the radial and tangential parts of the post-DD symbols, which is identical to the variances of the real and imaginary parts of the *de-rotated* post-DD symbols. The variances of the real and imaginary parts of the de-rotated post-DD symbols are evaluated as

$$\text{Var}\{\text{Re}\{x_n e^{-j\Delta_n}\}\} = E\{(\text{Re}\{x_n e^{-j\Delta_n}\})^2\} - (E\{\text{Re}\{x_n e^{-j\Delta_n}\}\})^2, \tag{3.81}$$

and

$$\text{Var}\{\text{Im}\{x_n e^{-j\Delta_n}\}\} = E\{(\text{Im}\{x_n e^{-j\Delta_n}\})^2\} - (E\{\text{Im}\{x_n e^{-j\Delta_n}\}\})^2, \tag{3.82}$$

respectively. From (3.47) it is clear that

$$E\{\text{Re}\{x_n e^{-j\Delta_n}\}\} = S \tag{3.83}$$

$$E\{\text{Im}\{x_n e^{-j\Delta_n}\}\} = 0. \tag{3.84}$$

For the evaluation of $E\{(\text{Re}\{x_n e^{-j\Delta_n}\})^2\}$ and $E\{(\text{Im}\{x_n e^{-j\Delta_n}\})^2\}$, it is convenient to use the notation given by (3.52) in Section 3.7.7. Using that notation and (3.51), the de-rotated symbols are expressed as

$$\begin{aligned} x_n e^{-j\Delta_n} = S &+ (a_I + ja_Q)(b'_I - jb'_Q)(\cos\Delta_n + j\sin\Delta_n) \\ &+ (b_I + jb_Q)(a'_I - ja'_Q)(\cos\Delta_n + j\sin\Delta_n) \\ &+ (b_I + jb_Q)(b'_I - jb'_Q)(\cos\Delta_n + j\sin\Delta_n). \end{aligned} \tag{3.85}$$

Consider first the evaluation of $E\{(\mathrm{Im}\{x_n e^{-j\Delta_n}\})^2\}$. After expanding terms, the imaginary part of (3.85) is found to be

$$\begin{aligned}\mathrm{Im}\{x_n e^{-j\Delta_n}\} &= (a_Q b'_I - a_I b'_Q + a'_I b_Q - a'_Q b_I + b'_I b_Q - b_I b'_Q)\cos\Delta_n \\ &\quad + (a_I b'_I + a_Q b'_Q + a'_I b_I + a'_Q b_Q + b_I b'_I + b_Q b'_Q)\sin\Delta_n. \quad (3.86)\end{aligned}$$

Squaring (3.86), one obtains

$$\begin{aligned}&\left[\mathrm{Im}\{x_n e^{-j\Delta_n}\}\right]^2 = \\ &\quad (a_Q b'_I - a_I b'_Q + a'_I b_Q - a'_Q b_I + b'_I b_Q - b_I b'_Q)^2 \frac{1+\cos 2\Delta_n}{2} \\ &\quad + [(a_Q b'_I - a_I b'_Q + a'_I b_Q - a'_Q b_I + b'_I b_Q - b_I b'_Q) \\ &\qquad \times (a_I b'_I + a_Q b'_Q + a'_I b_I + a'_Q b_Q + b_I b'_I + b_Q b'_Q)]\sin 2\Delta_n \\ &\quad + (a_I b'_I + a_Q b'_Q + a'_I b_I + a'_Q b_Q + b_I b'_I + b_Q b'_Q)^2 \frac{1-\cos 2\Delta_n}{2}. \quad (3.87)\end{aligned}$$

As in Section 3.7.7, the expected values of all "cross-terms" vanish. That is, due to the assumed independence and correlation properties of the signal and noise, the expected values of the following products resulting from the expansion of (3.87) vanish:

- All products with a single signal or noise term.
- All products in which the cross-correlation of I and Q signal terms and/or I and Q noise terms occur.
- All products in which the correlation of the n^{th} and $(n-1)^{\text{th}}$ signal and/or noise terms occurs in isolation. For example, $E\{b_I b'_I\} = 0$, but $E\{b_I b_I b'_I\}$ does not necessarily evaluate to zero.

Omitting these vanishing terms, the expected value of (3.87) simplifies to

$$\begin{aligned}&E\{(\mathrm{Im}\{x_n e^{-j\Delta_n}\})^2\} = \\ &\quad \frac{1}{2}\left[E\{a_Q^2(b'_I)^2 + a_I^2(b'_Q)^2 + (a'_I)^2 b_Q^2 + (a'_Q)^2 b_I^2 + (b'_I)^2 b_Q^2 + b_I^2(b'_Q)^2\}\right. \\ &\quad \left. + E\{a_I^2(b'_I)^2 + a_Q^2(b'_Q)^2 + (a'_I)^2 b_I^2 + (a'_Q)^2 b_Q^2 + b_I^2(b'_I)^2 + b_Q^2(b'_Q)^2\}\right.\end{aligned}$$

$$+ E\{a_Q^2 \cos 2\Delta_n\}(E\{(b_I')^2\} - E\{(b_Q')^2\}) + E\{a_I^2 \cos 2\Delta_n\}(E\{(b_Q')^2\} - E\{(b_I')^2\})$$
$$+ E\{(a_I')^2 \cos 2\Delta_n\}(E\{b_Q^2\} - E\{b_I^2\}) + E\{(a_Q')^2 \cos 2\Delta_n\}(E\{b_I^2\} - E\{b_Q^2\})$$
$$+ E\{(b_I')^2 b_Q^2 - b_I^2(b_I')^2 + b_I^2(b_Q')^2 - b_Q^2(b_Q')^2\} E\{\cos 2\Delta_n\}\Big]. \tag{3.88}$$

Noting that

$$E\{a_I^2\} = E\{a_Q^2\} = E\{(a_I')^2\} = E\{(a_Q')^2\} = S/2 \tag{3.89a}$$
$$E\{b_I^2\} = E\{b_Q^2\} = E\{(b_I')^2\} = E\{(b_Q')^2\} = \Xi/2, \tag{3.89b}$$

(3.88) further simplifies to

$$E\{(\mathrm{Im}\{x_n e^{-j\Delta_n}\})^2\} = S\Xi + \Xi^2/2. \tag{3.90}$$

Substituting (3.84) and (3.90) into (3.82), one obtains

$$\mathrm{Var}\{\mathrm{Im}\{x_n e^{-j\Delta_n}\}\} = S\Xi + \Xi^2/2 = \sigma^2_{Q,\mathrm{MSIV}}. \tag{3.91}$$

Consider next the evaluation of $E\{(\mathrm{Re}\{x_n e^{-j\Delta_n}\})^2\}$. The real part of (3.85) is found to be

$$\begin{aligned}\mathrm{Re}\{x_n e^{-j\Delta_n}\} &= S + (a_I b_I' + a_Q b_Q' + a_I' b_I + a_Q' b_Q + b_I b_I' + b_Q b_Q')\cos\Delta_n \\ &\quad - (a_Q b_I' - a_I b_Q' + a_I' b_Q - a_Q' b_I + b_I' b_Q - b_I b_Q')\sin\Delta_n \\ &= S + A\cos\Delta_n - B\sin\Delta_n, \end{aligned}\tag{3.92}$$

where, for convenience, A and B are defined as

$$A = a_I b_I' + a_Q b_Q' + a_I' b_I + a_Q' b_Q + b_I b_I' + b_Q b_Q' \tag{3.93}$$

and

$$B = a_Q b_I' - a_I b_Q' + a_I' b_Q - a_Q' b_I + b_I' b_Q - b_I b_Q', \tag{3.94}$$

respectively. Squaring (3.92), one obtains

$$\begin{aligned}[\mathrm{Re}\{x_n e^{-j\Delta_n}\}]^2 &= S^2 + A^2\cos^2\Delta_n + B^2\sin^2\Delta_n \\ &\quad + 2(SA\cos\Delta_n - SB\sin\Delta_n - AB\,\sin\Delta_n\cos\Delta_n) \\ &= S^2 + \tfrac{1}{2}[A^2 + B^2 + (A^2 - B^2)\cos 2\Delta_n] \\ &\quad + 2(SA\cos\Delta_n - SB\sin\Delta_n - \tfrac{1}{2}AB\,\sin 2\Delta_n). \end{aligned}\tag{3.95}$$

Using (3.93), (3.94), and the arguments presented earlier in the derivation of $E\{(\mathrm{Im}\{x_n e^{-j\Delta_n}\})^2\}$, one has

$$E\{A\cos\Delta_n\} = E\{B\sin\Delta_n\} = E\{AB\sin 2\Delta_n\} = 0.$$

Further, the expected values of all cross-terms that arise from the expansion of A^2 and B^2 vanish so that the expected value of (3.95) may be written as

$$\begin{aligned}
&E\{(\mathrm{Re}\{x_n e^{-j\Delta_n}\})^2\} = \\
&\quad S^2 + \frac{1}{2}\Big[E\{a_I^2(b_I')^2 + a_Q^2(b_Q')^2 + (a_I')^2 b_I^2 + (a_Q')^2 b_Q^2 + b_I^2(b_I')^2 + b_Q^2(b_Q')^2\} \\
&\quad + E\{a_Q^2(b_I')^2 + a_I^2(b_Q')^2 + (a_I')^2 b_Q^2 + (a_Q')^2 b_I^2 + (b_I')^2 b_Q^2 + b_I^2(b_Q')^2\} \\
&\quad + E\{a_I^2\cos 2\Delta_n\}(E\{(b_I')^2\} - E\{(b_Q')^2\}) + E\{a_Q^2\cos 2\Delta_n\}(E\{(b_Q')^2\} - E\{(b_I')^2\}) \\
&\quad + E\{(a_I')^2\cos 2\Delta_n\}(E\{b_I^2\} - E\{b_Q^2\}) + E\{(a_Q')^2\cos 2\Delta_n\}(E\{b_Q^2\} - E\{b_I^2\}) \\
&\quad + E\{b_I^2(b_I')^2 - (b_I')^2 b_Q^2 + b_Q^2(b_Q')^2 - b_I^2(b_Q')^2\}E\{\cos 2\Delta_n\}\Big]. \qquad (3.96)
\end{aligned}$$

Using (3.89), (3.96) simplifies to

$$E\{(\mathrm{Re}\{x_n e^{-j\Delta_n}\})^2\} = S^2 + S\Xi + \Xi^2/2. \qquad (3.97)$$

Substituting (3.83) and (3.97) into (3.81), one obtains

$$\mathrm{Var}\{\mathrm{Re}\{x_n e^{-j\Delta_n}\}\} = S^2 + S\Xi + \Xi^2/2 - S^2 = S\Xi + \Xi^2/2 = \sigma^2_{I,\mathrm{MSIV}}. \qquad (3.98)$$

Summing (3.91) and (3.98), solving for Ξ and taking the positive root, an expression for the MSIV TxDA estimate of the impairment power may be obtained as

$$\hat{\Xi}_{\mathrm{MSIV,TxDA}} = \sqrt{\hat{S}^2_{\mathrm{MSIV,TxDA}} + \hat{\sigma}^2_{I,\mathrm{MSIV}} + \hat{\sigma}^2_{Q,\mathrm{MSIV}}} - \hat{S}_{\mathrm{MSIV,TxDA}}. \qquad (3.99)$$

The MSIV TxDA SIR estimator is expressed as the ratio of $\hat{S}_{\mathrm{MSIV,TxDA}}$ to $\hat{\Xi}_{\mathrm{MSIV,TxDA}}$ as

$$\begin{aligned}
\hat{\rho}_{\mathrm{MSIV,TxDA}} &= \frac{\hat{S}_{\mathrm{MSIV,TxDA}}}{\sqrt{\hat{S}^2_{\mathrm{MSIV,TxDA}} + \hat{\sigma}^2_{I,\mathrm{MSIV}} + \hat{\sigma}^2_{Q,\mathrm{MSIV}}} - \hat{S}_{\mathrm{MSIV,TxDA}}} \\
&= \frac{1}{\sqrt{1 + \frac{\hat{\sigma}^2_{I,\mathrm{MSIV}} + \hat{\sigma}^2_{Q,\mathrm{MSIV}}}{\hat{S}^2_{\mathrm{MSIV,TxDA}}}} - 1}. \qquad (3.100)
\end{aligned}$$

In practice, expected values are replaced by time averages, so that (3.80) becomes

$$\hat{S}_{\text{MSIV,TxDA}} = \frac{1}{N_{sym}} \sum_{n=1}^{N_{sym}} \text{Re}\{x_n e^{-j\Delta_n}\} \tag{3.101}$$

and the sum of (3.91) and (3.98) may be evaluated as

$$\begin{aligned}\hat{\sigma}^2_{I,\text{MSIV}} + \hat{\sigma}^2_{Q,\text{MSIV}} &= \frac{1}{N_{sym}-1}\sum_{n=1}^{N_{sym}}(\text{Re}\{x_n e^{-j\Delta_n}\})^2 + (\text{Im}\{x_n e^{-j\Delta_n}\})^2 \\ &\quad - \frac{\left(\sum_{n=1}^{N_{sym}} \text{Re}\{x_n e^{-j\Delta_n}\}\right)^2 + \left(\sum_{n=1}^{N_{sym}} \text{Im}\{x_n e^{-j\Delta_n}\}\right)^2}{N_{\text{sym}}(N_{\text{sym}}-1)}.\end{aligned} \tag{3.102}$$

The in-service, MSIV RxDA SIR estimator has exactly the same form as (3.100), except that $\hat{S}_{\text{MSIV,TxDA}}$ is replaced by $\hat{S}_{\text{MSIV,RxDA}}$ expressed as

$$\hat{S}_{\text{MSIV,RxDA}} = \frac{1}{N_{\text{sym}}} \sum_{n=1}^{N_{sym}} \text{Re}\{x_n e^{-j\hat{\Delta}_n}\}, \tag{3.103}$$

and $\hat{\sigma}^2_{I,\text{MSIV}} + \hat{\sigma}^2_{Q,\text{MSIV}}$ is evaluated as

$$\begin{aligned}\hat{\sigma}^2_{I,\text{MSIV}} + \hat{\sigma}^2_{Q,\text{MSIV}} &= \frac{1}{N_{sym}-1}\sum_{n=1}^{N_{sym}}(\text{Re}\{x_n e^{-j\hat{\Delta}_n}\})^2 + (\text{Im}\{x_n e^{-j\hat{\Delta}_n}\})^2 \\ &\quad - \frac{\left(\sum_{n=1}^{N_{sym}} \text{Re}\{x_n e^{-j\hat{\Delta}_n}\}\right)^2 + \left(\sum_{n=1}^{N_{sym}} \text{Im}\{x_n e^{-j\hat{\Delta}_n}\}\right)^2}{N_{\text{sym}}(N_{\text{sym}}-1)}\end{aligned} \tag{3.104}$$

where, as before, $\hat{\Delta}_n$ is the estimated transmitted differential phase at time index, n.

The main difference between the general approach of the SIV algorithm of Section 3.7.7 and that of the MSIV algorithm is that the SIV algorithm operates on the Cartesian real and imaginary parts of the differentially-detected symbols whereas the MSIV algorithm operates on the radial and tangential components. Figure 3.19 illustrates this idea.

Comparison of Figure 3.19(b) with Figures 3.14 and 3.18 reveals that *the tangential variance of the post-DD symbols is not affected by signal envelope fluctuations caused by fast Rayleigh fading.* The amplitude fluctuations affect only the real parts of the de-rotated, post-DD symbols (the radial parts of the post-DD symbols before de-rotation). In contrast, referring to Figure 3.19(a) and Figures 3.14 and 3.18, the sample variances required by the SIV estimator are generated from the Cartesian real

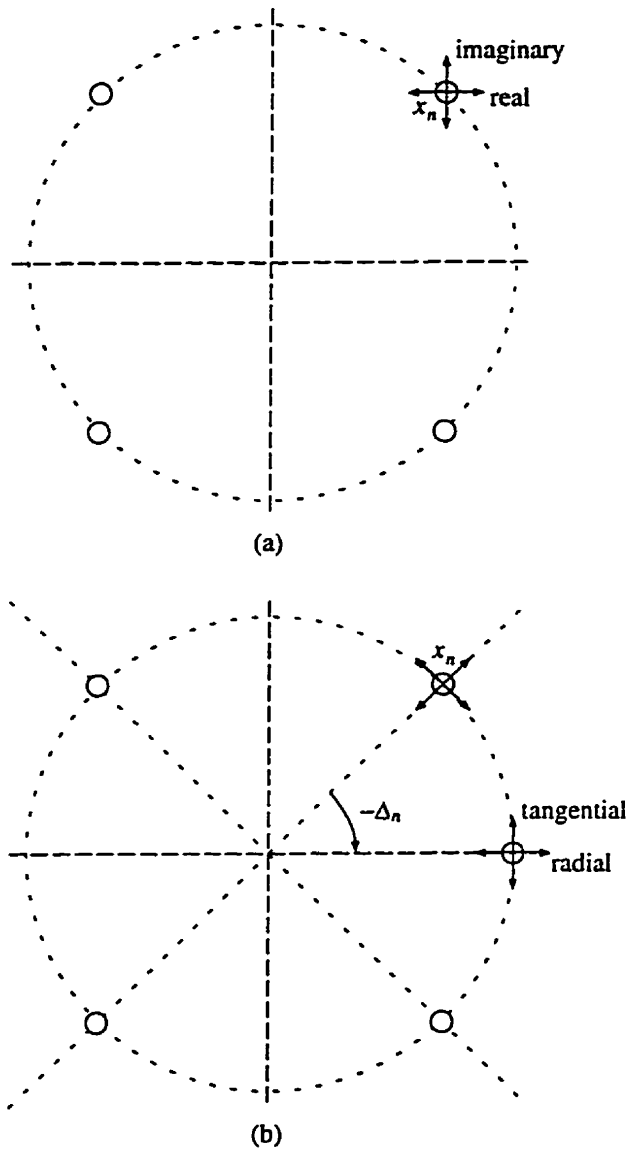

Figure 3.19: Illustration identifying the components of the post-DD samples used in (a) SIV, and (b) MSIV SIR estimation.

and imaginary parts of the post-DD symbols. Both the real and imaginary components are affected by signal envelope fluctuations caused by fast Rayleigh fading. It is not intuitively obvious which estimator should perform better, or even if there should be any performance difference at all.

3.8 Measures of Estimator Performance

This section identifies statistical tests which may be performed on the output of the simulated SIR estimators to evaluate their performances in the various test channels described in Section 3.6.2. Three useful performance metrics, the bias, variance and MSE, were used in Section 2.5.1 to evaluate the performances of the various SNR estimators in the real and complex AWGN channels of Chapter 2. As well, a theoretical bound, the CRB, was derived in Section 2.5.2 and compared to the variances and MSE's of the various SNR estimators. Similar performance metrics are required for the evaluation of the SIR estimators studied in this chapter.

Recall, the unbiased estimators of the sample bias, sample variance, and sample MSE of an estimator, $\hat{\rho}$, given in Section 2.5.1 as

$$\begin{aligned} \text{Bias}\{\hat{\rho}\} &= \frac{1}{N_t}\sum_{i=1}^{N_t}(\hat{\rho}_i - \rho) \\ s_{\hat{\rho}}^2 &= \frac{1}{N_t - 1}\sum_{i=1}^{N_t}(\hat{\rho}_i - \bar{\hat{\rho}})^2 \\ \text{MSE}\{\hat{\rho}\} &= \frac{1}{N_t}\sum_{i=1}^{N_t}(\hat{\rho}_i - \rho)^2 \end{aligned}$$

where $\hat{\rho}_i$ denotes one of N_t SNR estimates generated by a particular estimator, ρ is the true SNR, and the unbiased sample mean, $\bar{\hat{\rho}}$, is given by

$$\bar{\hat{\rho}} = \frac{1}{N_t}\sum_{i=1}^{N_t}\hat{\rho}_i.$$

Unfortunately, it is not possible to evaluate the performances of the SIR estimators as rigorously as the performances of the SNR estimators because of the fundamental challenge to identify the "true" SIR of a given channel. The simplest case is the

complex AWGN channel of Case 1 where the SIR is simply the SNR, S/N. Similarly, in the Case 4 channel in which CCI is dominant over every other impairment, the SIR may be expressed as S/C, where C represents the power of the CCI sequence, *assuming the interferer is time-synchronized with the desired signal* so that the symbol centres of the interferer align with the symbol centres of the desired signal. If the two signals are *not* time synchronized, then the degradation caused by the interferer is less and the SIR could be expressed as some number greater than S/C. The situation becomes more complicated when considering fading channels (especially the frequency-selective fading channel of Case 3), and more confusing still when the various impairments are combined (as in Case 5)—the impairments all have different statistics and different deleterious effects on the BER so that a composite SIR would have to be formed from some sort of weighted combination of the individual impairment parameters.

Instead of trying to create a universal definition of the SIR in order to be able to compute the bias and MSE performances of the various SIR estimators, it is conceded that any universal SIR definition will be arbitrary, so focus is instead placed on unambiguous performance metrics: the sample variance, and the correlation of the sample mean to the simulated, raw BER of the specified receiver (see Figure 3.10). The "best" estimator is judged to be the one with the smallest variance, and the best correlation of its mean to the BER, for all of the various channel conditions considered.

3.9 Simulation Results

The results of the evaluation of the performances of the SIR estimators discussed in Section 3.7 are presented here. The fading simulator implemented for these tests is discussed in Section 3.4, and the system models are presented in Section 3.6.

For the most part, the SIR estimators are evaluated graphically by plotting the variances of a collection of simulated SIR estimates as a function of some channel

parameter, and by plotting the means of the SIR estimates as a function of the simulated BER. The collection of SIR estimates is large enough to ensure that the error in the estimated variance and the error in the estimated BER is less than approximately 20% with 95% confidence.

As mentioned in Section 3.8, it is difficult to compute the bias and MSE of the SIR estimates since the "true" SIR of a particular system can be an ambiguous parameter. However, the Case 1 channel is simply the complex AWGN channel in which the SIR is unambiguously the SNR, so that bias and MSE results are presented for that particular case. In Case 4, the only impairment, in addition to a small amount of AWGN, is CCI so that the true SIR in that case is simply $S/(N+C)$, thus bias and MSE results are provided for Case 4 as well. For Cases 2, 3, and 5, in which the SIR is not straightforward to define, the estimator performances must be judged on the basis of the variances and the SIR/BER correlations of the SIR estimates alone.

The normalized, complex CRB given by (2.131) of Section 2.5.2 is provided for reference in the plots of variance and MSE for Case 1 since it is applicable to some of the estimators. The pre-MF CRB (described in Section 2.6) is applicable to the ML TxDA and ML RxDA estimators, and the post-MF CRB (also described in Section 2.6) is applicable to the SVR and M_2M_4 estimators. The other estimators studied in this chapter, namely the ADIQ, DDML, PML, SIV, and MSIV estimators, are post-DD estimators to which the CRB given by (2.131) of Section 2.5.2 does not apply because the statistics of the signal and noise after the DD are not consistent with the assumptions made in the derivation of the complex CRB in Section 2.5.2. No CRB is provided for reference in the plots of the variances for Cases 2–5 since the CRB derivations become complicated for impairment processes which do not have Gaussian statistics.

The results are presented in the following sections, grouped according to each of the five channels tested.

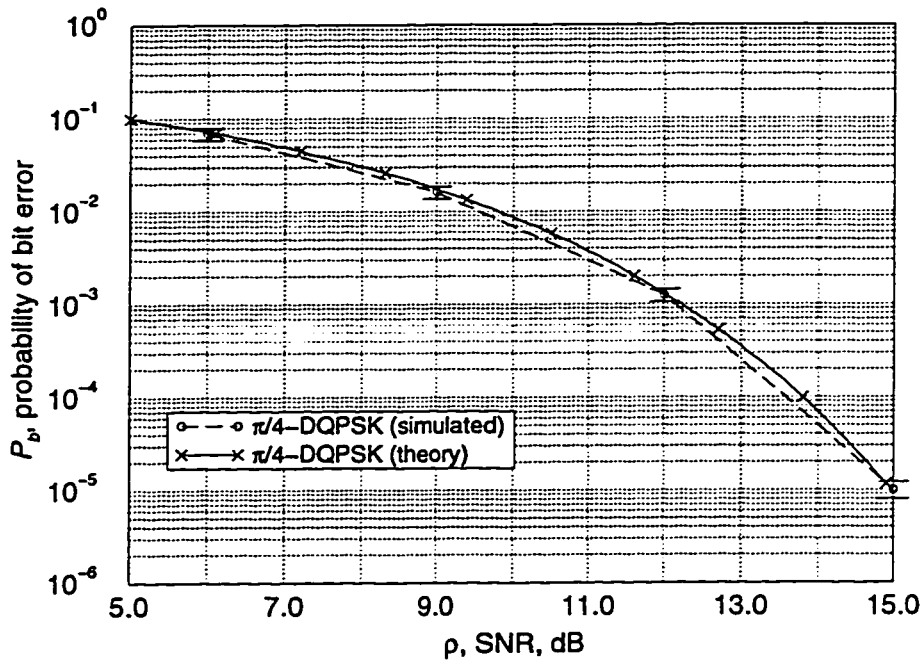

Figure 3.20: BER as a function of SNR in Case 1 channel.

3.9.1 Case1: Complex AWGN Channel

The simulated BER of this Case 1 channel is plotted in Figure 3.20 as a function of SNR, and the theoretical BER is plotted for comparison. The bit-error probability of $\pi/4$-DQPSK is essentially identical to that of conventional DQPSK which is given in [65, page 271] as

$$P_{\text{b,DQPSK}} = Q(a,b) - \tfrac{1}{2} I_0(ab) e^{-(a^2+b^2)/2},$$

where $Q(a,b)$ is the Marcum Q function given by

$$Q(a,b) = e^{-(a^2+b^2)/2} \sum_{k=0}^{\infty} \left(\frac{a}{b}\right)^k I_k(ab),$$

$I_k(x)$ is the k^{th}-order modified Bessel function of the first kind, and a and b are given by

$$a = \sqrt{\rho\left(1 - \frac{1}{\sqrt{2}}\right)},$$

and

$$b = \sqrt{\rho\left(1 + \frac{1}{\sqrt{2}}\right)},$$

respectively, where $\rho = 2E_b/N_0$ is the SNR.

The SIR estimators tested in this channel include the SVR, M_2M_4, PML RxDA, PML TxDA, SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, ADIQ, ML RxDA, and ML TxDA estimators. The statistical performances of these estimators are presented below.

Biases, Variances, and MSE's of Estimators in Case 1 Channel

The estimator biases are plotted in Figures 3.21 and 3.22, the variances are plotted in Figures 3.23 and 3.24, and the MSE's are plotted in Figures 3.25 and 3.26. Note that the bias and MSE of each estimator are plotted normalized to the true SNR, whereas the variance of each estimator is plotted normalized to the *estimated* SNR. It is more appropriate to normalize the variance to the estimated SNR since the normalization of the variance of a biased estimator to the true SNR can yield misleading results.

Due to the large number of estimators tested in the Case 1 channel, each set of results is split among three different graphs. In order to be able to compare estimators from different graphs, the SIV RxDA and SIV TxDA results and the post-MF CRB appear in all of the graphs of variance and MSE as common benchmarks.

As was found in Chapter 2, the ML TxDA estimator in the Case 1 channel has the best overall performance as it can be seen in Figure 3.22 that its bias is very small, and it can be seen in Figures 3.24 and 3.26 that its variance and MSE are practically indistinguishable from the CRB. It's not surprising that the ML TxDA estimator performs much better than the other estimators studied here since it is the only one

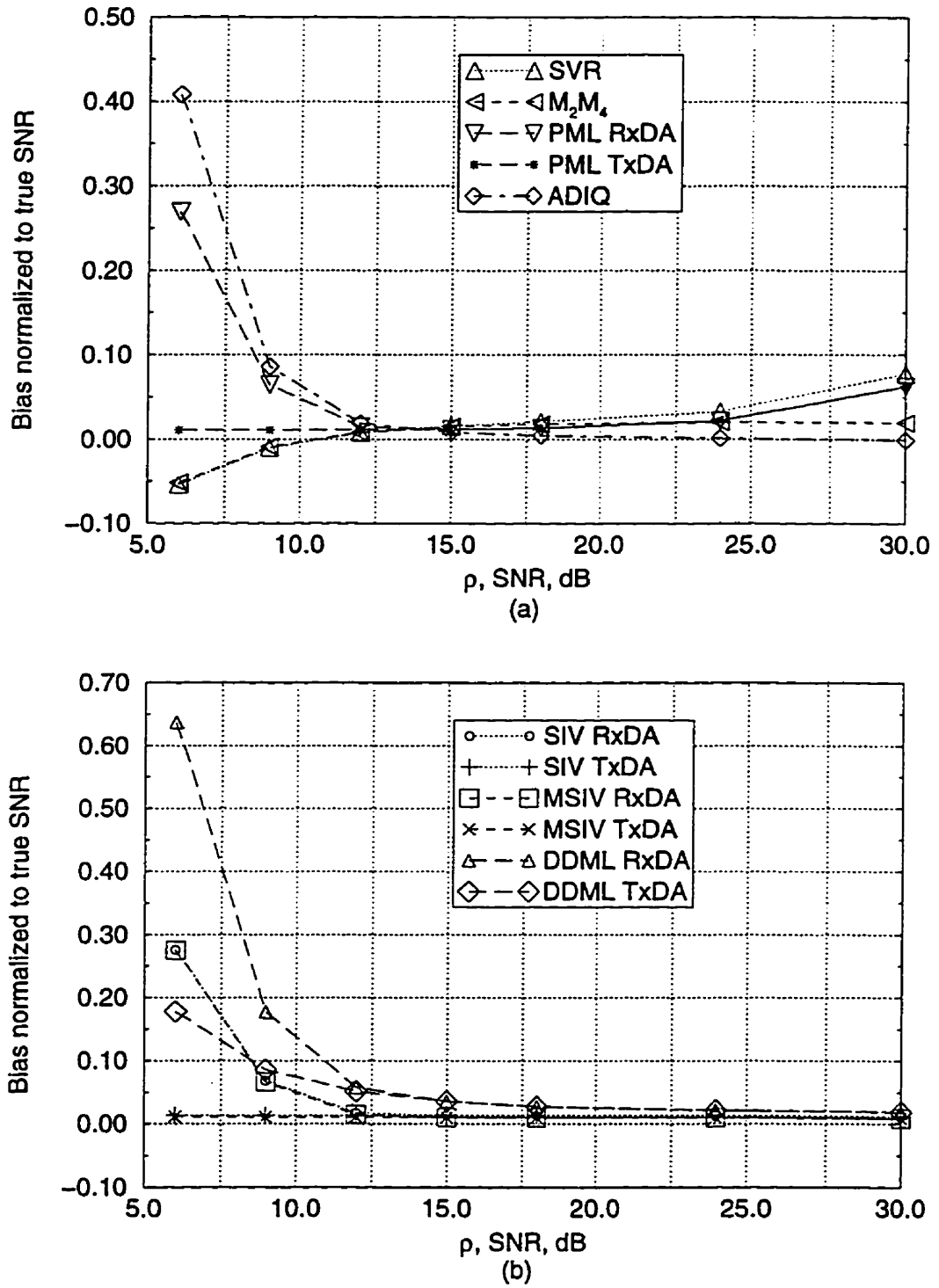

Figure 3.21: Normalized bias of SIR estimates generated by (a) SVR, M_2M_4, PML, and ADIQ estimation, (b) SIV, MSIV, and DDML estimation in Case 1 channel.

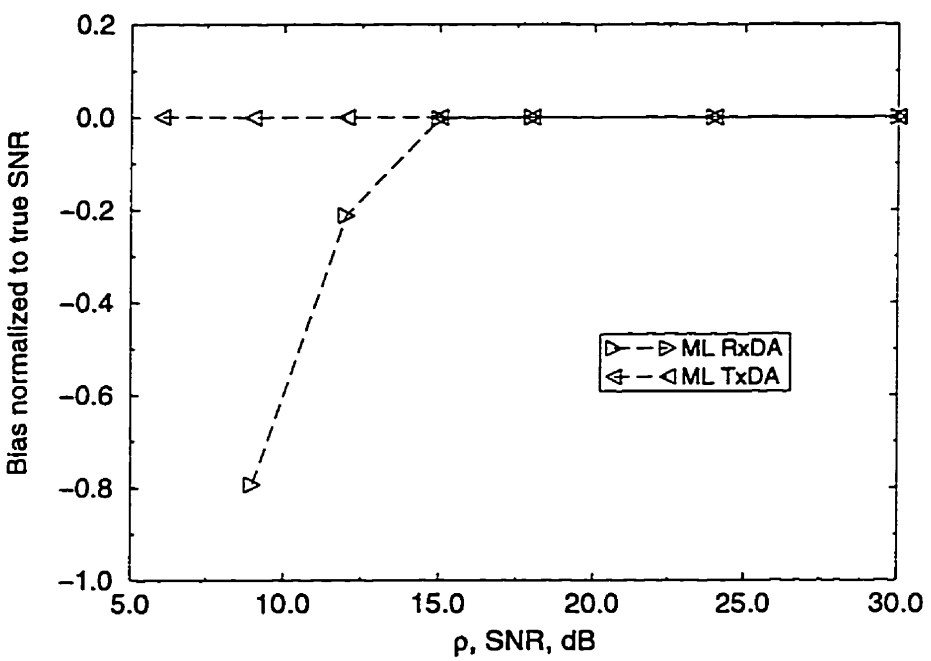

Figure 3.22: Normalized bias of SIR estimates generated by ML estimation (at the input to the MF) in Case 1 channel.

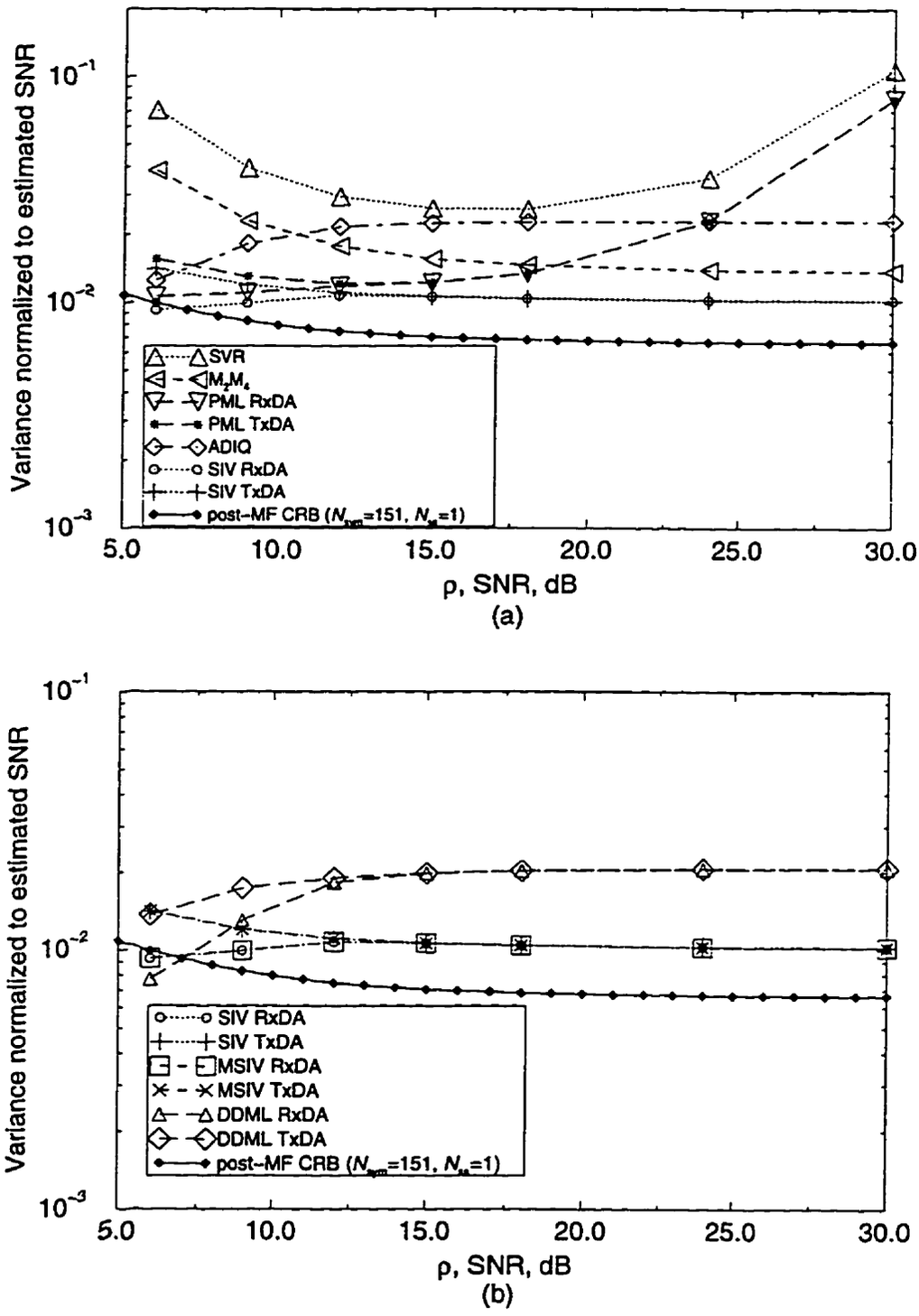

Figure 3.23: Normalized variance of SIR estimates generated by (a) SVR, M_2M_4, PML, ADIQ, and SIV estimation, (b) SIV, MSIV, and DDML estimation in Case 1 channel. The post-MF CRB is shown in both plots for reference.

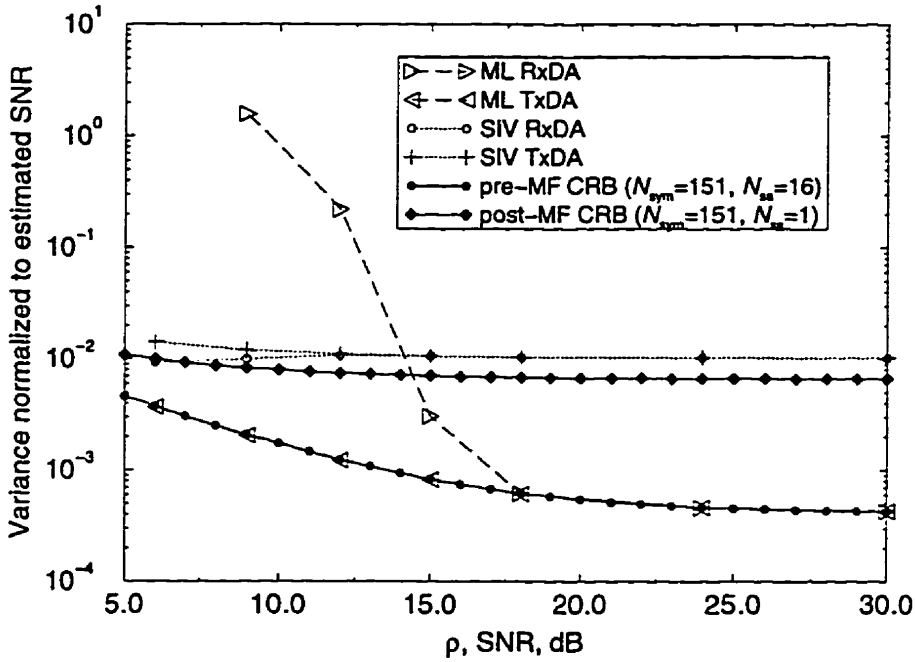

Figure 3.24: Normalized variance of SIR estimates generated by ML estimation (at the input to the MF) in Case 1 channel. The variance curves for the SIV estimators and the pre- and post-MF CRB's are shown for reference.

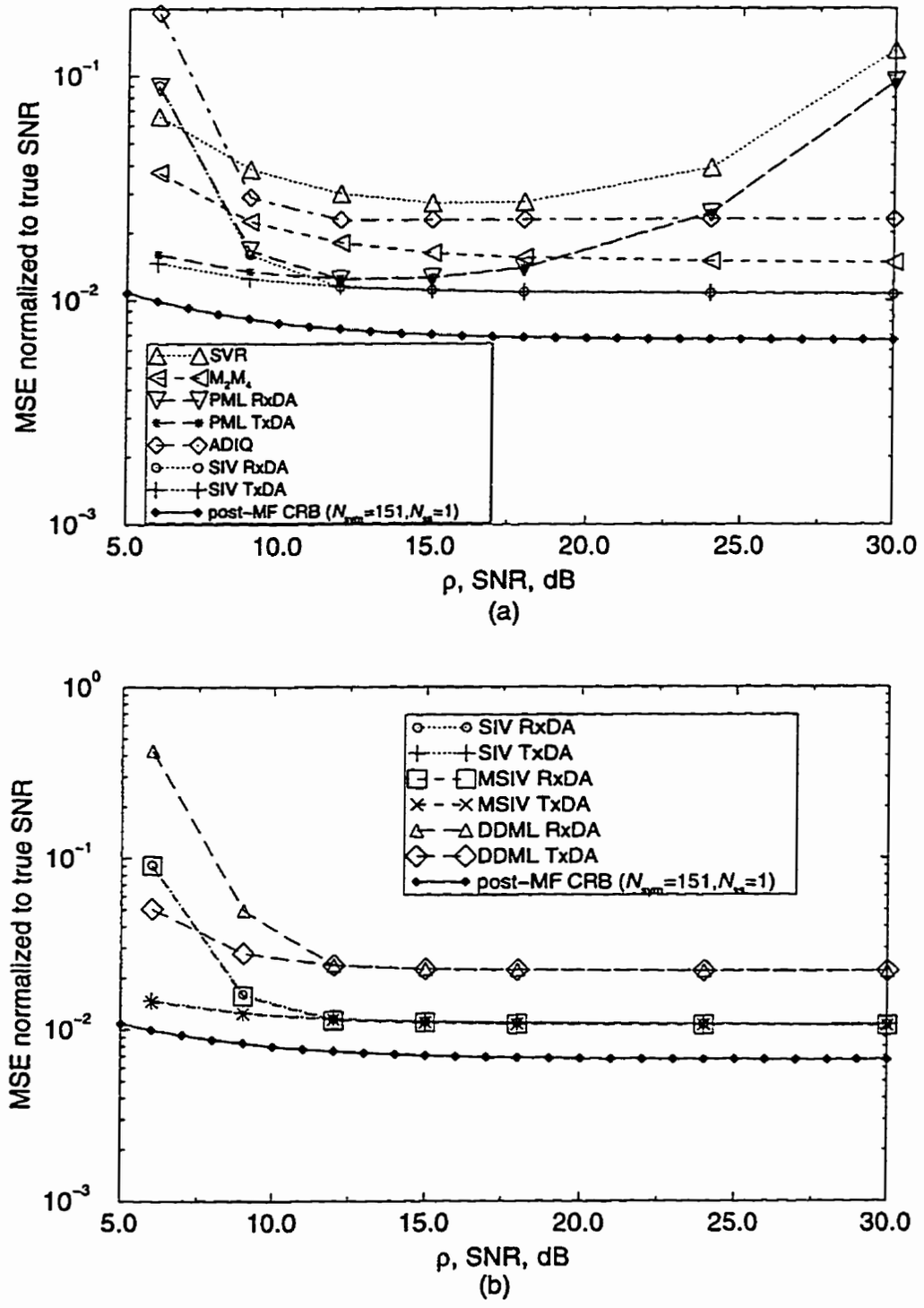

Figure 3.25: Normalized MSE of SIR estimates generated by (a) SVR, M_2M_4, PML, ADIQ, and SIV estimation, (b) SIV, MSIV, and DDML estimation in Case 1 channel. The post-MF CRB is shown in both plots for reference.

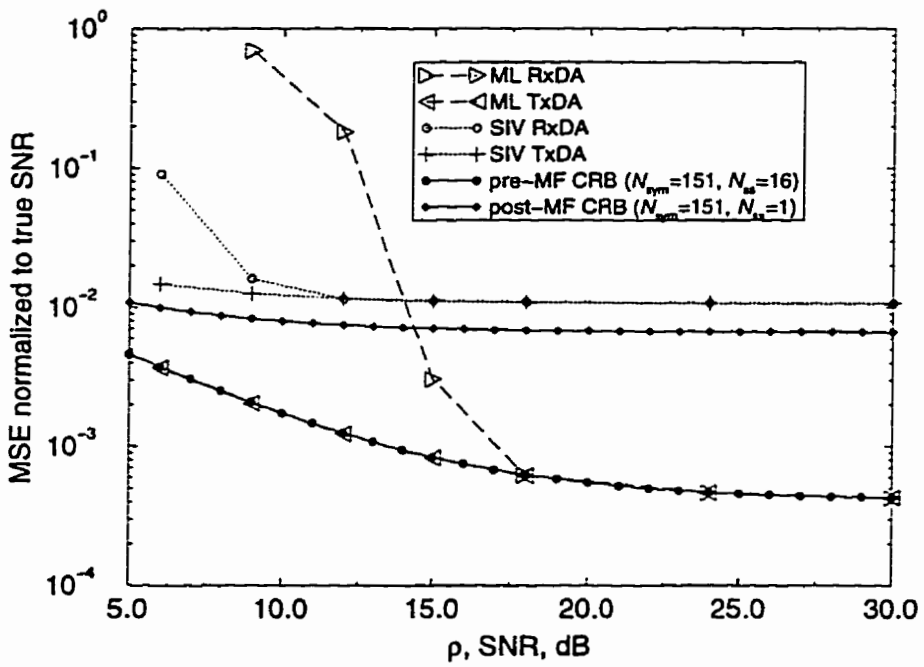

Figure 3.26: Normalized MSE of SIR estimates generated by ML estimation (at the input to the MF) in Case 1 channel. The MSE curves for the SIV estimators and the pre- and post-MF CRB's are shown for reference.

that employs multiple samples per symbol (in this case, sixteen samples per symbol). The ML RxDA estimator performs almost identically to its TxDA counterpart at values of SNR greater than about 18 dB, but its performance degrades sharply for values of SNR less than 18 dB. The reason for this extremely poor performance is that just one receiver error can have a serious effect on estimates produced by the ML RxDA algorithm. This effect was discussed in detail in Section 3.7.4.

Of the estimators that use just one sample per symbol to form SIR estimates, the SIV TxDA and MSIV TxDA estimators perform best as their biases and MSE's are smallest. From Figures 3.21(b), 3.23(b), and 3.25(b), one can see that the performances of the SIV TxDA and MSIV TxDA estimators are essentially identical. The MSIV TxDA estimator seems to have a slight bias and MSE advantage, but the difference is too small to be statistically significant. A similar relationship exists between the SIV RxDA and MSIV RxDA estimators. The performances of both the SIV RxDA and MSIV RxDA estimators deviate from the performances of their TxDA counterparts as the SNR is decreased below about 12–15 dB due to receiver errors, but the degradation is much less severe than that experienced by the ML RxDA estimator. Though the post-MF CRB is not applicable to the post-DD estimators, it is still interesting to compare the performances of the post-DD estimators to the theoretical performance bound of post-MF (pre-DD) estimators. As can be seen in Figures 3.23(b) and 3.25(b), though the SIV TxDA and MSIV TxDA estimators perform best of all of the one-sample-per-symbol estimators, there is a significant performance gap between their variances and MSE's and the post-MF CRB.

Of the estimators using just one sample per symbol which are also of the in-service type, the M_2M_4 estimator appears to have the best overall performance, which is most easily seen in Figure 3.25(a), but the SIV RxDA and MSIV RxDA estimators perform better for values of SNR greater than about 9 dB. The performance of the ADIQ estimator is next best. In Figure 3.21(a), it is particularly noteworthy that, for values of SNR greater than about 15 dB, the bias of the ADIQ is smallest among

all of the SIR estimates tested except for the ML estimators. However, the ADIQ estimator also has the second-largest bias (and MSE) at very low SNR.

Though the DDML estimators are derived using an ML approach (which is an optimum approach) their performances are not as good as the performances of the SIV and MSIV estimators. For example, it can be seen from Figures 3.21 and 3.25 that the biases and MSE's of the DDML estimates are larger than those of the SIV and MSIV estimates. Interestingly, even the TxDA form of the DDML estimator exhibits a significant bias at very low SNR, whereas the TxDA forms of the SIV and MSIV estimators do not. The asymptotic variance of the DDML at high SNR is about twice as large as the asymptotic variances of the SIV and MSIV estimators, as can be seen in Figure 3.23.

The bias, variance, and MSE of the SVR estimator all begin to rise at high SNR which is consistent with the observations made in Sections 2.6.1 and 2.6.2. It is indicated in Chapter 2 that the mechanism behind this degradation at high SNR is not fully understood but, by simulation, it has been observed that this phenomenon is common among some estimators based on higher-order moments. As also pointed out in Chapter 2, the variances and MSE's of the "best" estimators can only be as good as the CRB which itself does not progressively decrease with increasing SNR, but rather approaches a constant value. It is reasonable, therefore, to expect the variance and MSE of an inferior estimator either to approach a constant larger than the high-SNR CRB asymptote, or to rise with increasing SNR.

The performances of the PML TxDA and PML RxDA estimators also degrade with increasing SNR which is surprising since the PML, SIV, and MSIV estimators are closely related and the SIV and MSIV estimators are well-behaved for all values of SNR. In order to understand this high-SNR effect better, it may be insightful to derive the pdf of $\hat{\rho}_{\text{PML,TxDA}}$ and/or the pdf of $\hat{\rho}_{\text{PML,RxDA}}$ so that the bias, variance, and MSE may be computed analytically. This analysis is not performed here. On the other end of the SNR range, it is interesting to note in Figure 3.25(a) that the

MSE of the PML TxDA estimator is slightly inferior both to that of the MSIV TxDA estimator and to that of the SIV TxDA estimator, again, in spite of the fact that the algorithms are very similar. A derivation of the pdf may help to explain this behaviour as well.

It is interesting in Figure 3.23 that the variances of some estimators fall as the SNR becomes very small while the variances of other estimators rise. By comparison with Figure 3.21, it may be seen that those estimators whose variances "droop" at low SNR also have large low-SNR biases. This observation may be used in those test channels (Cases 2A, 2B, 3, and 5) for which bias and MSE results are not presented. That is, in the absence of plots of the bias and MSE, it may be deduced that a "drooping" variance indicates a large low-SNR bias.

Correlation of SIR to BER in Case 1 Channel

The correlation between the mean estimated SIR (simply the SNR in this case) and the probability of bit error, P_b, or BER, of the Case 1 channel is graphically portrayed in Figure 3.27 for the SVR, M_2M_4, PML RxDA, PML TxDA, ADIQ, SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, ML RxDA, and ML TxDA estimators. Most of the SIR estimators have very similar good correlations with the BER, with the exception of the ML RxDA estimator which sometimes generates negative SIR estimates at $\rho = 6$ dB (which is the reason why this point is omitted from the ML RxDA plot).

Based on Figure 3.27 alone, there is little to distinguish one estimator from another. Additional observations regarding the SIR/BER correlations are made in Section 3.9.6.

Additional Results for Case 1 Channel

In Sections 3.7.1 and 3.7.2, the kurtosis of the impairment process, k_ξ, is assumed to be equal to 1.5 as a compromise between the kurtosis of complex AWGN ($k_{\xi,\mathrm{AWGN}} = 2$)

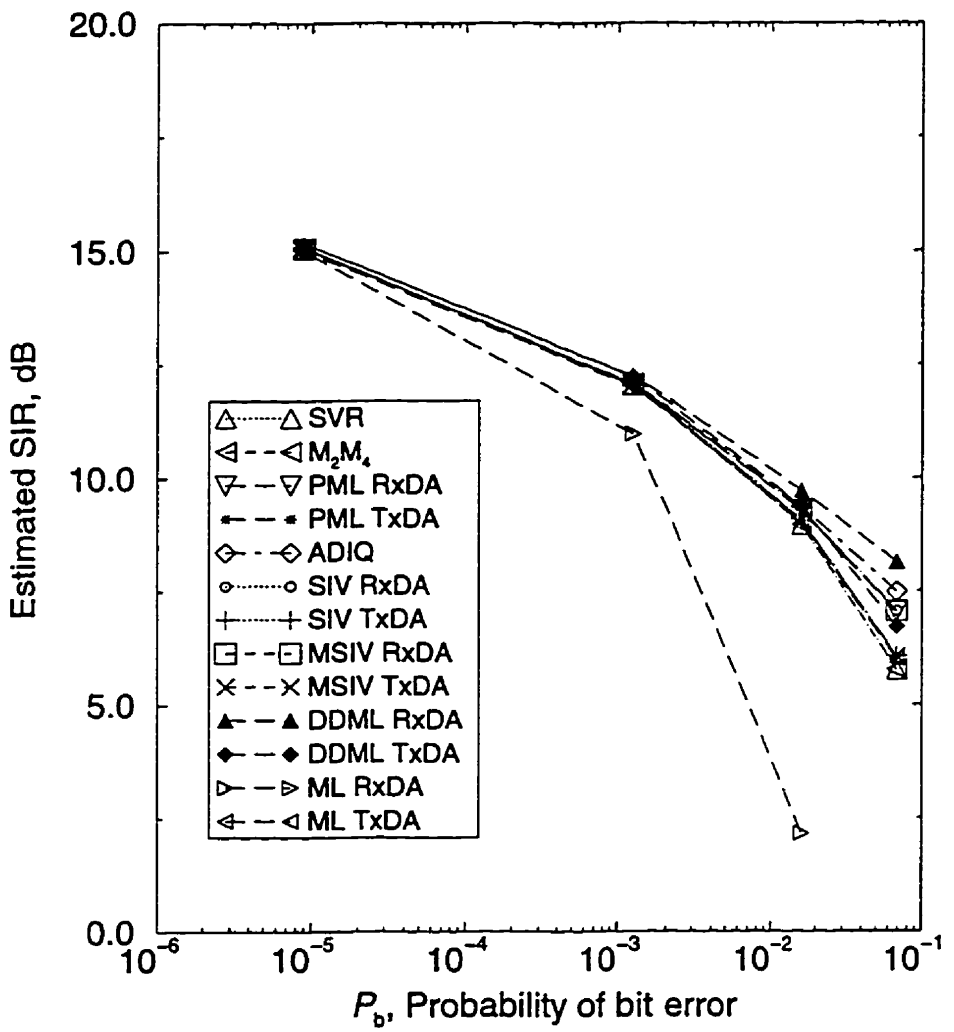

Figure 3.27: Correlation between mean estimated SIR (SNR) and probability of bit error (BER) in Case 1 channel for SVR, M_2M_4, PML, ADIQ, SIV, MSIV, DDML, and ML estimation.

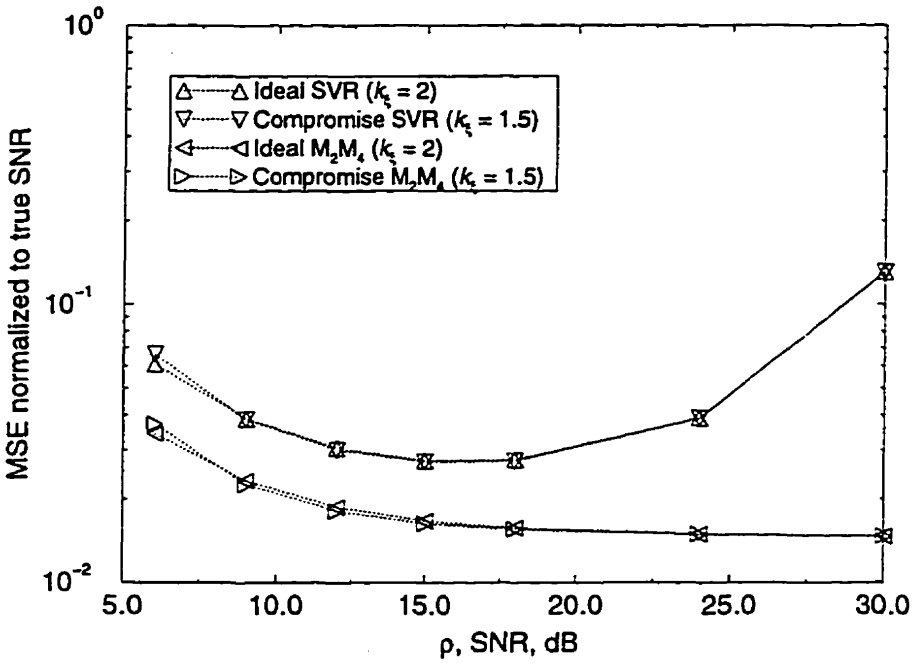

Figure 3.28: Comparison of normalized MSE's of ideal M_2M_4 and SVR estimators to those of compromise M_2M_4 and SVR estimators in Case 1 channel.

and that of a cochannel interferer ($k_{\xi,\mathrm{CCI}} = 1$) in the derivation of the M_2M_4 and SVR estimators. It is interesting to investigate the effect of this compromise on the performances of the M_2M_4 and SVR estimators in the complex AWGN channel of Case 1. It is sufficient to investigate the MSE performances alone.

The MSE's of "ideal" and "compromise" M_2M_4 and SVR estimators are plotted in Figure 3.28. The ideal estimators use $k_\xi = 2$ while the compromise estimators use $k_\xi = 1.5$. It is evident from Figure 3.28 that the ideal estimators outperform the compromise estimators only at very low SNR and otherwise perform essentially identically. This observation indicates that there may be little justification to motivate the implementation of an estimator that adapts itself to a particular channel based

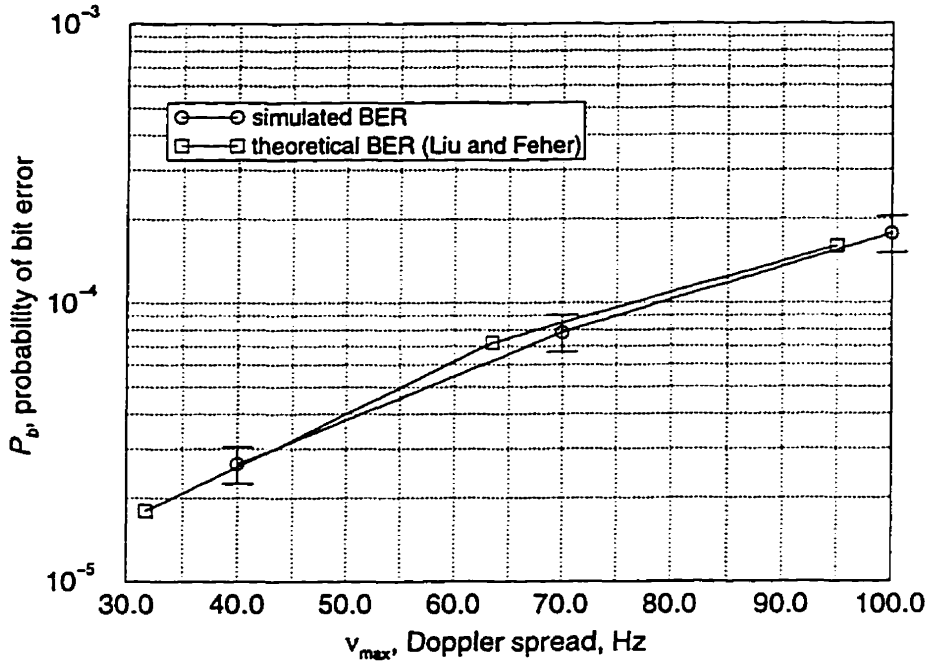

Figure 3.29: BER as a function of Doppler spread in Case 2A channel.

on the kurtosis of the impairment process of the channel.

3.9.2 Case2: Fast, Flat Fading Channel

As indicated in Section 3.6.2, two models of the Case 2 channel are investigated: one in which the Rayleigh amplitude and phase fluctuations are allowed to evolve over time (Case 2A), and another in which the channel conditions are replayed for each successive SIR estimate (Case 2B). Separate results are given for each model.

Case 2A: Time-evolving Rayleigh Fading

The simulated BER of the Case 2A channel is plotted in Figure 3.29 as a function of the Doppler spread, ν_{max}. For reference, a theoretical BER curve is also plotted

based on points obtained by the numerical analysis of Liu and Feher (see (19d) and Figure 7 of [11]). The plotted points are for $S/N = 70$ dB, which is large enough for the BER to be independent of the SNR (due to the error floor caused by the Rayleigh fading). There is good agreement between simulation and theory.

The SIR estimators tested in this channel include the SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ estimators. Results for the SVR, M_2M_4, PML RxDA and PML TxDA estimators are not presented since these algorithms fail in this Case 2A channel. Specifically, some of the SIR estimates produced by the M_2M_4 estimator are imaginary, and some of the SIR estimates produced by the SVR, PML RxDA, and PML TxDA estimators are negative. Since an imaginary or negative estimate of the SIR is meaningless, these four estimators are omitted from the performance comparison for Case 2A. It appears that these methods fail whenever a deep fade occurs.

The variances of the SIR estimates generated by the SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ estimators in the Case 2A channel are plotted in Figure 3.30, and the relationships between the mean estimated SIR and the BER of the Case 2A channel for the various estimators are portrayed in Figure 3.31.

In general, it is reasonable to expect the variance of SIR estimates to increase with decreasing Doppler spread since the SIR estimates begin to track the Rayleigh amplitude fluctuations as the rapidity of the fluctuations slows relative to the observation window of N_{sym} symbols. The SIV and MSIV estimators reflect a behaviour that is consistent with this hypothesis while the DDML and ADIQ estimators do not. Since the DDML and ADIQ estimators do not use the amplitude information of the received signal, it is reasonable to expect the variances of these particular estimators to have a weak relationship, or no relationship, to the Rayleigh amplitude fluctuations.

Comparing the variances of the SIV and MSIV SIR estimates to those of the ADIQ and DDML SIR estimates, the results plotted in Figure 3.30 indicate that

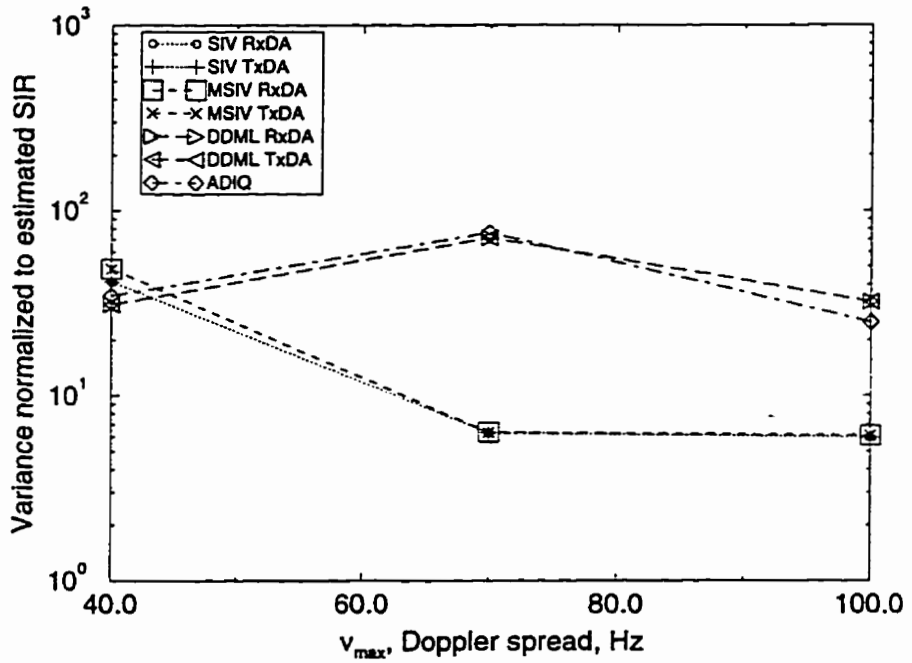

Figure 3.30: Normalized variance of SIR estimates generated by SIV, MSIV, DDML, and ADIQ estimation in Case 2A channel.

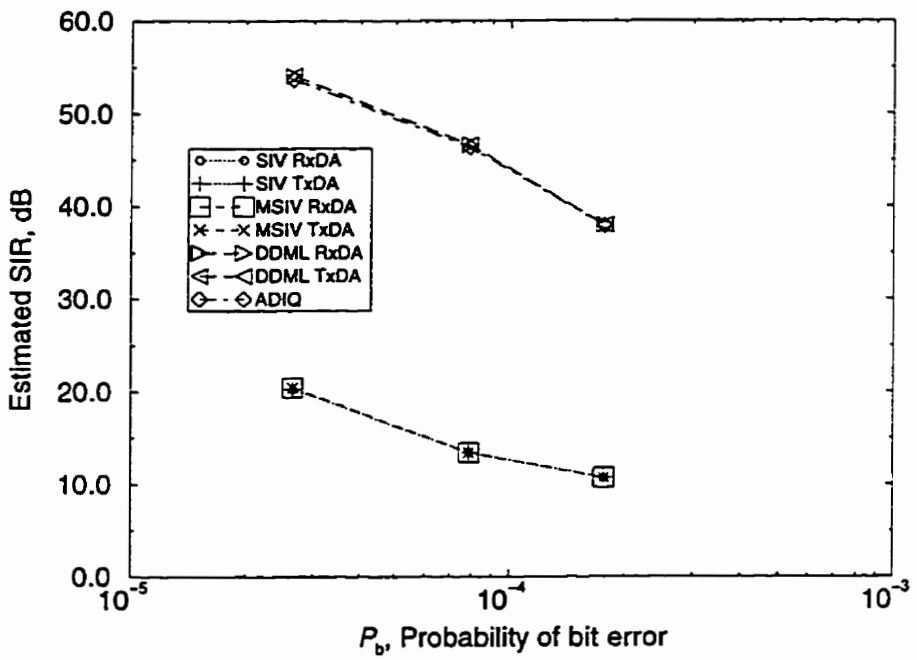

Figure 3.31: Correlation between mean estimated SIR and BER in Case 2A channel for SIV, MSIV, DDML, and ADIQ estimation.

the normalized variances of the SIV and MSIV SIR estimates are almost an order of magnitude smaller than those of the DDML and ADIQ SIR estimates at Doppler spreads of 70 and 100 Hz, but the variances of the SIR estimates of all seven estimators tested are almost equal at a Doppler spread of 40 Hz. One could expect, therefore, that SIV and MSIV estimators may perform better in applications in which the channel is degraded by relatively fast Rayleigh fading.

Referring to Figure 3.31, it is interesting that the estimator performance curves split neatly into two groups as they do in Figure 3.30. The SIV and MSIV curves form one group, and the DDML and ADIQ estimators form the other. More striking, however, is the fact that SIR estimates represented by these two groups of curves are separated by about 30 dB. As will be shown in Figures 3.50 to 3.55, the SIV and MSIV curves correspond to better SIR/BER correlations.

The reason why there is such a tremendously large difference between the absolute DDML and ADIQ SIR estimates and the absolute SIV and MSIV SIR estimates is that the amplitude of the input samples is ignored in both DDML and ADIQ estimation. In DDML estimation, only differential phases are considered. In ADIQ estimation, an instantaneous AGC removes all amplitude information. The effect of the absence of amplitude information is clarified below using the example of ADIQ estimation.

The effect of the instantaneous AGC of the ADIQ estimator can be appreciated by referring to Figure 3.14(c) which shows the post-DD symbols for a channel subjected to fast Rayleigh fading. Without AGC (as portrayed in Figure 3.14(c)), the amplitudes of the post-DD symbols vary greatly due to the random amplitude fluctuations (these fluctuations are reflected by the SIV and MSIV estimates). The instantaneous AGC of the ADIQ estimator normalizes the amplitudes of the post-DD symbols such that the amplitudes of all post-DD symbols lie on the unit circle. For example, after application of an instantaneous AGC operation, the post-DD symbols portrayed by Figure 3.14(c) would appear similar to the post-DD symbols of Figure 3.11(c), which are the post-DD symbols in a clear-sky channel. By this illustration, it is reasonable

to expect the ADIQ SIR estimator to give an overly optimistic estimate of the SIR in a fast Rayleigh fading channel.

It is interesting that there is little differentiation between the performances of the RxDA and TxDA SIV, MSIV, and DDML estimators in Figures 3.30 and 3.31. This observation indicates that receiver errors due to the channel conditions of the Case 2A channel do not have a great impact on the variances of the RxDA SIR estimates compared to the effect of the time-varying channel conditions.

Case 2B: Rayleigh Fading Replayed for each Successive SIR Estimate

The simulated BER of the Case 2B channel is plotted in Figure 3.32 as a function of the SNR. No theoretical BER curve is available in the published literature for reference since the Case 2B channel conditions are artificial and atypical (recall that the amplitude and phase fluctuations are replayed for each successive SIR estimate).

The SIR estimators tested in this channel are the same ones tested in the Case 2A channel, namely the SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ estimators. As in the Case 2A channel, results for the SVR, M_2M_4, PML RxDA and PML TxDA estimators are not presented in this Case 2B channel since *all* of the SIR estimates produced by the M_2M_4 estimator are imaginary, *all* of the SVR SIR estimates are negative, and a large percentage of the PML RxDA, and PML TxDA SIR estimates are negative. As mentioned earlier in the context of the Case 2A channel, "bad" SIR estimates seem to be associated with the deep nulls caused by Rayleigh fading. The Case 2B channel is designed to have the same deep null occur for each successive SIR estimate so it is not surprising that bad SIR estimates would be more prevalent in the Case 2B channel.

The variances of the SIR estimates generated by the SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ estimators in the Case 2B channel are plotted in Figure 3.33, and the relationships between the mean estimated SIR and the BER of the Case 2B channel for the various estimators

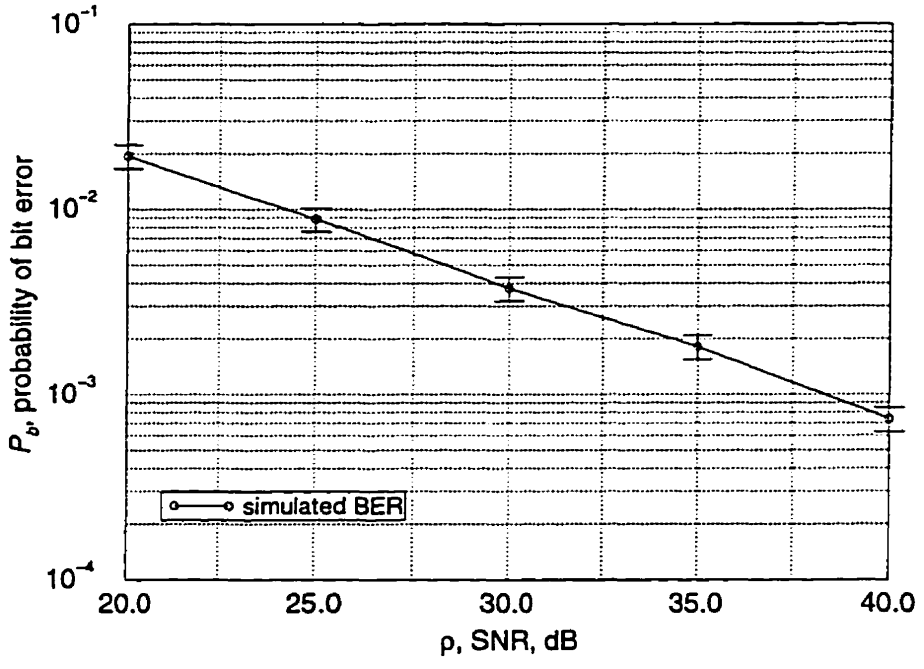

Figure 3.32: BER as a function of SNR in Case 2B channel.

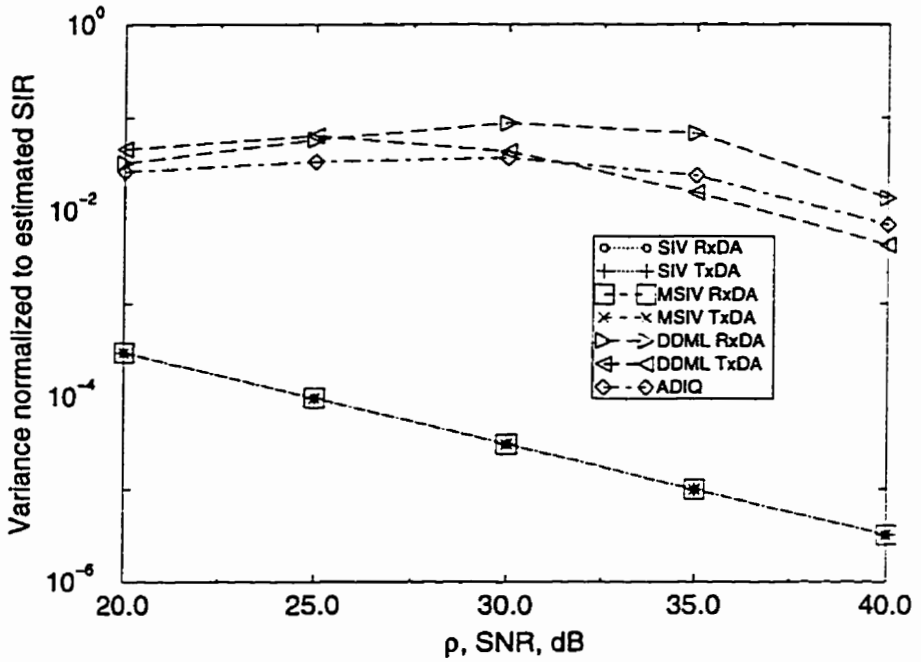

Figure 3.33: Normalized variance of SIR estimates generated by SIV, MSIV, DDML, and ADIQ estimation in Case 2B channel.

are portrayed in Figure 3.34.

From Figure 3.33, it is apparent that the variances of the DDML and ADIQ SIR estimates are about an order of magnitude larger than the variances of the SIV and MSIV SIR estimates, which is consistent with the relative performance results observed in Figure 3.30 for the Case 2A channel. The variances displayed in Figure 3.33 are due exclusively to the inherent performances of the SIR estimators and are not affected by time-evolving channel conditions as in Figure 3.30 which is the reason why, in an absolute sense, the variances seen in Figure 3.30 are much larger than those seen in Figure 3.33. It is not clear why, in a relative sense, the variances of the DDML and ADIQ SIR estimates are so much larger than the variances of the SIV

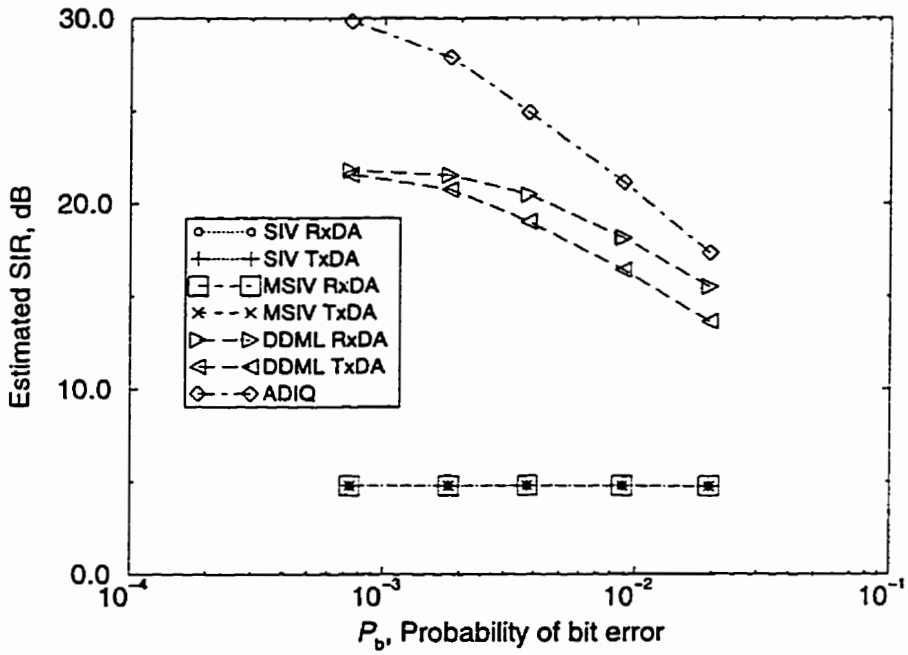

Figure 3.34: Correlation between mean estimated SIR and BER in Case 2B channel for SIV, MSIV, DDML, and ADIQ estimation.

and MSIV SIR estimates.

It should also be noted that, from Figure 3.33, the variances of the TxDA and RxDA forms of the SIV and MSIV estimators are close to identical even though the BER is quite high. This is a somewhat surprising result since in Figure 3.23 it is seen that the variances of the RxDA forms of the SIR estimators deviate from the variances of the TxDA forms at low SNR (high BER) in the Case 1 channel. In contrast, the variances of the RxDA and TxDA forms of the DDML estimator are different.

With respect to the correlation between SIR and BER, it is apparent from Figure 3.34 that the DDML and ADIQ SIR estimates reflect (in a relative sense) the BER of the channel quite well whereas the SIV and MSIV SIR estimates are flat throughout the range of tested BER. This observation indicates that the DDML and ADIQ SIR estimators are better SIR metrics in the context of the Case 2B channel from the point of view of SIR/BER correlation. Also, note that the DDML and ADIQ SIR estimates are much larger relative to the SIV and MSIV SIR estimates. The reason for these differences was attributed earlier in the discussion of the Case 2A channel to the fact that both the DDML and ADIQ estimators ignore the fluctuating amplitudes of the post-DD samples, and thus give optimistic estimates of the SIR.

3.9.3 Case 3: Slow, Frequency-Selective Fading Channel

The simulated BER of the Case 3 channel is plotted in Figure 3.35 as a function of the delay spread, τ_{rms}. No theoretical BER curve is plotted for reference as none could be found in the published literature corresponding precisely to the characteristics of the Case 3 channel. Recall that the SNR of the channel is set to 15 dB to increase the BER and shorten the simulation time since the BER due to multipath over the range of delay spreads tested can be very low. The optimum sampling instant in frequency-selective channels often does not correspond to the sampling instant that is optimum in the AWGN channel [123]. In this frequency-selective Case 3 channel,

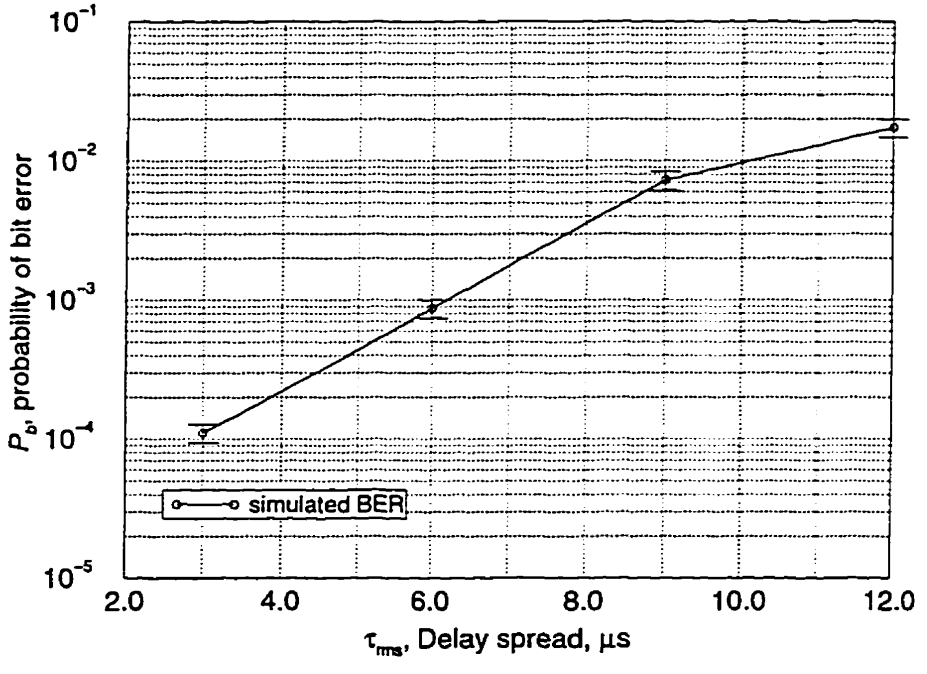

Figure 3.35: BER as a function of delay spread in Case 3 channel.

however, the optimum sampling instant is actually the same as that in the AWGN channel to within 1/16 of a symbol period (the resolution of the simulator).

The SIR estimators tested in this channel include the SVR, M_2M_4, PML RxDA, PML TxDA, SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ SIR estimators. The variances of the SIR estimates in the Case 3 channel are shown in Figure 3.36. The results are split between two graphs to avoid clutter, but the SIV variance curves are common to both graphs to facilitate comparison.

The behaviours of the SIR estimator variances in this Case 3 channel appear very similar to the behaviours of the SIR variances in the Case 1 channel, shown in Figure 3.23, except that the curves are reversed in the sense that increasing abscissa corresponds to increasing BER in Figure 3.36, while increasing abscissa corresponds to decreasing BER in Figure 3.23. Just as the variances of the RxDA SIR estimators in the Case 1 channel "droop" below the variances of their TxDA counterparts for low values of SNR (corresponding to high BER), the variances of the RxDA SIR estimators in the Case 3 channel droop at values of large τ_{rms} (again corresponding to high BER). Counter-intuitively, the smaller (drooping) variances of the RxDA estimates do not imply better estimator performances. For example, it can be seen in Figures 3.21 to 3.26 of Case 1 that a drooping variance at low SNR is associated with a growing estimator bias and a growing MSE. Thus it can be deduced that whenever the variance of an RxDA estimator droops below the variance of its TxDA counterpart, the reason is not because the performance of the RxDA estimator is superior to that of the TxDA estimator, but rather because the RxDA estimator is manifesting a bias which implies *inferior* performance. This deduction is useful in cases such as this Case 3 channel where bias and MSE curves are not available due to the difficulty of identifying the true SIR of the channel.

The TxDA forms of the DA SIR estimators perform best overall. Among the TxDA estimators, the SIV TxDA and MSIV TxDA estimators perform best, followed

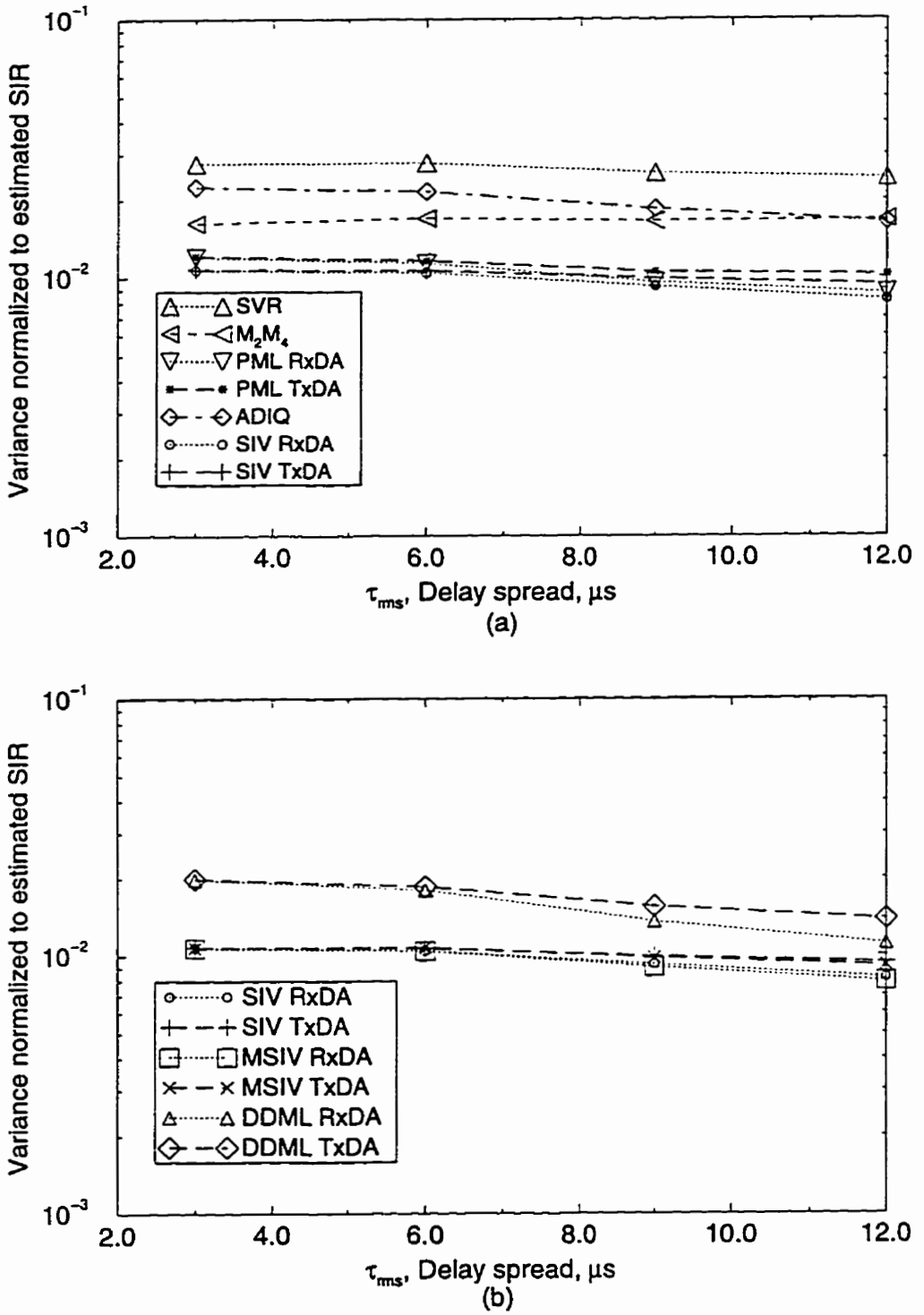

Figure 3.36: Normalized variance of SIR estimates generated by (a) SVR, M_2M_4, PML, ADIQ, and SIV estimation, (b) SIV, MSIV, and DDML estimation in Case 3 channel.

by the PML TxDA estimator and the DDML TxDA estimator. The variances of the SVR and M_2M_4 estimators (which are not DA estimators) are relatively unaffected by changing τ_{rms}.

The relationships between the mean estimated SIR and the BER of the Case 3 channel for the various estimators are portrayed in Figure 3.37. This plot shows more differentiation among the various estimators than is observed in Figure 3.27 of Case 1, but all estimators perform similarly with the exception of the SVR estimator whose SIR/BER curve is distinctly set apart from the main group.

3.9.4 Case 4: Single Cochannel Interferer

The simulated BER of the Case 4 channel is plotted in Figure 3.38 as a function of the ratio of desired signal power to CCI power, S/C. No theoretical BER curve is plotted for reference as none could be found in the published literature corresponding precisely to the characteristics of the Case 4 channel (all BER curves plotted as a function of CCI power found in the literature assume a flat, slowly fading channel, whereas the Case 4 channel assumes simply AWGN and CCI, and no fading at all—slow or fast). Recall that the SNR of the channel is set to 15 dB to increase the BER and shorten the simulation time since the BER due to CCI alone over most of the range of S/C tested can be very low up to a certain threshold and then degrade sharply.

The SIR estimators tested in this channel include the SVR, M_2M_4, PML RxDA, PML TxDA, SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ estimators. The biases, variances, and MSE's of these SIR estimates in the Case 4 channel are shown in Figures 3.39, 3.40, and 3.41. The biases and variances are normalized to the estimated SIR, and the MSE's are normalized to the true SIR which is given simply in this case by $S/(N+C)$.

The relative ranking of the estimators on the basis of the bias, variance and MSE results is consistent with the rankings of the estimators in Cases 1 and 3. The TxDA

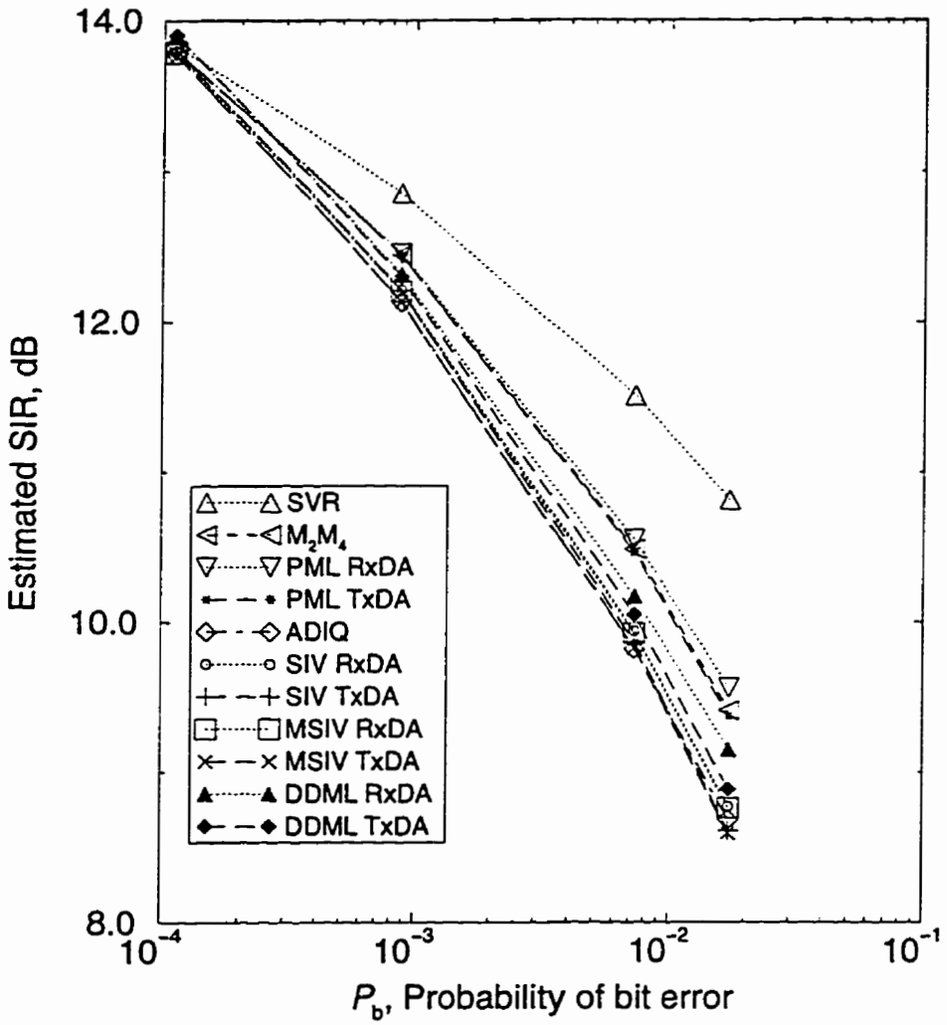

Figure 3.37: Correlation between mean estimated SIR and BER in Case 3 channel for SVR, M_2M_4, PML, ADIQ, SIV, MSIV, and DDML estimation.

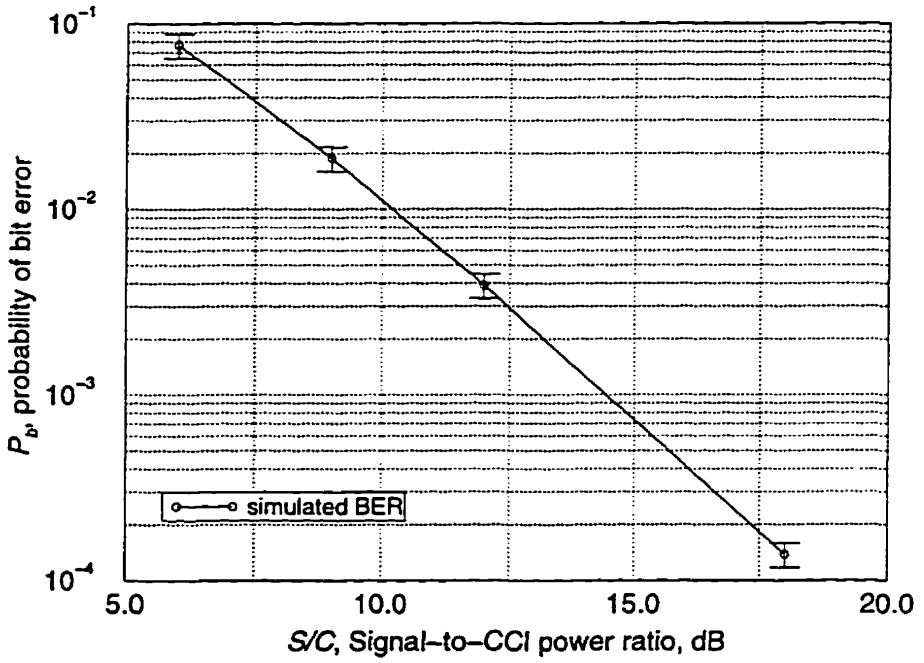

Figure 3.38: BER as a function of desired signal power to CCI power in Case 4 channel.

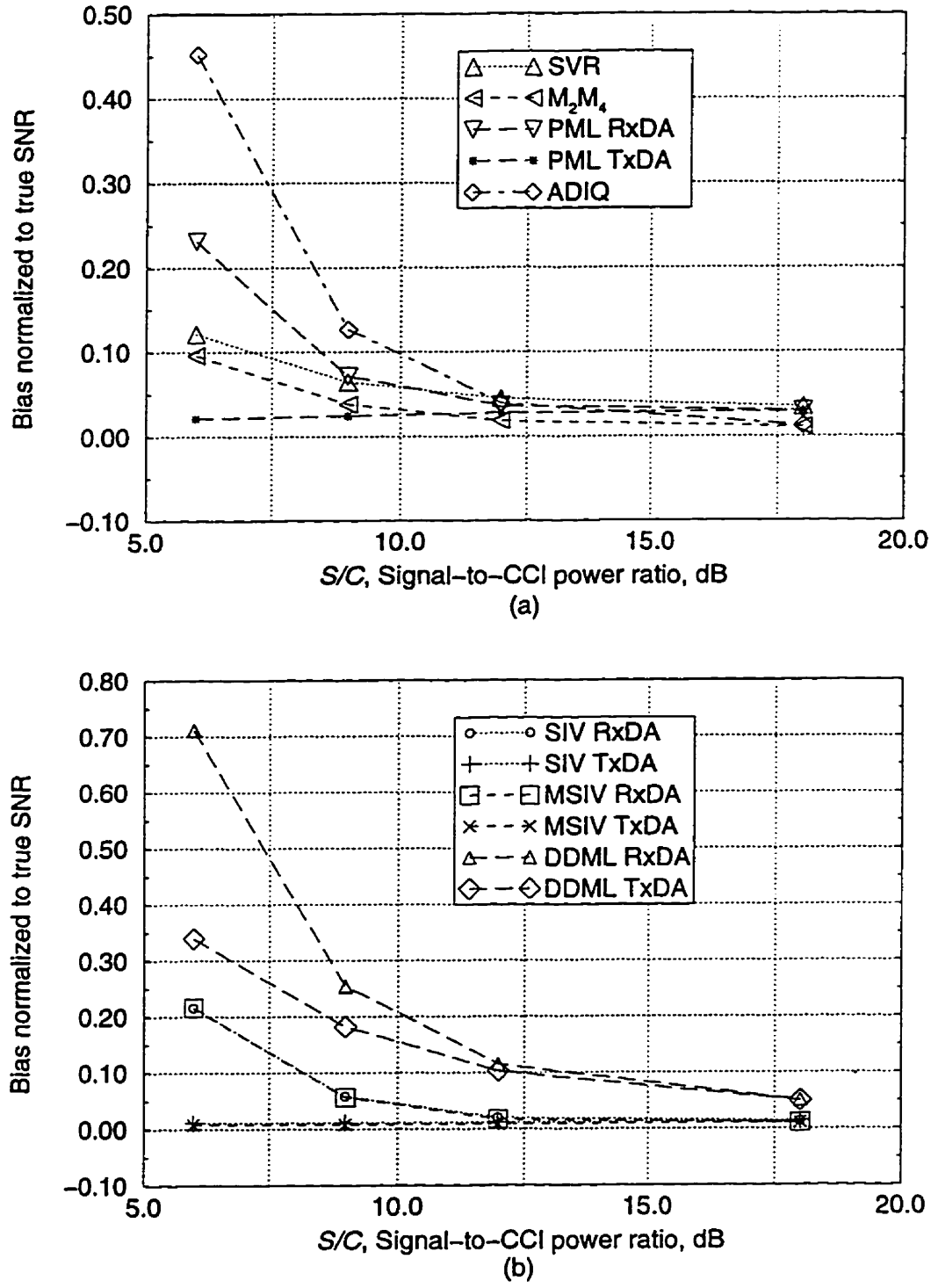

Figure 3.39: Normalized bias of SIR estimates generated by (a) SVR, M_2M_4, PML, and ADIQ estimation, (b) SIV, MSIV, and DDML estimation in Case 4 channel.

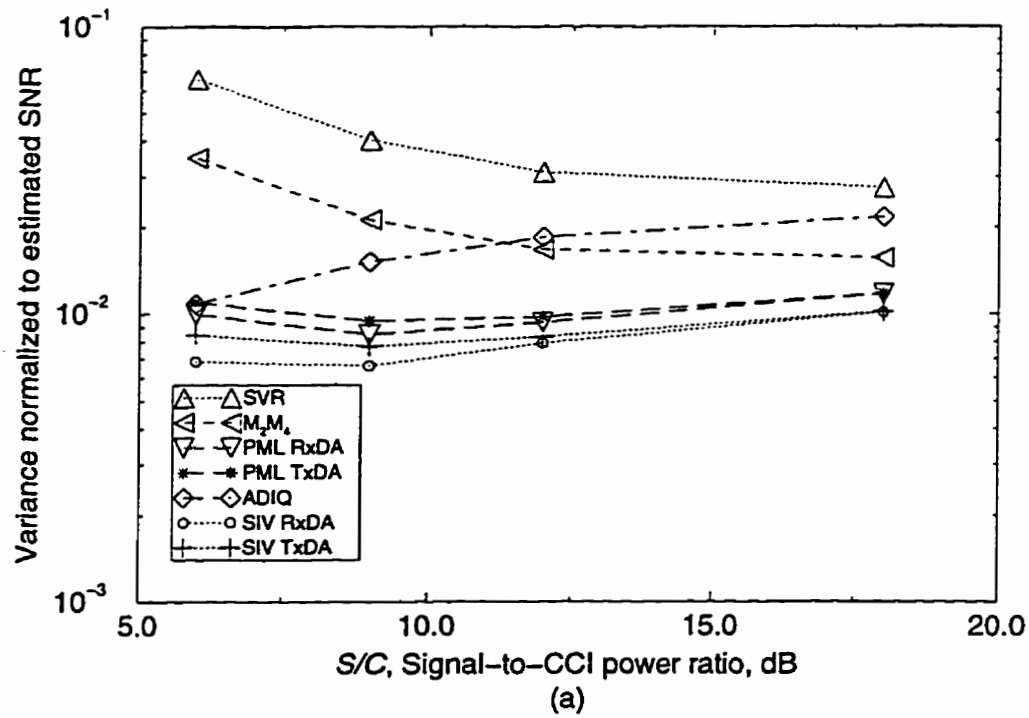

Figure 3.40: Normalized variance of SIR estimates generated by (a) SVR, M_2M_4, PML, ADIQ, and SIV estimation, (b) SIV, MSIV, and DDML estimation in Case 4 channel.

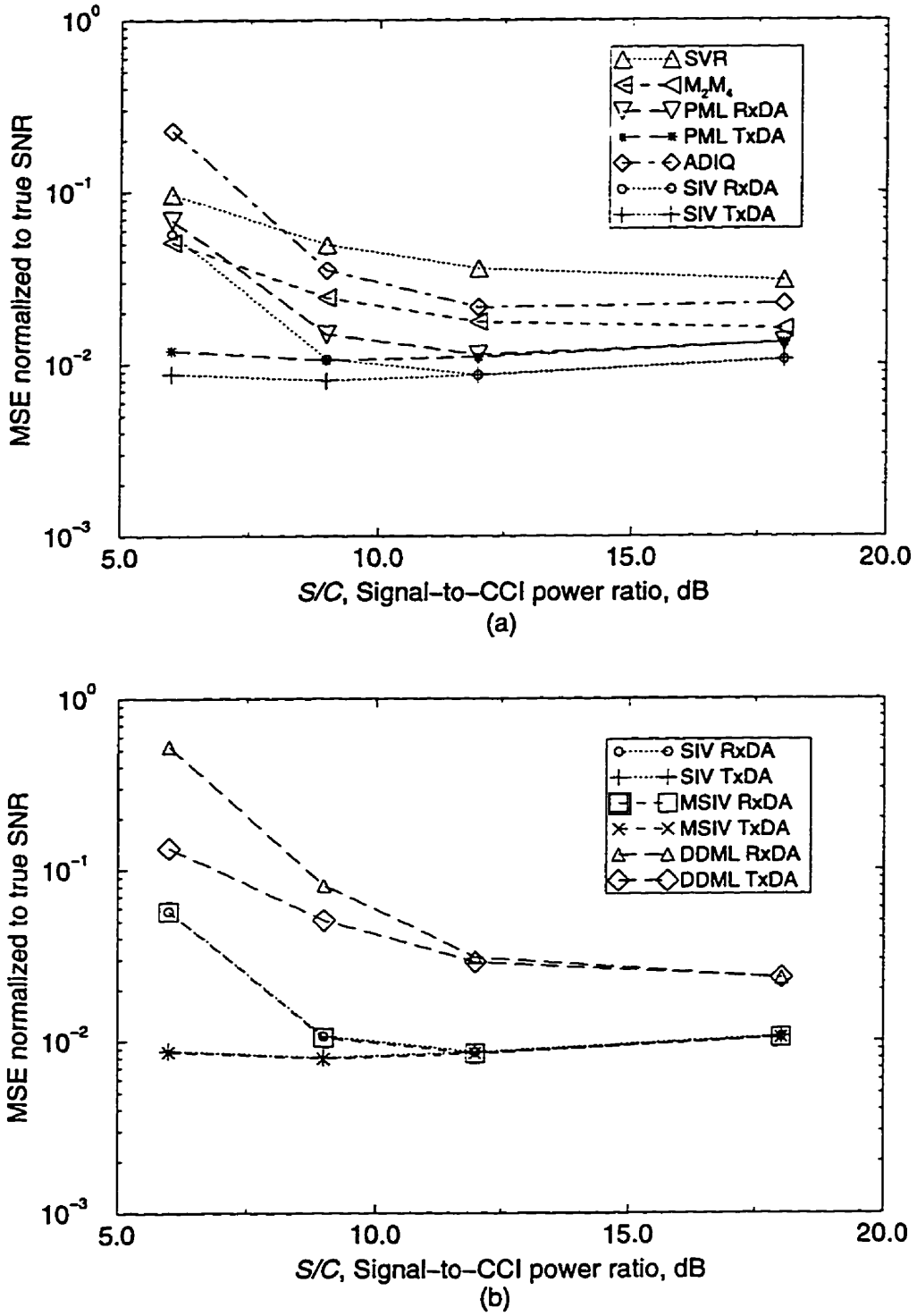

Figure 3.41: Normalized MSE of SIR estimates generated by (a) SVR, M_2M_4, PML, ADIQ, and SIV estimation, (b) SIV, MSIV, and DDML estimation in Case 4 channel.

forms of the SIV and MSIV estimators generate SIR estimates with the smallest overall biases, variances and MSE's, while the performances of the RxDA forms deviate slightly from those of their TxDA counterparts at low S/C. Specifically, the biases and MSE's of the RxDA SIR estimates become larger, and the variances of the RxDA SIR estimates become smaller relative to their TxDA counterparts at low S/C. These deviations are consistent with the comments made in Section 3.9.3 regarding the correlation between large RxDA estimator bias and MSE, and the droop in the variance of RxDA estimates below the variance of corresponding TxDA estimates. In addition note that, once again, the variances of the SIV RxDA and MSIV RxDA estimators and the variances of the SIV TxDA and MSIV TxDA estimators are essentially identical.

The performances of the PML RxDA and PML TxDA estimators are inferior to those of the SIV and MSIV estimators, but they are better than those of the SVR, M_2M_4, DDML RxDA, DDML TxDA, and ADIQ estimators over most of the range of S/C tested. The significant droop of the variance of the DDML RxDA estimates at low values of S/C is associated with the large bias seen in Figure 3.39(b).

The ADIQ estimator also exhibits a droop in the variance of its SIR estimates as it, too, is an RxDA type of estimator. Though no results of its TxDA form are presented for comparison, it is apparent from Figures 3.39(a) and 3.41(a) that a large bias and MSE is associated with the droop in variance.

The relationships between the mean estimated SIR and the BER of the Case 4 channel for the various estimators are portrayed in Figure 3.42. Similar to the Case 1 channel, there is little to distinguish one estimator from another in terms of SIR/BER correlation. In general, the SIR/BER correlation is good for all of the estimators tested.

3.9.5 Case 5: A "Typical" Mobile Radio Channel

The simulated BER of the Case 5 channel is plotted in Figure 3.43 as a function of the SNR. No theoretical BER curve is plotted for reference as there is none in

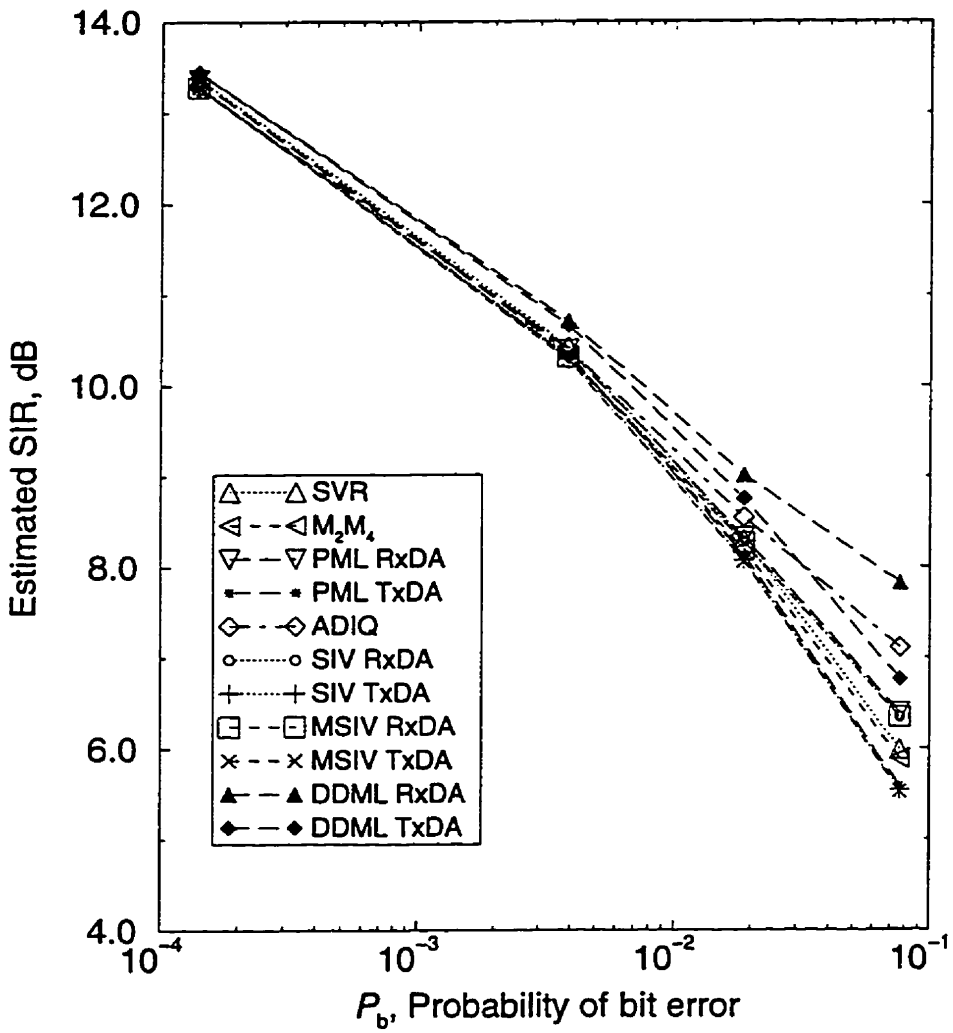

Figure 3.42: Correlation between mean estimated SIR and BER in Case 4 channel for SVR, M_2M_4, PML, ADIQ, SIV, MSIV, and DDML estimation.

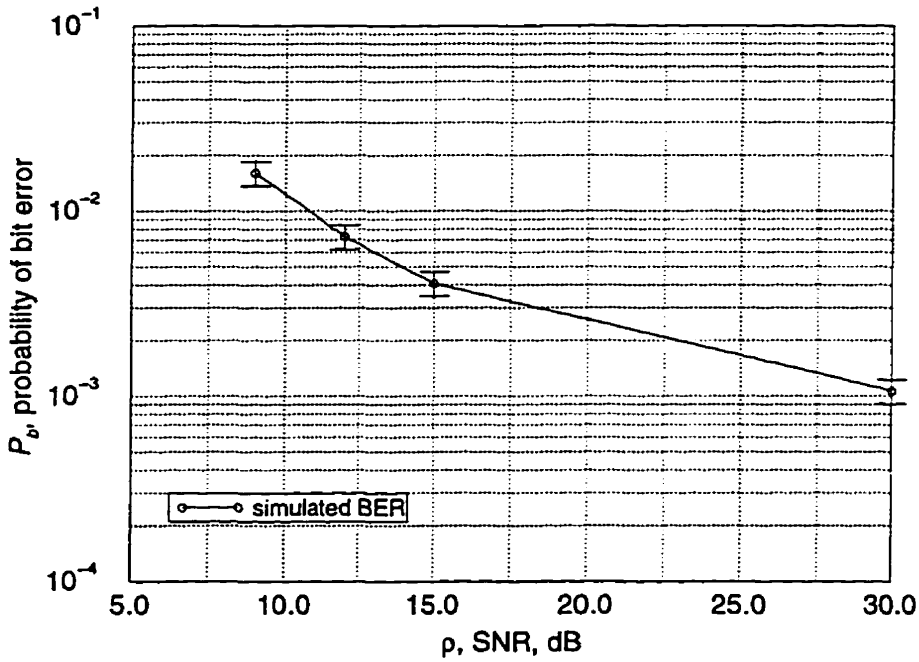

Figure 3.43: BER as a function of SNR in Case 5 channel.

the published literature corresponding precisely to the characteristics of the Case 5 channel. Recall that this channel comprises a number of impairments including a Doppler spread of 100 Hz, a delay spread of 3 μs, and a desired signal-to-CCI power ratio of 18 dB. The SNR is swept over the range $\{9, 12, 15, 30\}$ dB. The amplitude and phase fluctuations due to the Rayleigh fading are replayed for each successive SIR estimate similarly to the Case 2B channel. As mentioned in Section 3.9.3, for optimum BER performance in a frequency-selective channel it is sometimes necessary to use a different sampling instant than that used in an AWGN channel. This modification is unnecessary for the Case 5 channel, as it was also unnecessary for the Case 3 channel.

The SIR estimators tested in this channel include the SVR, M_2M_4, PML RxDA, PML TxDA, SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ estimators. All of the estimators generate "meaningful" estimates, even though Doppler spread is present in this Case 5 channel. Recall from Section 3.9.2 that the SVR, M_2M_4, PML RxDA, and PML TxDA estimators were described as generating "bad" estimates in Rayleigh-faded channels when a deep null occurs. The time index of the Case 5 channel is purposely designed so that no deep fade occurs over the interval chosen to be replayed for each successive SIR estimate (see Figure 3.13 of Section 3.6.2) so that "bad" estimates are avoided. In reality, channel conditions evolve over time and deep nulls are unavoidable in Rayleigh fading so that the SVR, M_2M_4, PML RxDA, and PML TxDA SIR estimators are unsuitable for "real-world" Rayleigh fading channels, but it is still interesting to investigate their performances in this specific, albeit artificial, composite channel.

The variances of the SIR estimates in the Case 5 channel are shown in Figure 3.44 in which it may be seen that there are two distinct groups of estimators: one comprises the SIV, MSIV and M_2M_4 estimators, the other comprises the SVR, DDML, ADIQ, and PML estimators. The distinction between these two groups is the general trends of the variances of the SIR estimates. In the former, the variances slope downward with increasing SNR while, in the latter, the variances are relatively flat throughout

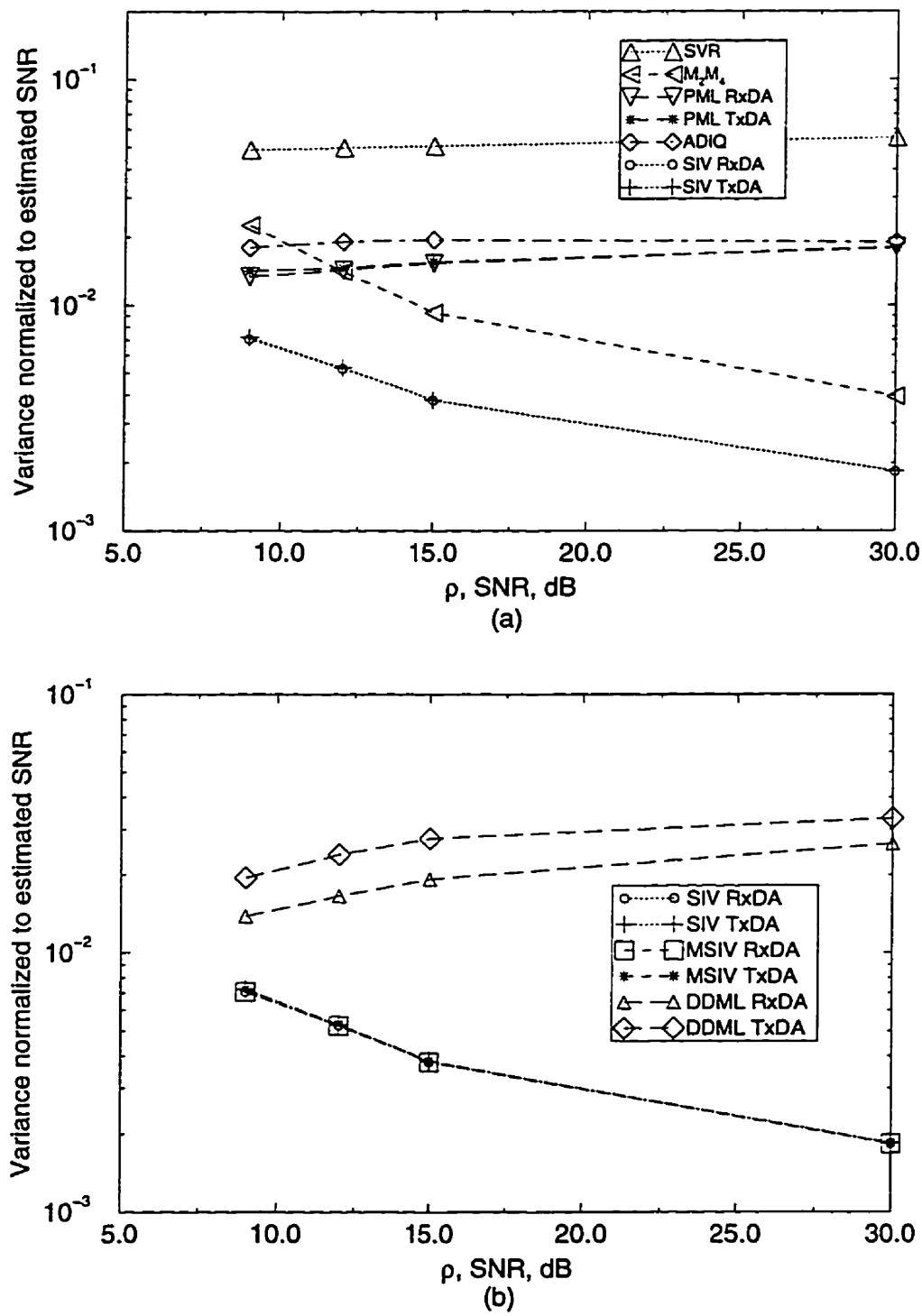

Figure 3.44: Normalized variance of SIR estimates generated by (a) SVR, M_2M_4, PML, ADIQ, and SIV estimation, (b) SIV, MSIV, and DDML estimation in Case 5 channel.

the range of tested SNR, or trending very slightly upward.

As in all of the other cases, the SIV and MSIV estimators exhibit the smallest variances. Note the the TxDA and RxDA forms of these estimators perform almost identically, even at an SNR of 9 dB where the BER is greater than 10^{-2}. This is a curious result since it seems reasonable to assume that receiver errors should degrade the performances of RxDA estimators relative to the performances of their TxDA counterparts. In fact, the variances of the RxDA estimators in Cases 1, 3, and 4 *are* degraded (see Figures 3.23, 3.36, and 3.40, respectively) which supports the assumption. The only other exception is Case 2. A common feature of the Case 2 and Case 5 channels is the presence of Rayleigh fading which suggests that one possible reason why receiver errors do not cause the variances of the RxDA estimates to diverge from the variances of the TxDA estimates is that the variances of the estimates are so large due to the Rayleigh fading that the effect of receiver errors is negligible.

In contrast, there is a noticeable difference between the variances of the RxDA and TxDA PML estimates, and a significant difference between the variances of the RxDA and TxDA DDML estimates. The gap between the variances of the DDML RxDA and DDML TxDA estimates throughout the range of tested SNR is particularly curious since the RxDA estimates have the smaller variance. From the discussion of Section 3.9.3, this drooping of the DDML RxDA estimates with respect to the TxDA estimates would seem to indicate that there is a strong associated bias throughout the tested range of SNR. One could then deduce that the DDML RxDA estimator is more sensitive to receiver errors than the other RxDA estimators in this type of channel.

The M_2M_4 SIR estimator is associated with the SIV and MSIV estimators in the sense that the variance of its SIR estimates has a similar downward slope with increasing SNR. The variance of the M_2M_4 estimates is about twice that of the SIV and MSIV estimates, but it is superior to the variances of the estimates of the other estimators over most of the range of tested SNR, with the exception of very low values

of SNR.

Among the estimators whose variances are relatively flat over the range of tested SNR, the variances of the PML RxDA and PML TxDA estimators are the smallest and are almost identical over the entire range of tested SNR. The performances of the ADIQ, DDML RxDA and DDML TxDA estimators are only slightly worse while the SVR estimator exhibits the largest variance of all SIR estimators tested.

The relationships between the mean estimated SIR and the BER of the Case 5 channel for the various estimators are portrayed in Figure 3.45. The two distinct groups of SIR estimators observed in Figure 3.44 also exist in Figure 3.45. The SIV, MSIV, and M_2M_4 estimators exhibit values of SIR that are smaller in magnitude and span a smaller range than the values of SIR exhibited by the SVR, DDML, ADIQ, and PML estimators so that, on the basis of Figure 3.45 alone, the latter group of estimators would be expected to perform better in applications than the former. However, the SIR/BER correlations are viewed from a different perspective in the next section which makes the relative efficacy of each estimator more clear.

3.9.6 Composite Plots of SIR/BER Correlation for each Estimator

It is informative to view the SIR/BER correlations from another perspective by plotting the results for all five test channels for each of the SIR estimators separately. Figures 3.46 to 3.56 portray these composite plots for the SVR, M_2M_4, PML RxDA, PML TxDA, SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ estimators. Of course, results from the Case 2A and Case 2B channels are not included for the SVR, M_2M_4, and PML estimators since, as mentioned earlier, these estimators do not produce meaningful SIR estimates under the channel conditions of Case 2A or Case 2B. Error bars are attached to each point, and the sizes of the error bars are determined by the square root of the estimated variance associated with each estimated SIR mean.

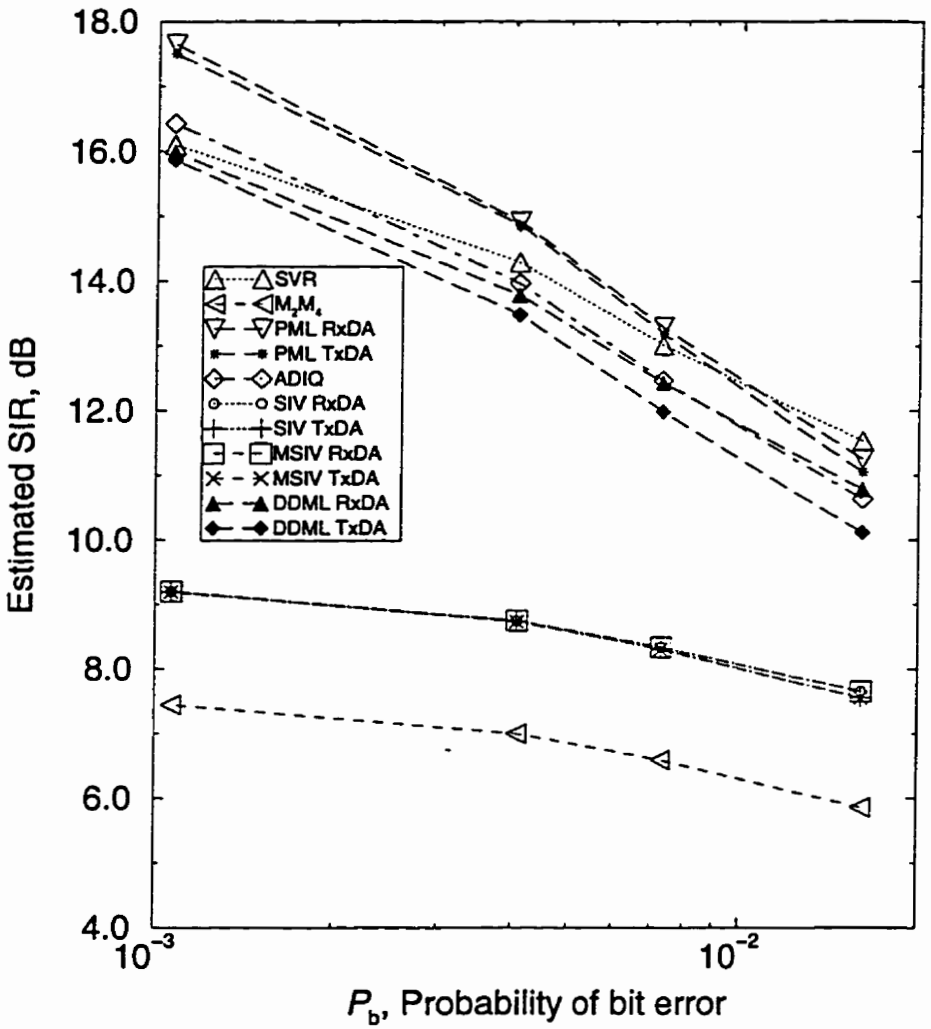

Figure 3.45: Correlation between mean estimated SIR and BER in Case 5 channel for SVR, M_2M_4, PML, ADIQ, SIV, MSIV, and DDML estimation.

In these plots, a good SIR/BER correlation is indicated by clusters of points grouped tightly together with some reasonably steep, monotonic slope. Ideally, the perfect SIR/BER correlation would be reflected by a plot of all SIR/BER points falling on the same curve; that is, the perfect SIR/BER correlation would be a one-to-one correspondence of the SIR to the BER for all of the test channels considered. In addition to one-to-one correspondence, good correlation also requires SIR estimates with a reasonably small variance.

The SIR/BER correlations for Cases 1, 3, and 4, are similarly good for all of the estimators under consideration. The main differences in the correlations arise in the context of Cases 2 and 5 where Rayleigh fading is present.

One of the criteria mentioned above for good SIR/BER correlation is that the clusters of SIR/BER points be grouped tightly together. From Figures 3.50 to 3.53, it may be seen that the SIV and MSIV SIR estimates satisfy this criterion. Though the error bars on the points for Case 2A are large, this is mainly due to the time-varying channel conditions rather than the inherent variance of the estimator. Unfortunately, it is difficult to separate the effect of the channel from the effect of the SIR algorithm itself on the variance, but the Case 2B model helps to accomplish this separation of effects by replaying the same amplitude and phase fluctuations for each successive SIR estimate, as described earlier, thus maintaining the true SIR constant (whatever its value may be). The SIR/BER correlation points for the Case 2B model are quite flat for the SIV and MSIV estimators and the error bars are small which indicates that the contribution of the SIV and MSIV algorithms to the variance is small, but the flatness of the plots indicates that the SIR/BER correlation is very poor *for the specific channel conditions of Case 2B* which are artificial.

In contrast, the SIR/BER points of the DDML and ADIQ SIR estimates (see Figures 3.54, 3.55 and 3.56) for Case 2B have a slope which reveals some correlation with the BER, but the points are located some distance away from the main cluster of SIR/BER points corresponding to Cases 1,3,4, and 5. The SIR/BER points of the

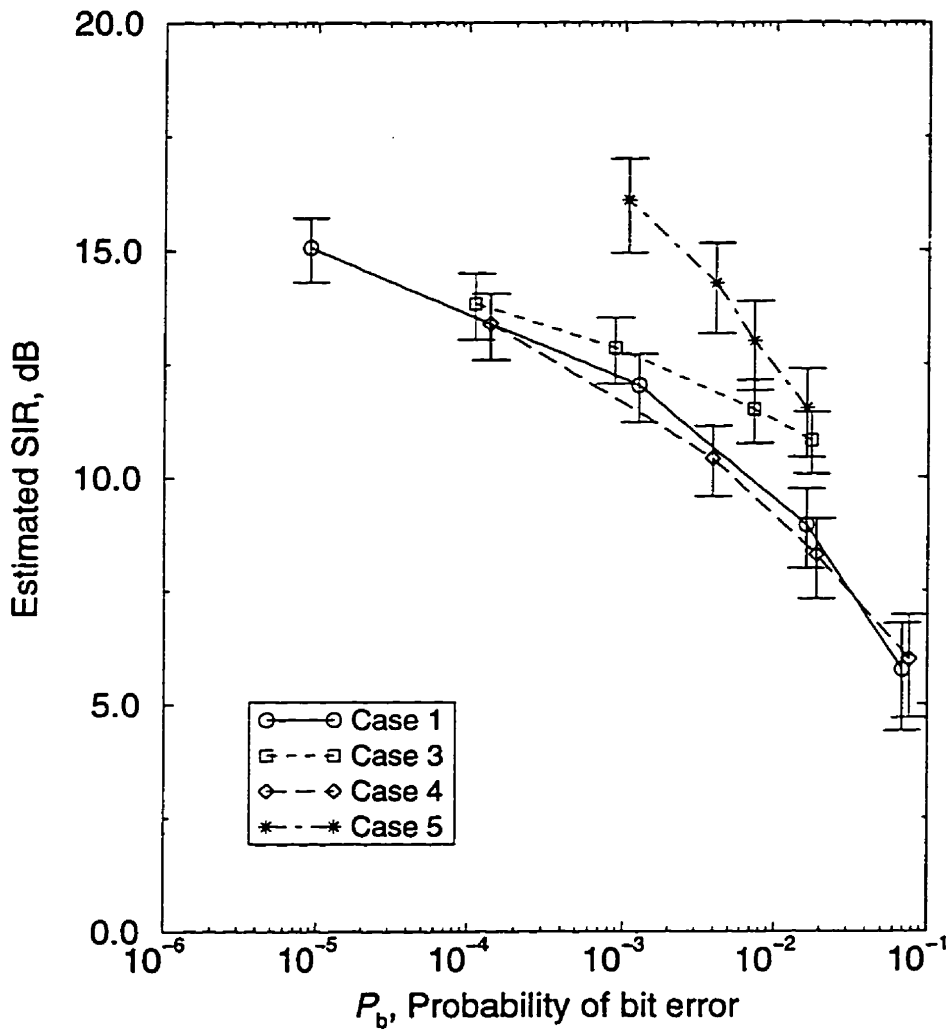

Figure 3.46: Composite plot of correlation between mean estimated SIR and BER for SVR estimation.

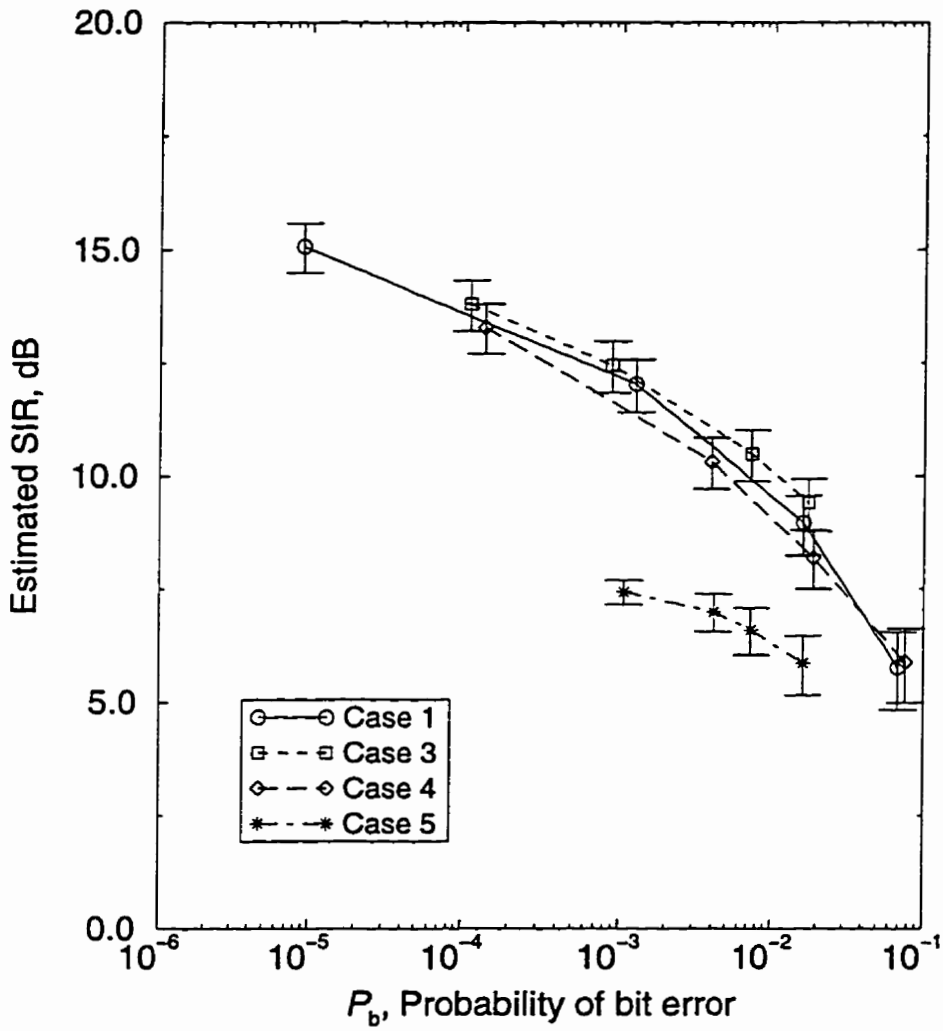

Figure 3.47: Composite plot of correlation between mean estimated SIR and BER for M_2M_4 estimation.

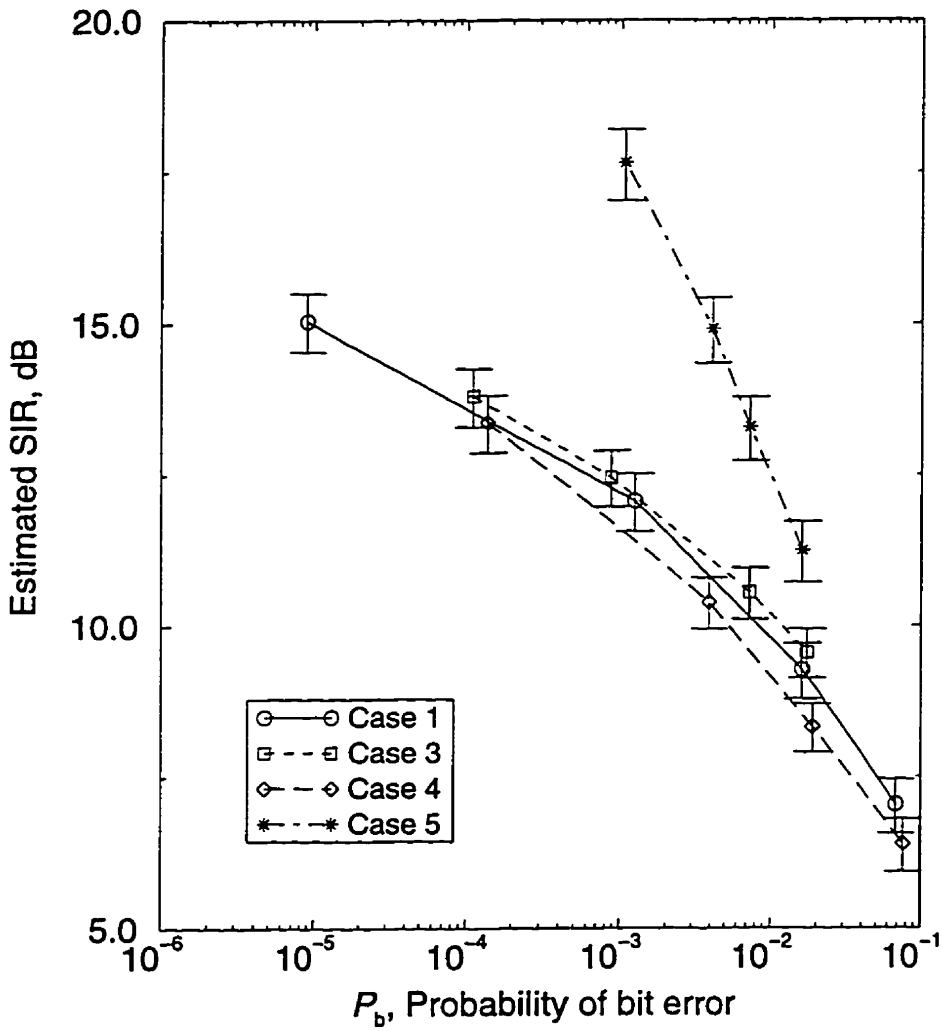

Figure 3.48: Composite plot of correlation between mean estimated SIR and BER for PML RxDA estimation.

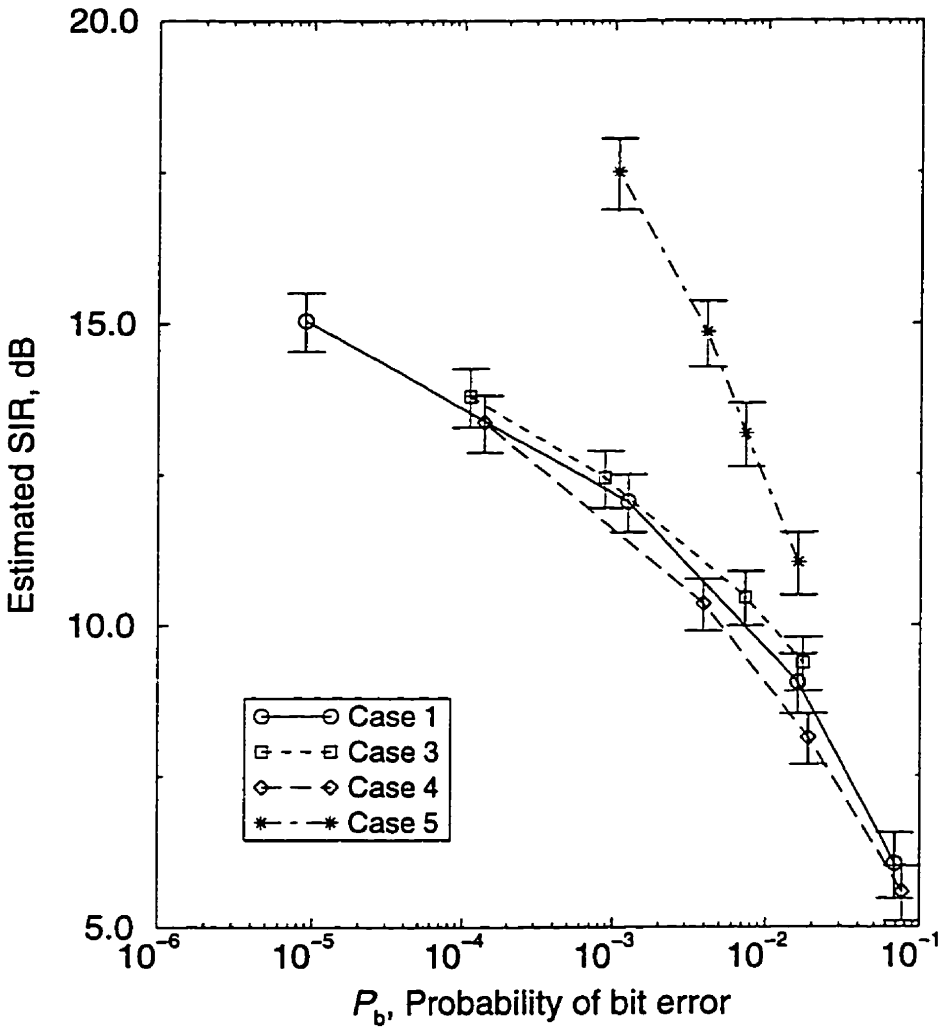

Figure 3.49: Composite plot of correlation between mean estimated SIR and BER for PML TxDA estimation.

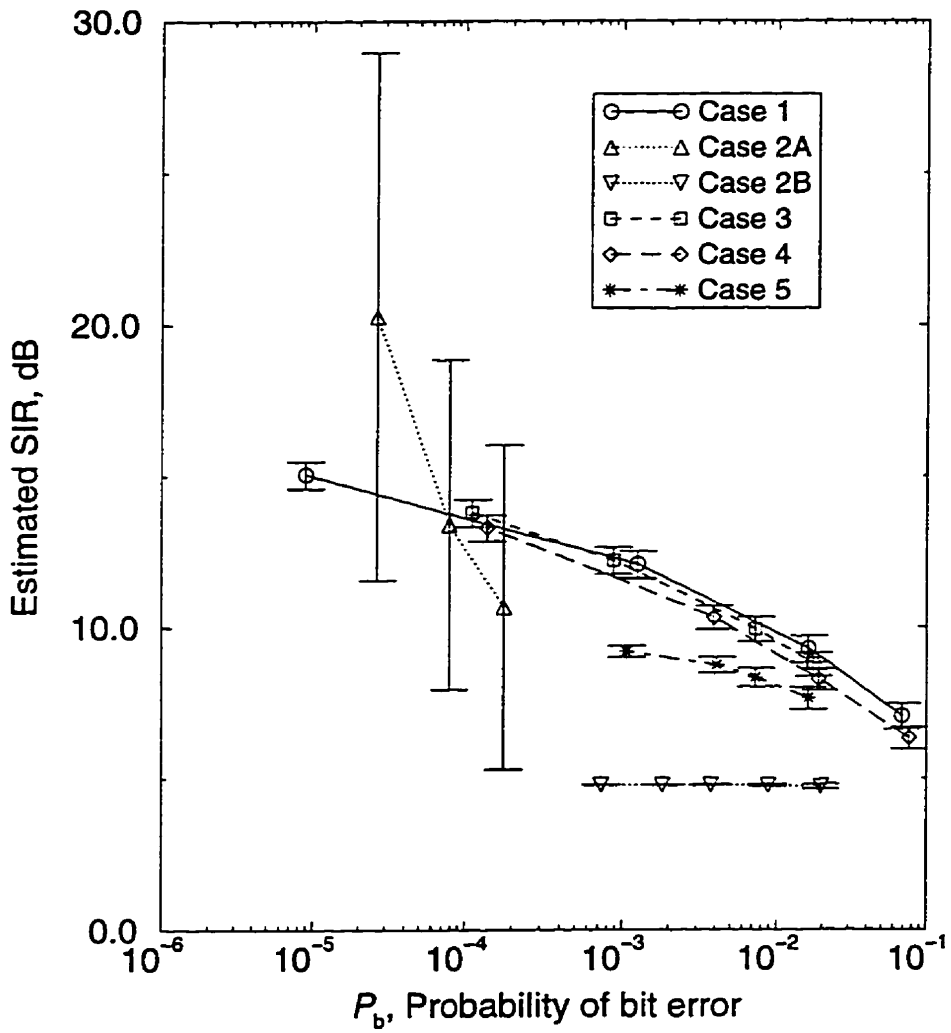

Figure 3.50: Composite plot of correlation between mean estimated SIR and BER for SIV RxDA estimation.

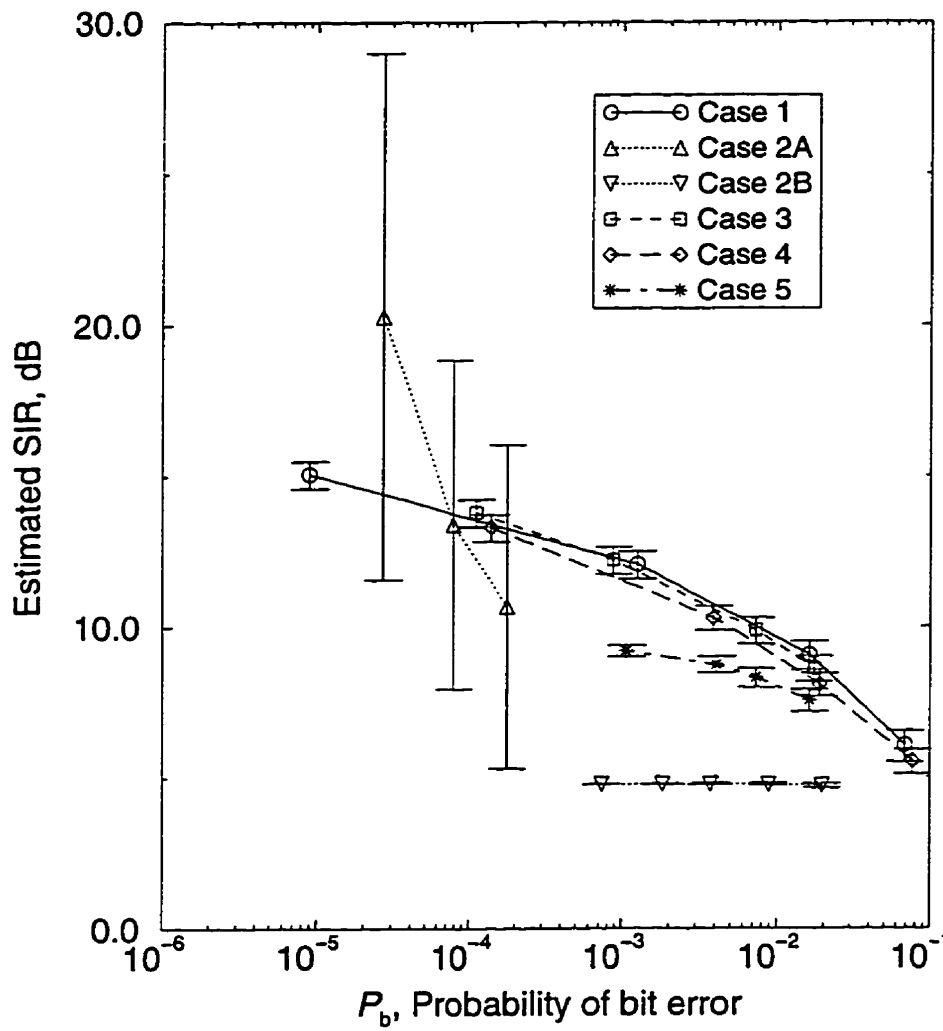

Figure 3.51: Composite plot of correlation between mean estimated SIR and BER for SIV TxDA estimation.

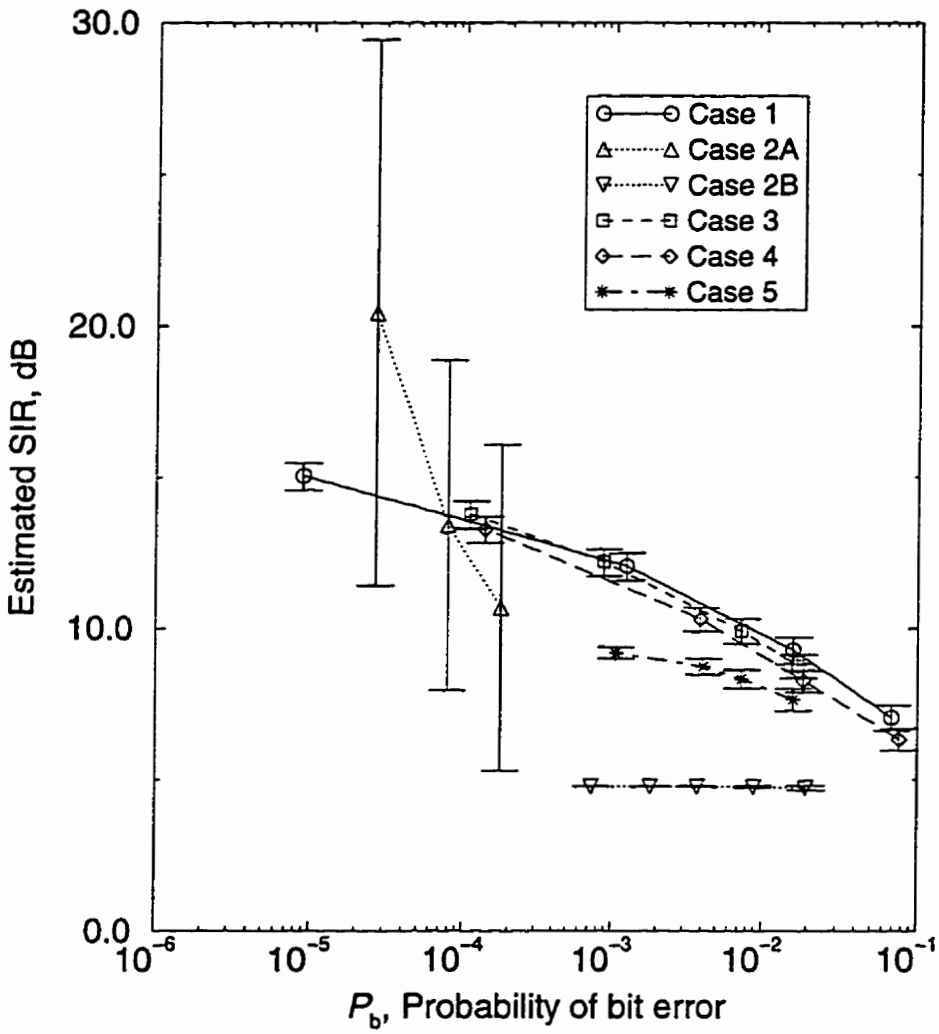

Figure 3.52: Composite plot of correlation between mean estimated SIR and BER for MSIV RxDA estimation.

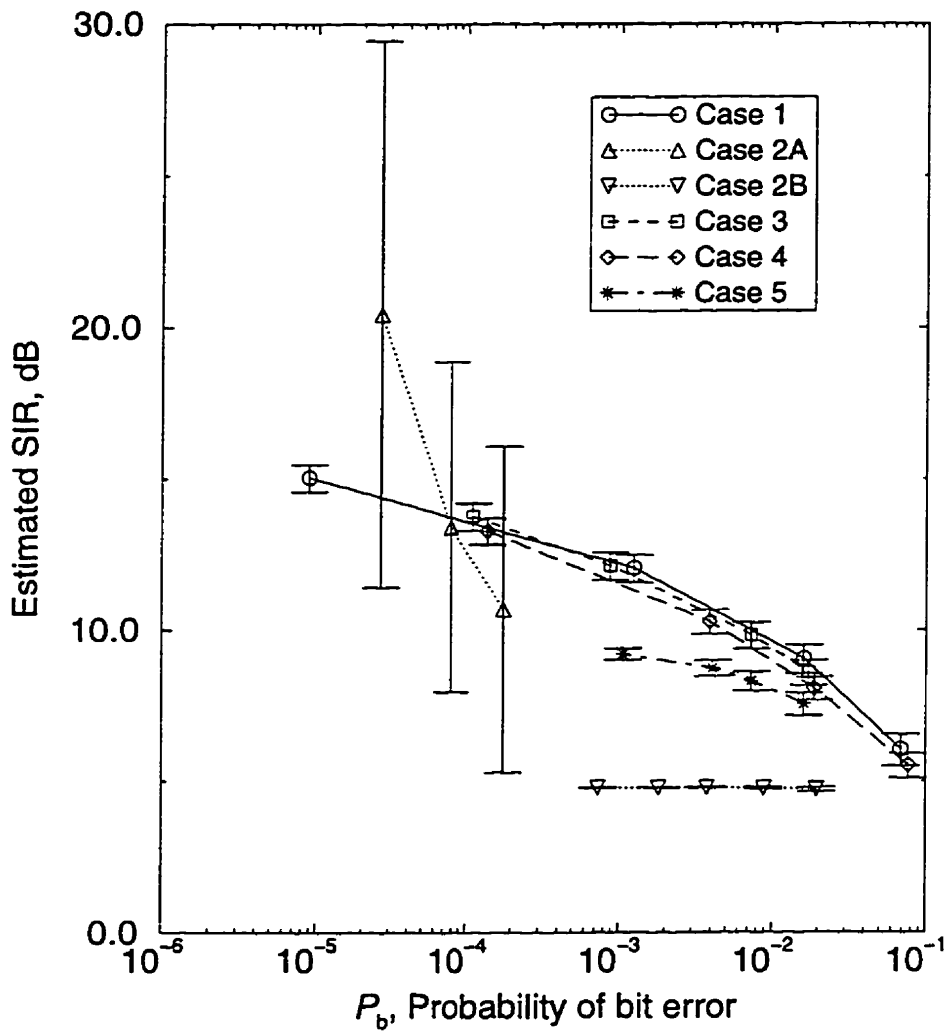

Figure 3.53: Composite plot of correlation between mean estimated SIR and BER for MSIV TxDA estimation.

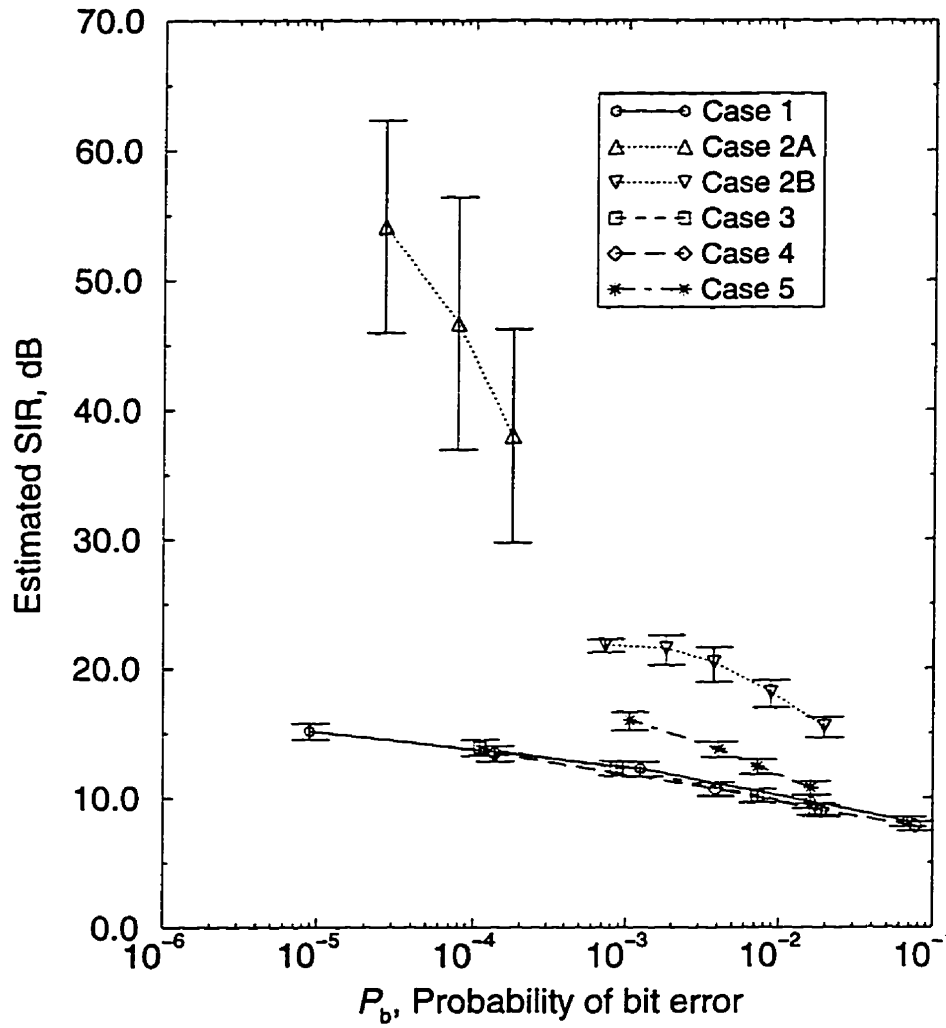

Figure 3.54: Composite plot of correlation between mean estimated SIR and BER for DDML RxDA estimation.

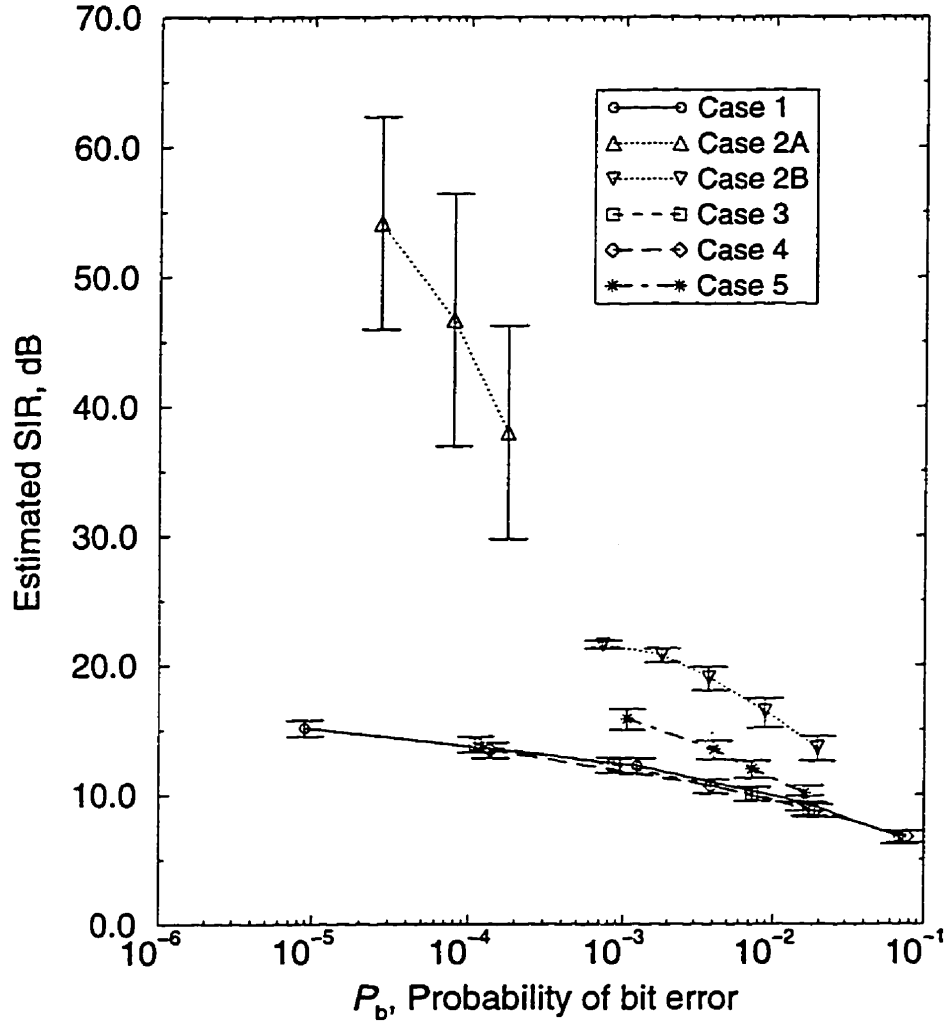

Figure 3.55: Composite plot of correlation between mean estimated SIR and BER for DDML TxDA estimation.

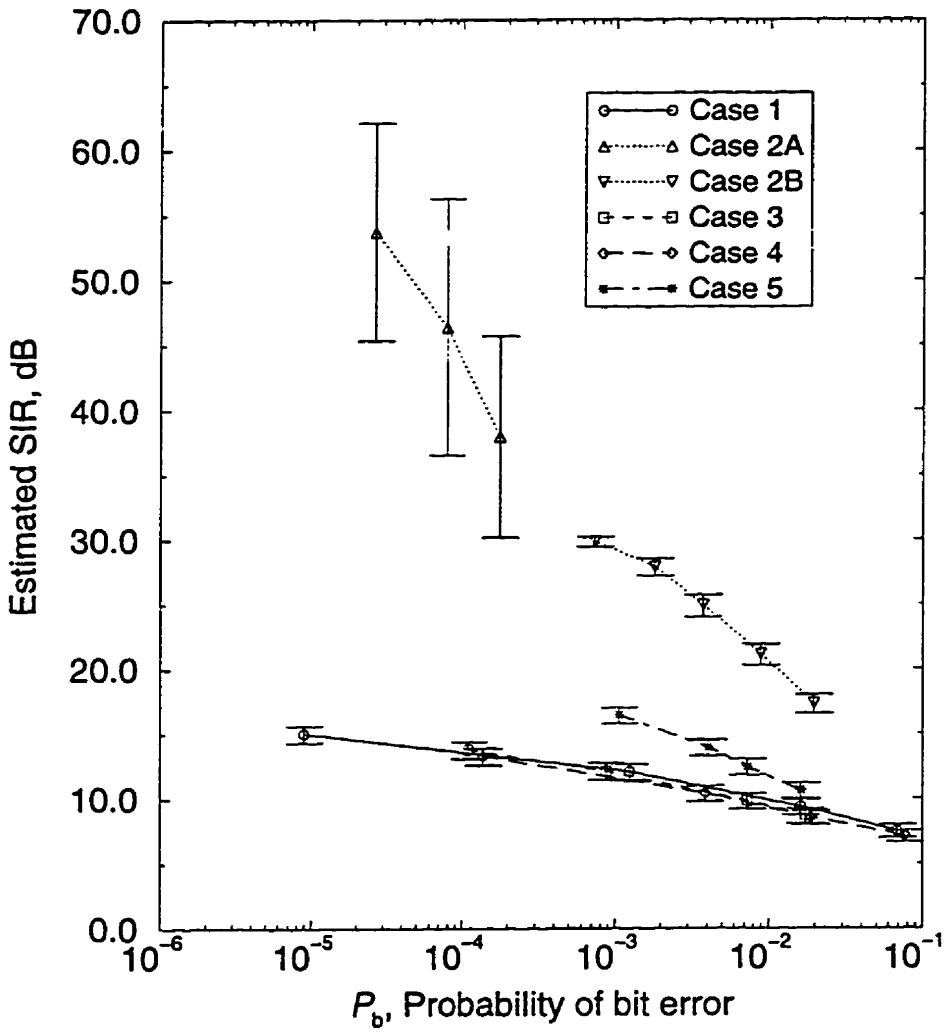

Figure 3.56: Composite plot of correlation between mean estimated SIR and BER for ADIQ estimation.

DDML and ADIQ SIR estimates for Case 2A are located even farther from the main cluster (by orders of magnitude). The error bars for the Case 2A points are large just as the error bars for the SIV and MSIV Case 2A SIR/BER points are large. Again, this large variance is mainly due to the random amplitude and phase fluctuations of the Rayleigh-faded channel.

The double grouping of SIR estimators identified in Section 3.9.5 in the context of the Case 5 results also manifests itself in Figures 3.46 to 3.56. The Case 5 SIR/BER points of the M_2M_4, SIV RxDA, SIV TxDA, MSIV RxDA, and MSIV TxDA SIR estimates fall below the main cluster of points in their respective plots, whereas the Case 5 SIR/BER points of the SVR, PML RxDA, PML TxDA, DDML RxDA, DDML TxDA, and ADIQ SIR estimates reside above the main cluster of points. The main effect of this difference is that SIR estimates made under frequency-selective conditions by the former group of estimators are slightly pessimistic, whereas SIR estimates made by the latter group of estimators are slightly optimistic.

3.10 Summary

Thirteen SIR estimators were described in Section 3.7: the SVR, M_2M_4, PML RxDA, PML TxDA, SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, ADIQ, ML RxDA, and ML TxDA SIR estimators. The PML, SIV, and MSIV estimators (both the RxDA and TxDA forms) are original. The performances of the SIR estimators were compared using $\pi/4$-DQPSK-modulated signals in the five different test channels described in Section 3.6.2.

An SIR estimator for wireless applications must be able to function well given any channel impairment including AWGN, CCI, and fast, frequency-selective fading. Of the thirteen estimators tested in one or more of the five test channels described in Section 3.6.2, only the SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ SIR estimators produce meaningful SIR estimates under all of the channel conditions tested.

Of these seven SIR estimators, none could be considered "perfect", but it seems reasonable to conclude from all of the plots presented in Sections 3.9.1 to 3.9.6 that the SIV and MSIV SIR estimators (both the RxDA and TxDA forms) exhibit the most favourable characteristics from the point of view of small variances and good SIR/BER correlations. The TxDA forms of the SIV and MSIV SIR estimators perform marginally better than their RxDA counterparts, as is most evident in plots of bias, variance, and MSE. There is very little to distinguish the performances of the SIV estimators from the performances of the MSIV estimators, and any difference is so small as to be statistically insignificant.

The DDML and ADIQ exhibit remarkably similar biases, variances, MSE's, and SIR/BER correlations so, based on these results, one would expect applications incorporating these estimators to perform similarly.

From the point of view of implementation complexity, the ADIQ SIR estimator is the simplest to implement of all of the estimators considered in Chapter 3 or Chapter 2, especially if the AGC operation could be simplified from an explicit division to a bit-shift operation, and if the correction formula given by (3.34) need not be applied, as is the case in applications which require only a *relative* measure of SIR.

If only a relative measure of SIR is required, the SIV and MSIV estimators may also be simplified, as it is evident from (3.75) or (3.100) that only the ratio

$$\frac{\hat{S}^2}{\hat{\sigma}_I^2 + \hat{\sigma}_Q^2}$$

need be computed instead of the full expression. Comparing the complexity of the SIV estimators to that of the MSIV estimators, the SIV estimators are easier to implement as the MSIV algorithm requires complex multiplication, whereas the SIV algorithm does not. Furthermore, the TxDA forms have the added complexity over the RxDA forms that known data must be identified and extracted from the received signal which is an issue of synchronization.

To implement the DDML estimators, the differential phase must be computed

(which could be a part of the detection hardware) and the sine function must somehow be implemented (perhaps as a lookup table). Alternatively, it is suggested in [28] that a reduced-complexity form of the DDML algorithm may be implemented to avoid the sine function, but at the cost of performance. Since the reduced-complexity form was not considered here, it is not possible to indicate the degree to which its performance could be expected to degrade.

Chapter 4

Application of SIR Estimation to Selection Diversity Combining

4.1 Motivation and Description of Application

The performances of various SIR estimators were investigated in Chapter 3 based on the statistical evaluation of simulated SIR estimates. Another strategy to evaluate SIR estimator performance is to compare the relative benefit of the incorporation of the SIR estimators into a specific application. Though the comparison results are specific to a particular application, they nonetheless provide additional insight into the efficacy of the various SIR estimation algorithms.

Postdetection selection diversity combining [23] is chosen here as an application which is of current interest in wireless communications [27, 2, 124, 125] since it is a relatively simple method to combat the effects of delay spread, AWGN, CCI, and the random amplitude and phase fluctuations due to Rayleigh fading. A postdetection selection diversity combiner (SDC) is classically described [23] as comprising L antennas (spaced far enough apart from each other to ensure that the fading affecting each of the L received signals is uncorrelated or nearly uncorrelated), L detectors, and a switch which connects the detector whose output samples have the *largest SNR* to the decision device. In practice [2], the detector whose samples have the *largest envelope* has been chosen to be connected to the decision device since measurement

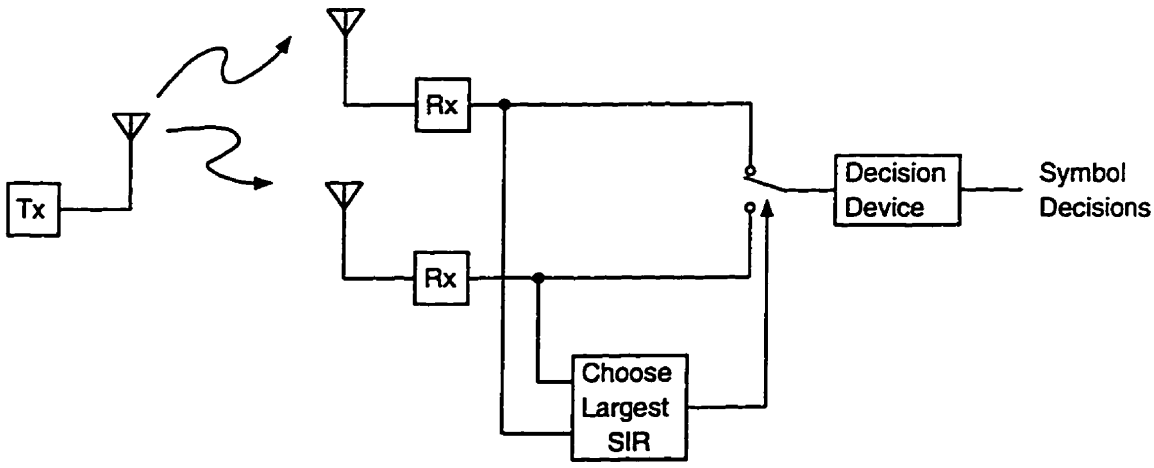

Figure 4.1: Simplified block diagram of postdetection selection diversity combiner.

of the SNR has been considered to be too costly historically. Using the measured SNR or SIR as the selection criterion rather than the received signal envelope (or total received power) should yield better diversity performance.

A simplified block diagram of the diversity system modeled here is illustrated in Figure 4.1 with $L = 2$. Each branch can be represented by the system model presented in Section 3.6.1. The fading between branches is assumed statistically similar but independent; that is, the Doppler and delay profiles are identical, but the sets of fading parameters, θ_m, ν_m, and τ_m, used in (3.9) of Chapter 3 are different for each branch.

The SIR estimators to be evaluated with this SDC are the ones identified in Chapter 3 that generate meaningful SIR estimates regardless of the channel conditions. These include the SIV RxDA, SIV TxDA, MSIV RxDA, MSIV TxDA, DDML RxDA, DDML TxDA, and ADIQ SIR estimators. In addition, two other types of estimators are implemented for comparison which assess received signal quality based on measurement of the *sum* of the signal and impairment components. In one implementation, the average power over a block of N_{sym} symbols is used as the selection criterion. This estimation method is referred to here as *AVGP* (AVeraGe Power) es-

timation. In the other implementation, branch selection is based on measurement of the *instantaneous* power which, in a discrete-time system, is equivalent to measurement of the envelope of each corrupted symbol. This estimation method is referred to here as *INSTP* (INSTantaneous Power) estimation. The SDC selects the branch with the largest AVGP or INSTP estimate. INSTP estimation is the one most often associated with postdetection selection diversity combining (see, for example, [2]).

The AVGP estimator used here computes the total power of the baud-spaced samples, y_n, at the output of the MF given by (3.21) of Chapter 3. The form of this estimator is simply expressed as

$$\hat{\varrho}_{\text{AVGP}} = \frac{1}{N_{\text{sym}}+1} \sum_{n=0}^{N_{\text{sym}}} |y_n|^2.$$

The INSTP estimator is even more simply expressed as

$$\hat{\varrho}_{\text{INSTP}} = |x_n| = |y_n y_{n-1}^*|,$$

which is the envelope of the post-DD samples. The INSTP estimator uses the post-DD samples in order to be consistent with [2].

It is also useful to have some bound on the best possible BER attainable by an $L = 2$ SDC given a certain set of channel conditions and certain block length, N_{sym}. A simple way to obtain such a bound is to run a simulation in which a copy of the transmitted data is used at the receiver to select the branch with the fewest bit errors for each block of N_{sym} symbols. A lower bound on the BER can thus be obtained regardless of the statistics of the channel impairments. These results are identified as the "Ideal" curves in the BER performance plots that follow in Section 4.3.

4.2 Verification of Selection Combiner

As a check of the implementation of the selection combiner pictured in simplified form in Figure 4.1, the simulated BER performance is compared to published results. The channel chosen for verification purposes is a slow, flat Rayleigh fading channel

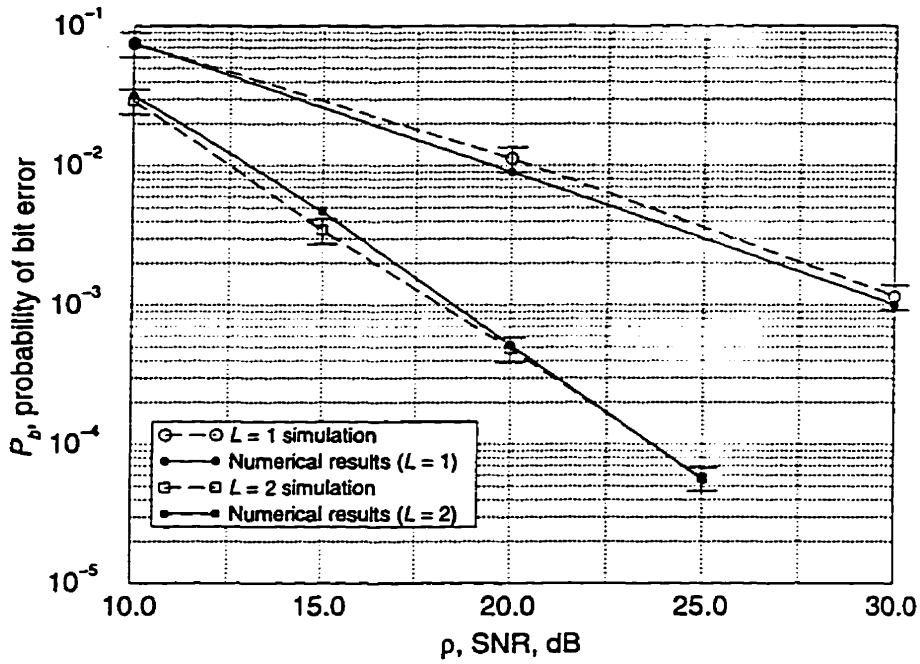

Figure 4.2: Verification of postdetection selection diversity combiner in a slow, Rayleigh fading channel using numerical results of [2].

corrupted by AWGN but no CCI. The simulated BER of this channel is plotted as a function of SNR in Figure 4.2 for $L = 1$ (no diversity) and $L = 2$, and is compared to the numerical results (equation (17) and Figure 4) of [2]. For the case of $L = 2$, it is assumed that the SDC is employed with INSTP estimation. These plots show good agreement between simulation and theory.

4.3 Simulation Results

The BER performance of selection diversity combining using various SIR estimators is displayed in Figures 4.3 to 4.5 for three different scenarios. The intention of this

chapter is not to provide simulation results of the BER performance for an exhaustive set of channel conditions, but rather to illustrate by example the relative improvement in BER attainable using SIR estimators in selection diversity combining for a few representative system configurations.

In the first scenario, the propagation medium is characterized by flat, fast Rayleigh fading with $\nu_{\max} = 100$ Hz. CCI is present at a level of $S/C = 14$ dB, and the SNR is swept from 10 to 50 dB in 10 dB steps. The block length is chosen to be $N_{\text{sym}} = 30$ symbols which is smaller than the block length of 150 symbols assumed for all tests in Chapter 3. The rationale behind the choice of the block length is described below.

For this selection diversity application, the block length should be chosen as small as possible to minimize the number of correctly detected bits that are discarded. Ideally, symbol quality should be assessed on a symbol-by-symbol basis ($N_{\text{sym}} = 1$) as in [2]. To illustrate, consider an $L = 2$ SDC with $N_{\text{sym}} = 100$. Assume 10 symbol errors occur in the *first* 50 symbols of the block corresponding to branch 1, 20 symbol errors occur in the *last* 50 symbols of the block corresponding to branch 2, and the SIR of branch 1 is greater than that of branch 2 (based on an observation interval of 100 symbols) so that the 100-symbol block of branch 1 is correctly selected. Though a block with 10 symbol errors is better than one with 20 symbol errors, if the block size had been 50 symbols in length rather than 100, it is possible that branch 2 could have been selected for the first 50 symbols and branch 1 could have been selected for the last 50 symbols so that, overall, there would have been no symbol errors at all. Based on a 100-symbol block size, the selection of branch 1 effectively discards the correctly detected symbols of branch 2 which could be used to minimize the BER. It is easy to extend this argument to show that smaller and smaller block sizes should yield increasingly better BER performances to the point that, theoretically, the best selection diversity BER performance should be attained with the smallest possible block size; that is, $N_{\text{sym}} = 1$. This requirement must be balanced with the fact that the performances of the SIR estimators improve with *increasing* block length. As a

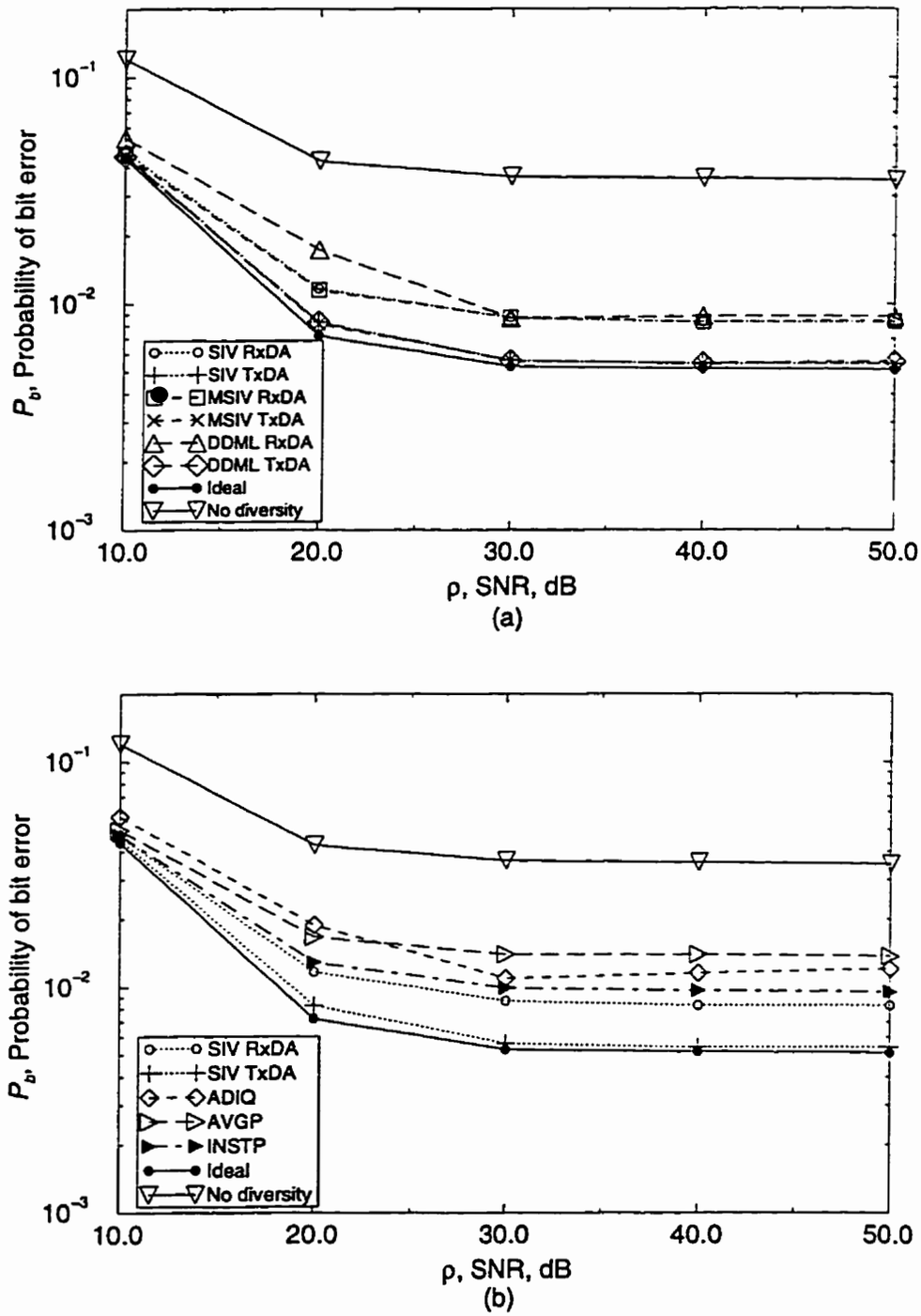

Figure 4.3: BER performance of selection diversity combining as a function of SNR in a flat fading channel with $\nu_{\max} = 100$ Hz, $S/C = 14$ dB, and $N_{\text{sym}} = 30$ using (a) SIV, MSIV, and DDML SIR estimation, and (b) SIV, ADIQ, AVGP, and INSTP estimation.

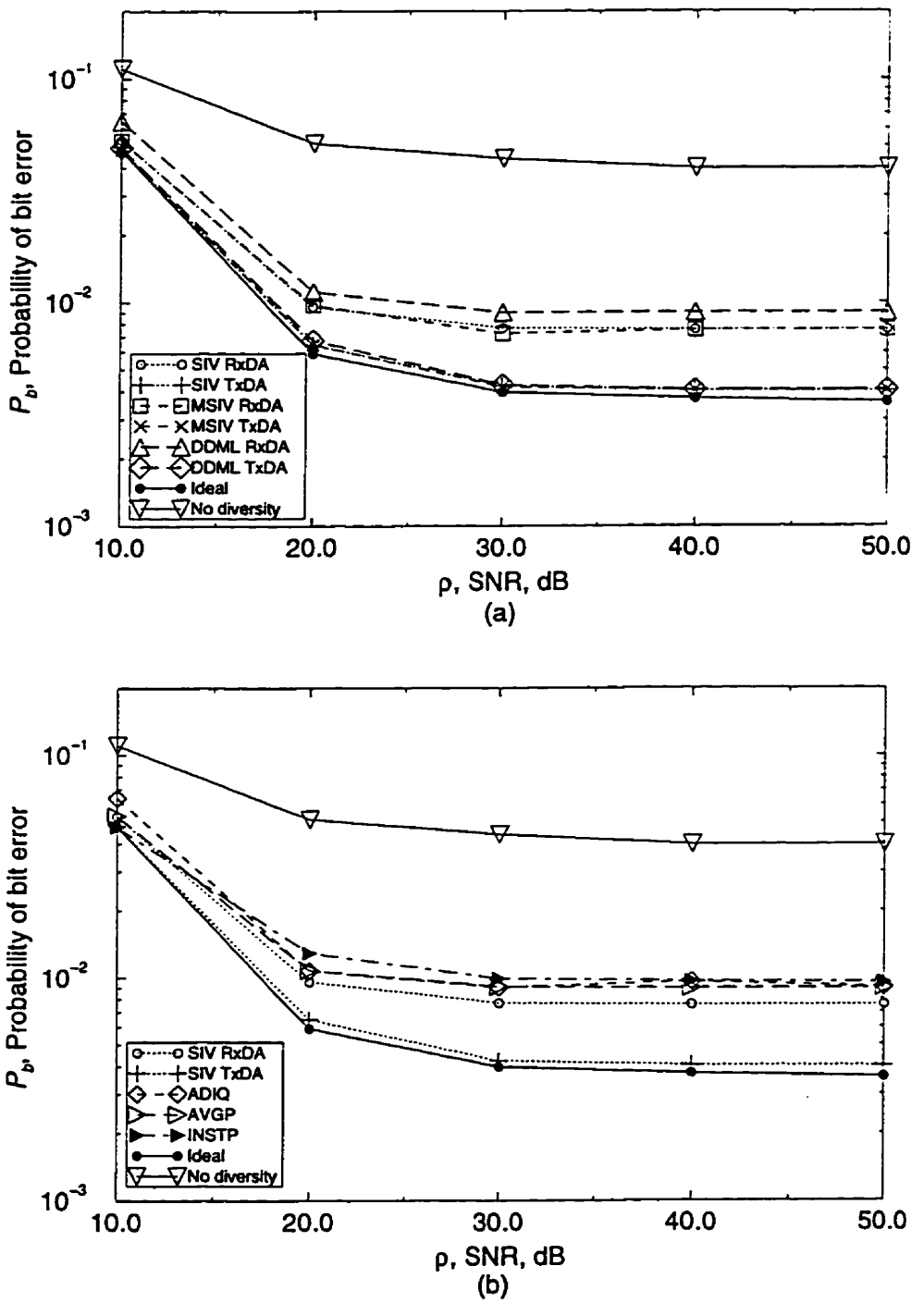

Figure 4.4: BER performance of selection diversity combining as a function of SNR in a flat fading channel with $\nu_{\max} = 100$ Hz, $S/C = 14$ dB, and $N_{\text{sym}} = 10$ using (a) SIV, MSIV, and DDML SIR estimation, and (b) SIV, ADIQ, AVGP, and INSTP estimation.

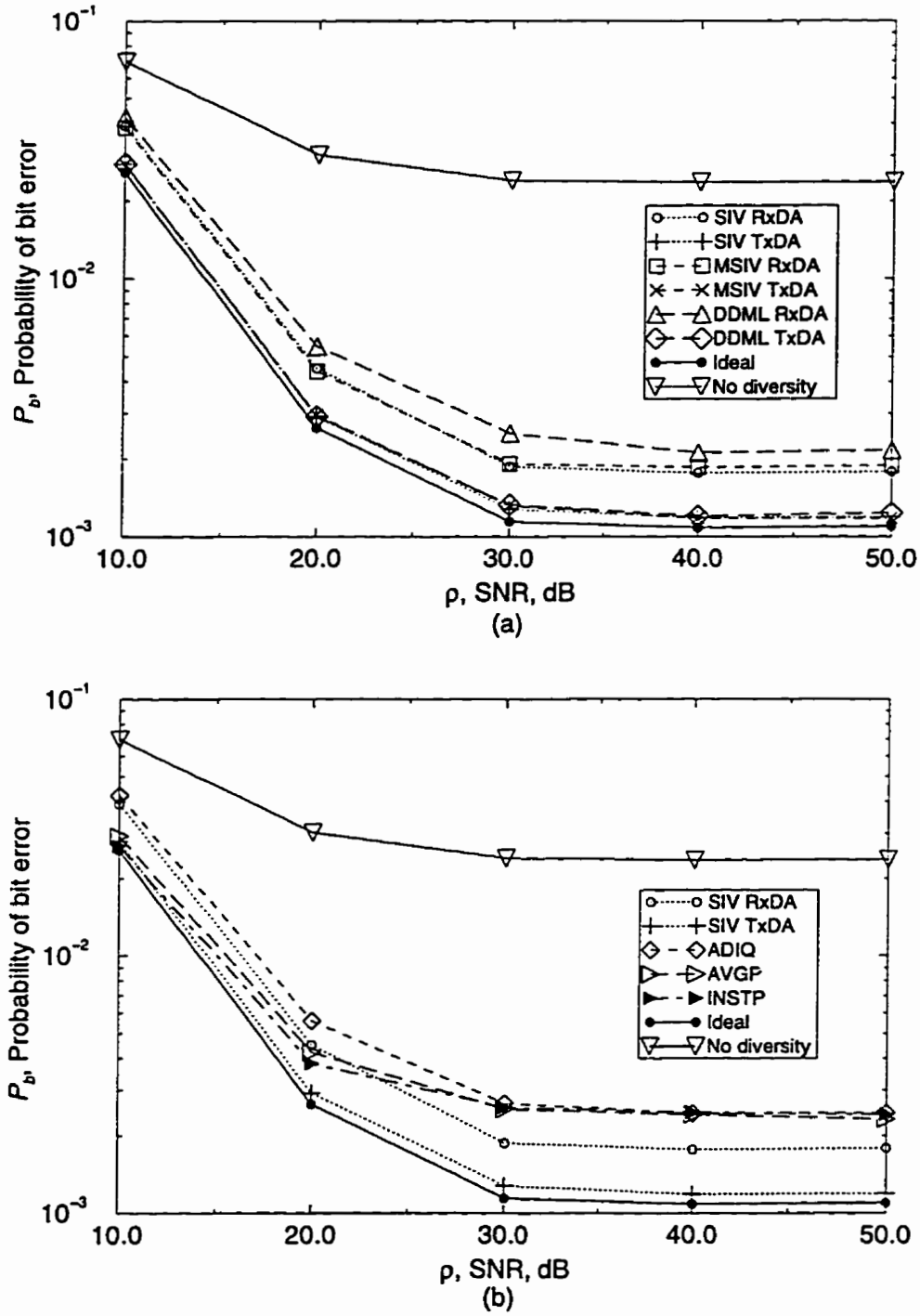

Figure 4.5: BER performance of selection diversity combining as a function of SNR in a frequency-selective fading channel with $\tau_{\text{rms}} = 3$ μs, $\nu_{\text{max}} = 40$ Hz, $S/C = 20$ dB, and $N_{\text{sym}} = 10$ using (a) SIV, MSIV, and DDML SIR estimation, and (b) SIV, ADIQ, AVGP, and INSTP estimation.

first attempt to achieve this balance, a block length of 30 symbols was chosen.

The BER results for the first scenario are displayed in Figure 4.3. The BER performance curves are split between two graphs to reduce congestion, but the SIV, "Ideal", and "No diversity" curves are common to both to facilitate comparison. From Figure 4.3 it can be seen that the difference in performance among the various SDC implementations is not huge, but it is significant. Distinctions arise for values of SNR greater than 30 dB where the error floor is visible. Specifically, the SDC's employing SIV TxDA, MSIV TxDA, and DDML TxDA SIR estimation attain almost ideal BER performance for $N_{\text{sym}} = 30$. The BER performances of the SDC's employing the SIV RxDA and MSIV RxDA SIR estimators are next best. These BER performances are about 17% better than that of a conventional SDC operating with INSTP estimation. Next best is the BER performance of an SDC employing DDML RxDA SIR estimation. The high-SNR performance of the DDML RxDA-assisted SDC is similar to that of the SIV RxDA- and MSIV RxDA-assisted SDC's and better than that of the INSTP-assisted SDC, but the low-SNR (less than about 30 dB) performance is inferior to that of the INSTP-assisted SDC. The BER performance of the ADIQ-assisted SDC is next best and, not surprisingly, the SDC using AVGP estimation performs least well of all of the SDC implementations. These performance rankings are reasonably consistent with the SIR estimator rankings observed in Chapter 3, though the DDML estimators perform better in this SDC application than expected.

The second scenario is almost identical to the first in that the propagation medium, the level of CCI, and the range of values of SNR are the same. The only difference is that the block length is decreased from $N_{\text{sym}} = 30$ to $N_{\text{sym}} = 10$ to see the effect of a shorter observation interval. From Figure 4.4, it can be seen that the performances of almost all of the SDC implementations improve slightly relative to the INSTP implementation which is identical in both plots (recall that the INSTP estimator measures the post-DD signal envelope on a symbol-by-symbol basis so that its estimates are independent of N_{sym}). The DDML RxDA-assisted SDC is the only SIR-

based implementation whose BER performance remains relatively unchanged. As a result, the SIV RxDA- and MSIV RxDA-assisted SDC's exhibit a marginally superior BER performance relative to that of the DDML RxDA-assisted SDC. As well, the SIV RxDA- and MSIV RxDA-assisted SDC implementations maintain their small, but distinct, BER performance advantage over all of the other SDC implementations using in-service quality estimators. The SDC implementations using the TxDA forms of the SIV, MSIV, and DDML estimators maintain their superior rankings, achieving BER performances that are very close to ideal.

In the final scenario, the propagation medium is characterized by frequency-selective, fast Rayleigh fading with $\tau_{\text{rms}} = 3$ μs and $\nu_{\text{max}} = 40$ Hz (decreased from $\nu_{\text{max}} = 100$ Hz). The signal-to-CCI ratio is increased from $S/C = 14$ dB to 20 dB. As in the other two scenarios, the SNR is swept from 10 to 50 dB in 10 dB steps. The block length is maintained at $N_{\text{sym}} = 10$ symbols, consistent with that of the second scenario. Relative to the BER performances with diversity in the first two scenarios, the BER performances with diversity in this frequency-selective scenario are better since the improvement in performance due to the smaller Doppler spread and lower interference level is more significant than the degradation in performance due to the added frequency selectivity. The high-SNR rankings are unchanged in this frequency-selective scenario as compared to the previous flat-fading case ($N_{\text{sym}} = 10$ in both cases).

4.4 BER Performance Summary

There is some correlation between the ranking of the SIR estimators in terms of relative SDC BER performance improvement (for the scenarios considered in Section 4.3) and the ranking of the SIR estimators developed in Chapter 3 based on application-independent statistical tests. Specifically, it is not surprising that the SDC's employing SIV TxDA and MSIV TxDA SIR estimation perform well, approaching ideal performance for a given block length, as these estimators were judged to perform best

in Chapter 3. The SDC's incorporating the RxDA forms of the SIV and MSIV estimators perform best of all the SDC's using in-service estimators. The differences in BER performance between SIV-based SDC implementations and MSIV-based SDC implementations are negligible.

The DDML TxDA-assisted SDC exhibits similar BER performance to the SIV TxDA- and MSIV TxDA-assisted SDC's, even though the DDML TxDA estimator appears to be inferior to the SIV and MSIV TxDA estimators based on the bias, variance, MSE, and SIR/BER correlation results of Chapter 3. The BER performance of the DDML RxDA-assisted SDC, however, is not as good as the performances of the SIV RxDA- and MSIV RxDA-assisted SDC's, and the difference becomes more pronounced with decreasing block length.

In Section 3.10, it was stated that applications incorporating DDML and ADIQ estimators should perform similarly since their bias, variance, MSE, and SIR/BER correlation performances are very similar. In fact, the DDML RxDA- and ADIQ-assisted SDC's exhibit similar BER performances using a block length of ten symbols, but the BER performance of the DDML RxDA-assisted SDC is better using a block length of thirty symbols. The DDML TxDA-assisted SDC performs significantly better than either the DDML RxDA- or ADIQ-assisted SDC's regardless of block length. It would be interesting to see if the relative rankings of the SIR estimators of Chapter 3 change if the block length is reduced from $N_{\text{sym}} = 150$ to $N_{\text{sym}} = 30$ or 10. Clearly, it is advisable to evaluate the SIR estimators in the context of each intended application in order to determine the estimator that is most suitable.

Recall that the INSTP-assisted SDC, which is the most common SDC implementation used in practice, makes selections on a symbol-by-symbol basis which, as mentioned earlier, is the optimal selection strategy to avoid discarding correct symbols. However, its BER performance is only better than that of the ADIQ- and AVGP-assisted SDC's given a block length of 30 symbols. Given a block length of 10 symbols, the BER performances of the ADIQ- and AVGP-assisted SDC's improve to the point

that they become comparable to the performance of the INSTP-assisted SDC (recall that the performance of the INSTP-assisted SDC is independent of N_{sym}). Further investigation is required to determine the optimal block length, and to determine whether the optimal block length depends on the channel conditions.

Though the BER performance differences among the various SDC implementations is small, the simulation results of Figures 4.3 to 4.5 nonetheless indicate that the use of SIR estimation can yield a performance improvement over SDC implementations that are based on measurement of the total received power.

Chapter 5

Conclusions

5.1 Summary

In this thesis, SNR estimation in the AWGN channel and SIR estimation in general mobile radio channels were investigated, and certain chosen SIR estimators were implemented in a postdetection selection diversity combining application.

In Chapter 2, the problem of SNR estimation in the classical AWGN channel was defined, and computer simulation was used to study the relative performances of various SNR estimation algorithms found in the literature. Both BPSK-modulated signals in real AWGN and 8-PSK signals in complex AWGN were considered. The structures of all of the SNR estimators except one (the SSME series of algorithms) were derived for both the real and complex cases (in the literature, the SNR estimators are typically presented either for the real or complex channel, but not both). Almost all of the published estimators required some modification to be used in the assumed channel model of Chapter 2. The SNR estimators studied form estimates either from known transmitted data, estimated received data, or measured statistical properties of the received signal. The latter two types of estimators (which do not require knowledge of the transmitted data) are referred to as in-service estimators. The SNR estimators were simulated under identical channel conditions using common performance metrics to facilitate a fair performance comparison. The performance metrics used are the sample bias, sample variance, and sample MSE. In addition,

theoretical bounds (CRB's) on the variance in both real and complex channels were derived and compared with the results of the simulations.

The choice of the "best" SNR estimator was found to depend on the specific application in which the estimator is to be used. In the case that known data is available at the receiver, the performances of the ML TxDA and SNV TxDA SNR estimators were found to exhibit the smallest biases, variances, and MSE's, and came closest to their respective CRB's. Their performances fall slightly short of the theoretical optimum, but it is difficult to justify attempts to find estimators that come closer to the theoretical bounds. Of the in-service SNR estimators, the "best" choice was found to depend on factors such as the block length, the number of available samples per symbol, the type and order of modulation used, the range of SNR over which the estimator is to operate, and the complexity of the implementation.

In Chapter 3, the problem of SIR estimation in general mobile radio channels was defined, and computer simulation was used to study the relative performances of various SIR estimation algorithms found in the literature and developed in this thesis. The basis of the channel model is the IS-54 North American digital cellular standard which specifies $\pi/4$-DQPSK as the modulation scheme. The various impairment processes incorporated into the channel model include fast, frequency-selective Rayleigh fading, CCI, and AWGN. The amplitude and phase fluctuations caused by the channel were modeled as a Rayleigh fading process having a Doppler profile described by the Jakes spectrum. The multipath was assumed to have an exponential delay profile. Several fading simulators were considered as candidates to model these effects, and the Monte Carlo method for the modeling of GWSSUS channels was chosen. Statistical tests were performed on the fading simulator to verify its functionality. Five different test channels were presented, each emphasizing a particular channel impairment. The published SIR estimators were described followed by three new algorithms presented for the first time in this thesis. These SIR estimators were simulated in the five test channels and their performances were compared based on

their measured variances and the correlations between their mean SIR's and the BER of the channel.

It was difficult to determine the "best" SIR estimator as the true SIR is difficult to determine in some of the five test channels. However, it was clear that certain SIR estimators (the PML, SVR, and M_2M_4 SIR estimators) do not work well in fast Rayleigh fading environments so that the field of candidates for the "best" SIR estimator that can operate under any channel conditions was narrowed to the SIV, MSIV, DDML, and ADIQ estimators. Based primarily on the SIR/BER correlations, it appeared that the SIV and MSIV SIR estimators outperformed the DDML and ADIQ algorithms.

Chapter 4 investigated the performances of the SIV, MSIV, DDML, and ADIQ estimators from a different perspective in the specific application of postdetection selection diversity combining. In this application, the SIR estimators were used to select the better of two diversity branches based on observation intervals of N_{sym} symbols. For comparison, SDC's incorporating an average power (AVGP) estimator and an instantaneous power (INSTP) estimator (which is used most often in practice) were also simulated, and a simulated bound on the lowest attainable BER was provided.

There is some correlation between the ranking of the SIR estimators based on their relative improvement of the BER in the SDC application with the ranking in Chapter 3 based on application-independent statistical tests. The SIV TxDA-assisted and MSIV TxDA-assisted SDC implementations yielded BER curves that were just slightly inferior to the simulated lower bound. The DDML TxDA-assisted SDC performed similarly which was slightly surprising considering the results of Chapter 3 which suggested that the SIV and MSIV estimators should outperform the DDML estimator. Of the in-service estimators, the incorporation of the SIV RxDA and MSIV RxDA algorithms yielded the best SDC BER performances, but the curves were not too far from the BER curves of the SDC's that incorporated the remaining in-service estimators (the DDML RxDA, ADIQ, AVGP, and INSTP estimators). The

relative rankings were found to be sensitive to the length of the observation interval.

5.2 Suggestions for Further Study

1. In Section 2.3.2 of Chapter 2, the reduced-bias forms of the complex ML TxDA and ML RxDA estimators are given by (2.72) and (2.73), respectively. It was indicated in Section 2.3.2, that the factor used in the reduced-bias forms of the complex ML SNR estimators was obtained heuristically. In a manner similar to the derivation of Thomas [47] for the real AWGN case, the pdf of the ML TxDA estimator for complex AWGN could be derived which could then, in turn, be used to obtain the proper factor analytically rather than heuristically.

2. A difficulty that arises in any comparison of published algorithms is that new techniques continually appear in the literature. Also, it is difficult to form a comprehensive list of every published technique since the algorithms are sometimes inconspicuous in the existing literature. As examples, SIR (and SNR) techniques which would be interesting to add to the list of tested estimators include those described in [59], [70, page 346], and [60]. The references [59] and [70] appear to describe the same technique. These estimators would be of particular interest if they prove to operate well under all fading conditions (particularly fast Rayleigh fading) so that they could be added to the list of "robust" estimators which now include only the SIV, MSIV, DDML, and ADIQ estimators.

3. The pdf's of the SIR estimators (especially the SVR and PML estimators which exhibit rising biases and variances at high SNR) could be evaluated to obtain a better insight into the operation of the estimators.

4. The topic of BER estimation was not considered in this thesis since focus was placed on techniques that yield estimates that are not just useful as relative

measures of quality, but also as absolute quantities required as input parameters to some algorithms. It would still be interesting, however, to compare the performances of SNR or SIR estimators to the performances of BER estimators (such as the algorithm described in [40]) in specific applications that only require relative measures of quality (for example, postdetection selection diversity combining).

5. A more rigorous investigation of the BER performance of SIR estimation applied to diversity combining could be performed. Simulations could be run to determine the optimum block length given several different channel configurations. Other diversity combining techniques (such as maximal-ratio combining) could also be simulated for comparison. Before embarking on such a project, however, the degree to which BER performance could theoretically be improved over the conventional combiner implementations which operate by measuring the total signal-plus-impairment power rather than the SIR should first be determined.

6. The SIV, MSIV, DDML, ADIQ, and AVGP algorithms are all examples of "interval" estimators since their estimates are based on a block of symbols, whereas the INSTP algorithms is an example of a "point" estimator whose estimates are made on a symbol-by-symbol basis. It would be interesting to convert the SIV, MSIV, DDML, and ADIQ interval estimators to point-estimator forms and compare SDC implementations using these point estimators with SDC implementations using interval forms.

7. Since the performance advantage of using SIR estimators versus total signal-plus-impairment power estimators in the selection diversity combining application is relatively small, different applications could be tried in which the advantage of SIR estimation is more pronounced. As suggestions, a couple of examples of potential applications include timing and frequency recovery (as explored in [28] using only the DDML TxDA SIR estimator) and Viterbi decoding

in fast, frequency-selective fading channels (see [35]).

8. A logical extension of this work is to implement certain SIR estimators (or simplified versions of these estimators) in a hardware implementation for a particular application, and perform lab and field tests.

Appendix A

Notes on Probability of Error and Noise Power Spectral Density

This appendix is included to help clarify some of the assumptions made in this thesis specifically regarding noise power spectral density and the relationship between discrete SNR and error probability. It is a constant source of frustration to try to interpret plots of error probability published in papers and textbooks since there is a significant lack of consistency in the way axes are labeled and interpreted. Often, what one author refers to as the SNR is actually E_b/N_0 or vice-versa, for example. Also, there is little mention made in the literature of the distinction between one-dimensional (real) and two-dimensional (complex) channels and the implications on the interpretation of the noise power spectral density (PSD), N_0. The following development is not meant to be rigorous, but attempts to tie together, clarify, and highlight pertinent concepts presented and discussed in Chapters 3 and 4 of [65]. The notation used here is modified somewhat from that given in [65].

The assumed, idealistic system model is first identified which is slightly different from the model described in Section 2.2, but the results derived here are general. The main difference between the model described here and that in Section 2.2 is that the transmitted signal here consists of a single pulse-shaped symbol which is nonzero only over a baud interval whereas, in Section 2.2, the signal is a pulse-shaped sequence of symbols, each symbol spanning several baud intervals. The baud-

constrained assumption is made to simplify the analysis. Following the description of the system model, the probability of error for a single, binary transmitted pulse in complex AWGN (two-dimensional problem) is derived. Finally, the probability of error for a single, binary transmitted pulse in real AWGN (one-dimensional problem) is derived and compared to the complex case. The purpose of deriving expressions for the probability of error for these two cases is to justify the noise PSD model assumed in each case. A summary of the main points that arise in the derivations is given in Section A.4.

A.1 System Description and Properties

Assume a system in which one of two possible bandpass[1], pulse-shaped waveforms, $s_m(t)$, $m \in \{1,2\}$, is transmitted and corrupted by bandpass white noise, $n(t)$, to form one of the two possible received signals

$$y(t) = s_1(t) + n(t)$$

or

$$y(t) = s_2(t) + n(t).$$

The transmitted signal is nonzero only over the interval, $0 \leq t \leq T$, and the noise has a two-sided PSD given by

$$S_n(f) = \begin{cases} N_0/2 & f_c - \frac{B}{2} \leq |f| \leq f_c + \frac{B}{2} \\ 0 & \text{otherwise,} \end{cases} \tag{A.1}$$

where f_c is the centre frequency and B is the bandwidth. It is in the context of the bandpass signal, $y(t)$, that the noise PSD is defined and measured in practice.

The signal and noise may be expressed in terms of lowpass equivalents as

$$s_m(t) = \mathrm{Re}\left\{u_m(t)e^{j2\pi f_c t}\right\} \tag{A.2}$$

[1]Strictly speaking, a function limited in time cannot be bandlimited, but it is assumed here that there is insignificant energy outside some bandwidth, B.

and

$$n(t) = \mathrm{Re}\left\{z(t)e^{j2\pi f_c t}\right\} \tag{A.3}$$

respectively, where $u_m(t)$ and $z(t)$ are the respective lowpass, complex equivalents of the signal and noise. Expressions for the energy and correlation of the transmitted signal, and the autocorrelation of the noise are provided below for use in the subsequent development.

The energy of the transmitted signal is

$$\begin{aligned} \mathcal{E}_m &\approx \int_0^T s_m^2(t)\,dt \\ &\approx \int_0^T \left(\mathrm{Re}\left\{u_m(t)e^{j2\pi f_c t}\right\}\right)^2 dt. \end{aligned} \tag{A.4}$$

Let $u_m(t) = X(t) + jY(t)$ and substitute into (A.4) to obtain

$$\begin{aligned} \mathcal{E}_m &= \int_0^T \left(\mathrm{Re}\left\{(X(t)+jY(t))e^{j2\pi f_c t}\right\}\right)^2 dt \\ &= \frac{1}{2}\int_0^T \left(X^2(t)+Y^2(t)+(X^2(t)-Y^2(t))\cos 4\pi f_c t - 2X(t)Y(t)\sin 4\pi f_c t\right) dt \\ &\approx \frac{1}{2}\int_0^T |u_m(t)|^2\,dt, \end{aligned} \tag{A.5}$$

where it is assumed that f_c is much larger than the highest frequency components of $u_m(t)$ so that the integrals of the sine and cosine terms in (A.5) are approximately zero. The correlation of the two possible lowpass-equivalent transmitted signals of equal energy, $\mathcal{E}$, is

$$\varphi = \frac{1}{2\mathcal{E}}\int_0^T u_1(t)u_2^*(t)dt = \frac{1}{2\mathcal{E}}\int_0^T u_2(t)u_1^*(t)dt. \tag{A.6}$$

It is shown on page 155 of [65] that the autocorrelation of the noise is given by

$$R_n(\tau) = \mathrm{Re}\left\{R_z(\tau)e^{j2\pi f_c \tau}\right\} \tag{A.7}$$

where R_z is the autocorrelation of the complex, lowpass-equivalent of the noise given by

$$R_z(\tau) = \tfrac{1}{2}E\{z^*(t)z(t+\tau)\}. \tag{A.8}$$

The relationship between the PSD of the bandpass noise, $S_n(f)$, and the PSD of the lowpass-equivalent noise, $S_z(f)$, is found by taking the Fourier transform of (A.7) as

$$\begin{aligned} S_n(f) &= \mathcal{F}\{R_n(\tau)\} \\ &= \int_{-\infty}^{\infty} R_n(\tau)e^{-j2\pi f\tau}d\tau \\ &= \int_{-\infty}^{\infty} \mathrm{Re}\{R_z(\tau)e^{j2\pi f_c\tau}\}e^{-j2\pi f\tau}d\tau \\ &= \frac{1}{2}\int_{-\infty}^{\infty} R_z(\tau)e^{-j2\pi(f-f_c)\tau}d\tau + \frac{1}{2}\left[\int_{-\infty}^{\infty} R_z(\tau)e^{-j2\pi(-f-f_c)\tau}d\tau\right]^* \\ &= \tfrac{1}{2}[S_z(f-f_c) + S_z^*(-f-f_c)] \\ &= \tfrac{1}{2}[S_z(f-f_c) + S_z(-f-f_c)] \end{aligned} \tag{A.9}$$

where $S_z^*(-f-f_c) = S_z(-f-f_c)$ because $S_n(f)$ is real—specifically, it is given by (A.1). The implication of (A.9) and (A.1) is that the PSD of $z(t)$ is

$$S_z(f) = \begin{cases} N_0 & |f| \le \frac{B}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{A.10}$$

As an illustration of the model developed here, the probability of error for a single transmitted, binary pulse is given below for both the complex AWGN channel and the real AWGN channel (the derivation below follows the development of [65] for binary signaling in complex AWGN).

A.2 Probability of Error for a Single, Binary Transmitted Symbol in Complex AWGN

Consider the lowpass-equivalent, received signal expressed as

$$r(t) = \alpha e^{-j\phi}u_m(t) + z(t),\ m \in \{1,2\},\ 0 \le t \le T \tag{A.11}$$

where α explicitly represents an attenuation factor, ϕ explicitly represents a phase shift in the carrier, and $u_m(t)$ and $z(t)$ are the lowpass-equivalents of the bandpass signals, $s_m(t)$ and $n(t)$, as described in Section A.1. Assume the energies of the two possible transmitted pulses are identical; that is, $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{E}$. It is shown

on page 242 of [65] that the correlator implementation of a MF receiver performs optimum reception on one of the two possible signals identified by (A.11) by forming the decision variables

$$U_m = \mathrm{Re}\left\{e^{j\phi}\int_0^T r(t)u_m^*(t)dt\right\},\ m \in \{1,2\} \tag{A.12}$$

and selecting the largest of the two. Substituting (A.11) into (A.12) for each of the two possible transmitted symbols, one obtains

$$\begin{aligned} U_1 &= \mathrm{Re}\left\{e^{j\phi}\int_0^T [\alpha e^{-j\phi}u_1(t) + z(t)]u_1^*(t)dt\right\} \\ &= \mathrm{Re}\left\{\alpha\int_0^T |u_1(t)|^2 dt + e^{j\phi}\int_0^T z(t)u_1^*(t)dt\right\} \\ &= \mathrm{Re}\{2\alpha\mathcal{E} + N_1\} \\ &= 2\alpha\mathcal{E} + N_{1r} \end{aligned}$$

and, similarly,

$$\begin{aligned} U_2 &= \mathrm{Re}\left\{e^{j\phi}\int_0^T [\alpha e^{-j\phi}u_1(t) + z(t)]u_2^*(t)dt\right\} \\ &= \mathrm{Re}\left\{\alpha\int_0^T u_1(t)u_2^*(t)dt + e^{j\phi}\int_0^T z(t)u_2^*(t)dt\right\} \\ &= \mathrm{Re}\{2\alpha\mathcal{E}\varphi + N_2\} \\ &= 2\alpha\mathcal{E}\varphi_r + N_{2r}, \end{aligned}$$

where

$$\begin{aligned} N_1 &= e^{j\phi}\int_0^T z(t)u_1^*(t)dt \\ N_2 &= e^{j\phi}\int_0^T z(t)u_2^*(t)dt \\ N_{1r} &= \mathrm{Re}\{N_1\} \\ N_{2r} &= \mathrm{Re}\{N_2\} \\ \varphi_r &= \mathrm{Re}\{\varphi\}, \end{aligned}$$

and use is made of (A.5) and (A.6).

If $U_1 > U_2$, the receiver decides $u_1(t)$ was transmitted; if $U_2 > U_1$, the receiver decides $u_2(t)$ was transmitted. If $u_1(t)$ is transmitted, the probability of the receiver making an error is

$$P\{U_2 > U_1\} = P\{U_1 - U_2 < 0\} = P\{V < 0\}$$

where $V = U_1 - U_2 = 2\alpha\mathcal{E}(1 - \varphi_r) + N_{1r} - N_{2r}$. Since N_{1r} and N_{2r} are Gaussian, V is also Gaussian and the probability of error, $P\{V < 0\}$ may be expressed in terms of the mean, m_V, and variance, σ_V^2, of V as

$$P\{V < 0\} = \int_{-\infty}^{0} p(v)dv = \frac{1}{\sqrt{2\pi}\sigma_V}\int_{-\infty}^{0} e^{-\frac{(v-m_V)^2}{2\sigma_V^2}}\,dv = \frac{1}{2}\mathrm{erfc}\left(\frac{m_V}{\sqrt{2}\sigma_V}\right) \qquad \text{(A.13)}$$

where

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-t^2}\,dt.$$

Since $u_1(t)$ and $u_2(t)$ are transmitted with equal probability, the binary probability of error is

$$\begin{aligned} P_e &= \tfrac{1}{2}P\{V < 0\} + \tfrac{1}{2}P\{V > 0\} \\ &= \frac{1}{2}\mathrm{erfc}\left(\frac{m_V}{\sqrt{2}\sigma_V}\right). \end{aligned} \qquad \text{(A.14)}$$

The mean of V is evaluated as

$$m_V = E\{V\} = 2\alpha\mathcal{E}(1 - \varphi_r). \qquad \text{(A.15)}$$

The variance of V is evaluated as

$$\begin{aligned} \sigma_V^2 &= E\{(V - E\{V\})^2\} \\ &= E\{(N_{1r} - N_{2r})^2\} \\ &= E\{N_{1r}^2\} + E\{N_{2r}^2\} - 2E\{N_{1r}N_{2r}\}. \end{aligned} \qquad \text{(A.16)}$$

Recognizing that $N_{1r} = \mathrm{Re}\{N_1\} = \frac{1}{2}(N_1 + N_1^*)$ allows $E\{N_{1r}^2\}$ to be evaluated as

$$E\{N_{1r}^2\} = \tfrac{1}{4}E\{(N_1^2 + 2N_1N_1^* + (N_1^*)^2)\}. \qquad \text{(A.17)}$$

Letting $N_1 = X_1 + jY_1$, it can be shown that

$$E\{N_1^2\} + E\{(N_1^*)^2\} = 2E\{X_1^2 - Y_1^2\} = 0,$$

(making the reasonable assumption that the in-phase and quadrature noise powers are equal) so that

$$\begin{aligned} E\{N_{1r}^2\} &= \tfrac{1}{2}E\{N_1 N_1^*\} \\ &= \frac{1}{2}E\left\{\int_0^T \int_0^T z(t)z^*(\tau)u_1^*(t)u_1(\tau)\,dt\,d\tau\right\} \\ &= \frac{1}{2}2N_0\int_0^T |u_1(t)|^2 dt \\ &= \frac{1}{2}(2N_0)(2\mathcal{E}) \\ &= 2N_0\mathcal{E} \end{aligned} \tag{A.18}$$

using (A.8), (A.10), and the fact that

$$\begin{aligned} R_z(\tau) &= \mathcal{F}^{-1}\{S_z(f)\} \\ &\approx N_0\delta(\tau), \end{aligned} \tag{A.19}$$

where $\delta(\tau)$ is the delta function. Equation (A.19) is only approximately true since the bandwidth, B, of the bandpass white noise is finite. Very similarly, $E\{N_{2r}^2\} = 2N_0\mathcal{E}$. Finally, the remaining term is evaluated as

$$\begin{aligned} E\{N_{1r}N_{2r}\} &= \tfrac{1}{4}E\{(N_1 + N_1^*)(N_2 + N_2^*)\} \\ &= \tfrac{1}{4}(E\{N_1N_2\} + E\{N_1^*N_2^*\} + E\{N_1N_2^*\} + E\{N_1^*N_2\}). \end{aligned} \tag{A.20}$$

Letting $N_1 = X_1 + jY_1$ and $N_2 = X_2 + jY_2$, it can be shown that

$$E\{N_1N_2\} + E\{N_1^*N_2^*\} = 2E\{X_1X_2 - Y_1Y_2\} = 0,$$

so that

$$E\{N_{1r}N_{2r}\} = \tfrac{1}{4}(E\{N_1N_2^*\} + E\{N_1^*N_2\})$$

$$
\begin{aligned}
&= \frac{1}{4}\left(E\left\{\int_0^T\int_0^T z(t)z^*(\tau)u_1^*(t)u_2(\tau)\,dt\,d\tau\right\}\right.\\
&\quad \left. + E\left\{\int_0^T\int_0^T z^*(t)z(\tau)u_1(t)u_2^*(\tau)\,dt\,d\tau\right\}\right)\\
&= \frac{1}{4}\left(2N_0\int_0^T u_1^*(t)u_2(t)\,dt + 2N_0\int_0^T u_1(t)u_2^*(t)\,dt\right)\\
&= \frac{1}{4}(4N_0\mathcal{E}\varphi + 4N_0\mathcal{E}\varphi)\\
&= 2N_0\mathcal{E}\varphi. \qquad \text{(A.21)}
\end{aligned}
$$

Equation (A.16) then evaluates to

$$\sigma_V^2 = 2N_0\mathcal{E} + 2N_0\mathcal{E} - 2(2N_0\mathcal{E}\varphi) = 4N_0\mathcal{E}(1-\varphi). \qquad \text{(A.22)}$$

Using (A.15) and (A.22), (A.14) becomes

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\alpha^2\mathcal{E}}{2N_0}(1-\varphi)}\right) = Q\left(\sqrt{\frac{\alpha^2\mathcal{E}}{N_0}(1-\varphi)}\right) \qquad \text{(A.23)}$$

using the identity

$$\text{erfc}(x) = 2Q(\sqrt{2}x).$$

If $u_1(t)$ and $u_2(t)$ are antipodal signals, then $\varphi = -1$ and the probability of error for binary, antipodal signaling in complex AWGN is expressed as

$$P_{\text{e,complex}} = Q\left(\sqrt{\frac{2\alpha^2\mathcal{E}}{N_0}}\right), \qquad \text{(A.24)}$$

which is the classic result [65, 62] where $\alpha^2\mathcal{E}$ is often identified as E_b, the energy per bit.

A.3 Probability of Error for a Single, Binary Transmitted Symbol in Real AWGN

The derivation of the probability of error for binary signaling in real AWGN is a classic problem which is derived here in a parallel manner as for the complex case discussed in Section A.2. However, some of the expressions given in the description

of the system model in Section A.1 need to be modified for this real case. Consider the received signal given by

$$r(t) = \alpha u_m(t) + z(t),\ m \in \{1,2\},\ 0 \leq t \leq T \tag{A.25}$$

where all quantities are real. The energy of the transmitted signal is

$$\mathcal{E}_m = \int_0^T u_m^2(t)dt, \tag{A.26}$$

the correlation of the two possible transmitted signals is

$$\varphi = \frac{1}{\mathcal{E}} \int_0^T u_1(t)u_2(t)dt, \tag{A.27}$$

and the autocorrelation of the real AWGN process is given by

$$R_z(\tau) = E\{z(t)z(t+\tau)\}. \tag{A.28}$$

The PSD of the noise in this real case is

$$S_z(f) = \begin{cases} N_0/2 & |f| \leq \frac{B}{2} \\ 0 & \text{otherwise} \end{cases} \tag{A.29}$$

so that

$$R_z(\tau) \approx \frac{N_0}{2}\delta(\tau) \tag{A.30}$$

where, again, the expression is only approximate since the bandwidth, B, of the bandpass white noise is finite. The magnitude of the PSD of the real noise is different from that of the complex noise by a factor of two. Effectively, for this real case, the signal may be considered to exist only in one of the in-phase or quadrature channels; the complex noise that is orthogonal to the signal may be ignored so that only half of the total complex noise power need be considered in the real case. It is shown in the derivation below that assuming (A.29) for the PSD of real noise yields the correct expression for the probability of error.

The decision variables in this real case are

$$\begin{aligned} U_1 &= \int_0^T [\alpha u_1(t) + z(t)]u_1(t)dt \\ &= \alpha \int_0^T u_1^2(t)dt + \int_0^T z(t)u_1(t)dt \\ &= \alpha\mathcal{E} + N_1 \end{aligned}$$

and

$$\begin{aligned} U_2 &= \int_0^T [\alpha u_1(t) + z(t)]u_2(t)dt \\ &= \alpha \int_0^T u_1(t)u_2(t)dt + \int_0^T z(t)u_2(t)dt \\ &= \alpha\mathcal{E}\varphi + N_2, \end{aligned}$$

where

$$N_1 = \int_0^T z(t)u_1(t)dt$$

and

$$N_2 = \int_0^T z(t)u_2(t)dt.$$

As before, form the difference $V = U_1 - U_2$. The probability of error in terms of this variable is given by (A.14), so all that remains is to find the mean and variance of V.

The mean of V evaluates to

$$m_V = E\{V\} = \alpha\mathcal{E}(1 - \varphi), \tag{A.31}$$

and the variance of V is evaluated as

$$\sigma_V^2 = E\{(V - E\{V\})^2\} = E\{N_1^2\} + E\{N_2^2\} - 2E\{N_1N_2\}. \tag{A.32}$$

The first term in (A.32) is

$$E\{N_1^2\} = E\left\{\int_0^T \int_0^T z(t)z(\tau)u_1(t)u_1(\tau)\,dt\,d\tau\right\} = \frac{N_0}{2}\int_0^T u_1^2(t)dt = \frac{N_0\mathcal{E}}{2}, \tag{A.33}$$

and the second term evaluates similarly to $E\{N_2^2\} = N_0\mathcal{E}/2$. The cross-correlation may be evaluated as

$$E\{N_1N_2\} = E\left\{\int_0^T \int_0^T z(t)z(\tau)u_1(t)u_2(\tau)\,dt\,d\tau\right\} = \frac{N_0}{2}\int_0^T u_1(t)u_2(t)dt = \frac{N_0\mathcal{E}\varphi}{2}. \tag{A.34}$$

Using these results, (A.32) becomes

$$\sigma_V^2 = \frac{N_0\mathcal{E}}{2} + \frac{N_0\mathcal{E}}{2} - 2\frac{N_0\mathcal{E}\varphi}{2} = N_0\mathcal{E}(1-\varphi). \tag{A.35}$$

Substituting (A.31) and (A.35) into (A.14), the probability of error may be expressed as

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\alpha^2\mathcal{E}}{2N_0}(1-\varphi)}\right) = Q\left(\sqrt{\frac{\alpha^2\mathcal{E}}{N_0}(1-\varphi)}\right). \tag{A.36}$$

Assuming antipodal signals, as before, so that $\varphi = -1$, the probability of error for antipodal signaling in real AWGN is expressed as

$$P_{e,\text{real}} = Q\left(\sqrt{\frac{2\alpha^2\mathcal{E}}{N_0}}\right), \tag{A.37}$$

which is identical to the result given by (A.24). Again, the quantity, $\alpha^2\mathcal{E}$ is often identified as the energy per bit, E_b.

A.4 Summary of Main Points

1. The two-sided noise PSD of $n(t)$ is defined (and measured, in practice) in the bandpass context shown in Figure A.1(a). For the sake of illustration, it is assumed that the receiver bandpass filter is ideal, centred at f_c, and of bandwidth, B. The two-sided PSD has an amplitude of $N_0/2$.

2. The signal, $z(t)$, which is the lowpass-equivalent of $n(t)$, is complex. It is important to be aware that this is true even if the transmitted signal has a one-dimensional lowpass-equivalent (such as BPSK). The PSD of the lowpass-equivalent of the noise is illustrated in Figure A.1(b) having an amplitude of N_0, and is used in Section 2.2 of Chapter 2 to show that the relationship between the SNR, ρ, and E_s/N_0 is

$$\rho = \frac{E_s}{N_0}$$

for (in general) complex signals in complex noise.

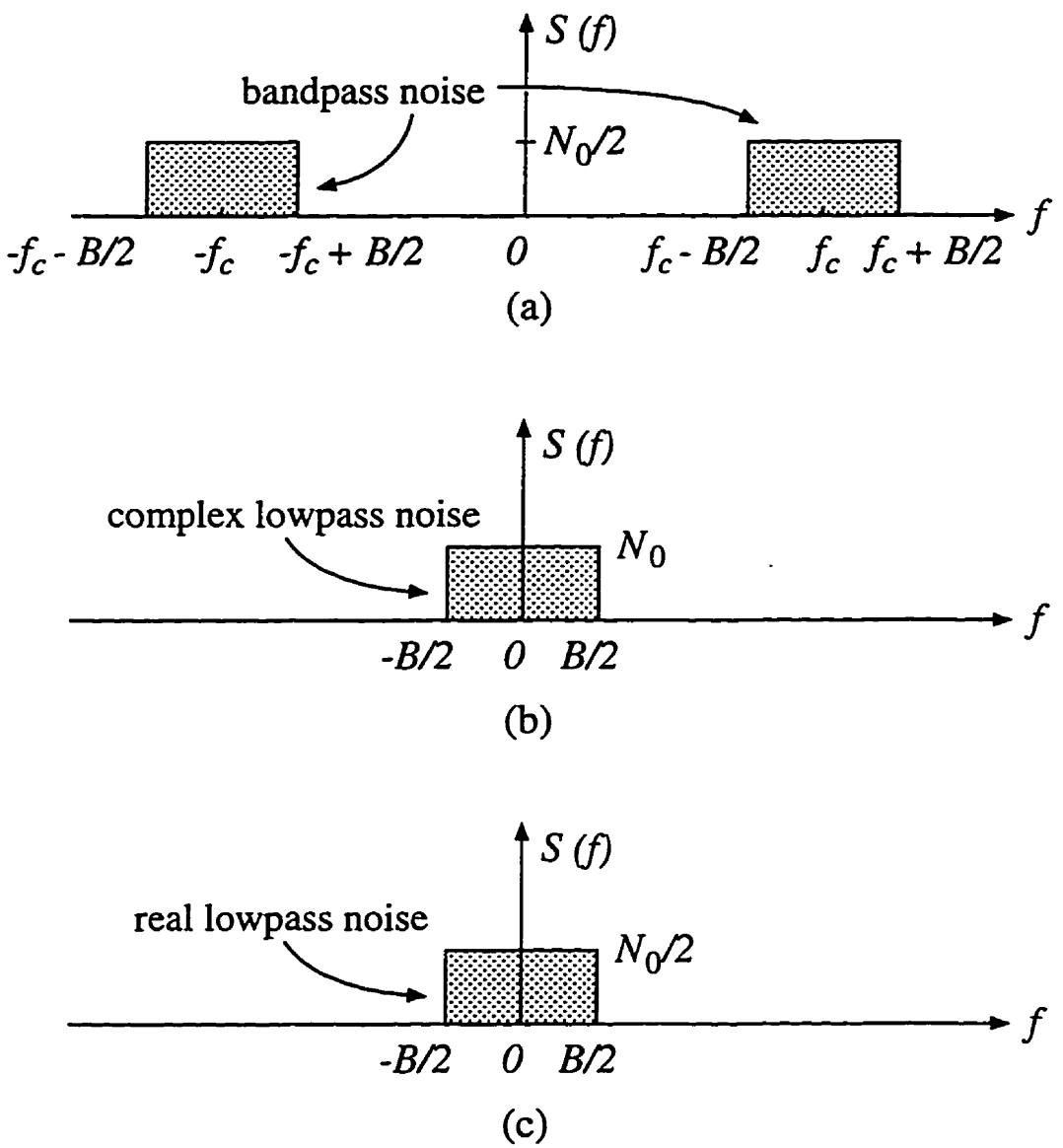

Figure A.1: PSD of (a) bandpass white Gaussian noise, (b) lowpass-equivalent complex Gaussian noise, and (c) lowpass-equivalent real Gaussian noise.

3. In a system where a real signal (such as a BPSK-modulated signal at baseband) is transmitted and corrupted by real AWGN, it is necessary to use a PSD for the noise, illustrated by Figure A.1(c), which has an amplitude of $N_0/2$. The real-valued signal and real-valued noise of this system effectively can be considered to exist solely in either the in-phase or quadrature channel of the complex lowpass equivalent; that is, in this real case, the lowpass-equivalent of the noise is still complex, but the portion of the noise that is orthogonal to the signal is ignored. This PSD is used in Section 2.2 of Chapter 2 to show that the relationship between the SNR, ρ, and E_b/N_0 is

$$\rho = 2\frac{E_b}{N_0}$$

for real, antipodal signals in real AWGN.

4. Note that, in this appendix, the energy and autocorrelation functions of complex quantities employ a factor of 1/2 (see, for example, (A.8)) whereas the energy and autocorrelation functions of real quantities do not (see, for example, (A.28)). In the main text of Chapters 2 and 3, the factor of 1/2 that appears in the computation of the autocorrelations of complex quantities is dropped (see, for example, (2.8b) of Chapter 2) so that the autocorrelation expressions given are applicable to both real and complex signals without regard to the factor of 1/2. Strictly speaking, the factor of 1/2 is required in the complex case in order to maintain consistency between the bandpass and lowpass-equivalent forms as is evident by (A.5), for example; however, whether or not the factor of 1/2 is employed makes no difference when computing the *ratio* of signal power to noise power as long as the use of the factor of 1/2 is consistent among the signal and noise expressions. Further, if the derivation of the probability of error given in Section A.2 for signals in complex AWGN is performed dropping the factors of 1/2 that appear in the energy and autocorrelation expressions, the exact same end result is obtained for the probability of error.

Appendix B

Design of Root-Raised Cosine (RRC) Filter Based on the Frequency Sampling Technique

A method is outlined below to synthesize the FIR filter tap coefficients (discrete impulse response) of a digital filter from a desired analog frequency response using the frequency sampling technique [126]. This method is used to find the tap coefficients of a RRC filter implemented in Chapters 2, 3, and 4 as a pulse-shaping filter in the transmitter and a MF in the receiver.

Consider an analog frequency response, $H(e^{j\omega})$, which is to be approximated by an FIR filter. Let H_k denote samples of the analog frequency response as

$$H_k = H(e^{j\omega})|_{\omega=\omega_k}$$

where $\omega_k = \frac{2\pi k}{T_S L}$, $k \in \{0, 1, \ldots, L-1\}$, L is the number of frequency samples (and the number of tap coefficients generated by the frequency sampling technique), and T_S is the sample period ($1/T_S$ is the sampling frequency). Represent H_k in polar form as

$$H_k = A_k e^{-j\theta_k}$$

where A_k is the magnitude of H_k, and θ_k is the phase.

FIR filters can be designed to have linear phase [126] if θ_k is constrained to be

$$\theta_k = \left(\frac{L-1}{2}\right)\left(\frac{2\pi k}{L}\right) = \frac{L-1}{L}\pi k.$$

The impulse response (or tap coefficients), h_n, may be found by taking the inverse discrete Fourier transform (IDFT) of H_k as

$$h_n = \frac{1}{L} \sum_{k=0}^{L-1} H_k e^{j\frac{2\pi}{L}kn}. \tag{B.1}$$

For h_n to be real, make

$$H_k = (H_{L-k})^*$$

where $*$ denotes the complex conjugate.

The design procedure may be summarized as follows:

1. Compute L samples of the magnitude of the analog frequency response according to

$$A_k = |H(e^{j\omega})|_{\omega_k = \frac{2\pi k}{T_S L}},\ k \in \{0, 1, \ldots, L-1\}$$

2. For h_n to be real and have linear phase, construct H_k as

$$\left.\begin{array}{rcl} H_k & = & A_k e^{-j\frac{L-1}{L}k\pi} \\ H_{L-k} & = & A_k e^{j\frac{L-1}{L}k\pi} \end{array}\right\} \quad k \in \{0, 1, \ldots, M\}$$

where

$$M = \begin{cases} \frac{L}{2} - 1, & L \text{ even} \\ \frac{L-1}{2}, & L \text{ odd.} \end{cases}$$

3. Compute the IDFT of H_k using either (B.1) above, or an FFT algorithm [126, 71].

4. In the time domain, apply a rectangular window of appropriate width, or some other window, to h_n to ensure that the impulse response goes smoothly to zero at the tails.

As an example of the frequency sampling technique, the method described above is applied below to the synthesis of an FIR RRC filter. The magnitude of the full raised-cosine frequency response, $A_{\mathrm{RC}}(\omega)$, is given in [65] as

$$A_{\mathrm{RC}}(\omega) = \begin{cases} T & 0 \le |\omega| \le \frac{(1-\alpha)\pi}{T} \\ \frac{T}{2}\left[1 - \sin\left(\frac{T}{2\alpha}(\omega - \frac{\pi}{T})\right)\right] & \frac{(1-\alpha)\pi}{T} \le |\omega| \le \frac{(1+\alpha)\pi}{T} \end{cases} \tag{B.2}$$

where T is the baud rate, and α is the rolloff factor. The magnitude response of the RRC filter is simply the square root of (B.2) expressed as

$$A_{\sqrt{\text{RC}}}(\omega) = \sqrt{A_{\text{RC}}(\omega)}. \tag{B.3}$$

A sample frequency of

$$F_S = \frac{1}{T_S} = \frac{N_{ss}}{T}$$

is chosen where N_{ss} is the number of samples per symbol. The analog magnitude response, given by (B.3), is sampled at the discrete frequencies specified by

$$\omega_k = 2\pi \frac{N_{ss}}{T} \frac{k}{L}, \; k \in \{0, 1, \ldots, L-1\}.$$

For $N_{ss} = 16$, $L = 128$ taps, and $\alpha = 0.5$, the sampled magnitude response,

$$A_{\sqrt{\text{RC}},k} = A_{\sqrt{\text{RC}}}(\omega)\Big|_{\omega=\omega_k},$$

appears as in Figure B.1 which shows the discrete frequency response up to half the Nyquist rate ($F_S/2 = N_{ss}/(2T) = 8/T$). The corresponding impulse response is displayed in Figure B.2(a). Notice that the tails of the impulse response are not zero. A rectangular window (applied in the time-domain) can be used to truncate the first and last few samples in order to have the tails of the impulse response go to zero as shown in Figure B.2(b). A more gradual window may be a better choice to avoid the resulting discontinuity of the truncation process.

As a check on the results, the discrete self-convolution of the RRC impulse response should yield the impulse response of the full raised-cosine filter. From Figure B.3, it can be seen that the amplitude of the impulse response is zero for every $N_{ss} = 16$ samples (each sample is represented by a vertical line) which indicates that this impulse response satisfies Nyquist's criterion for intersymbol-interference-free transmission, thus giving confidence in the tap coefficients obtained for the RRC filter.

Note that an alternate and easier method to evaluate the tap coefficients is to use samples of the continuous-time impulse-response, $h(t)$, which is given in equation (6)

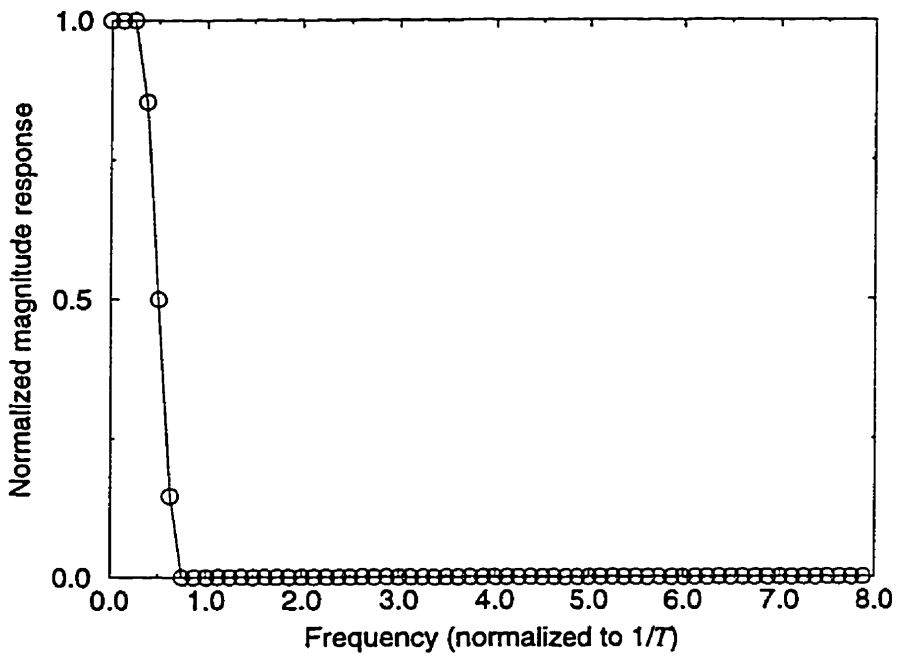

Figure B.1: Magnitude response of RRC filter with $N_{ss} = 16$, $L = 128$ taps, and $\alpha = 0.5$.

of [28] for the root raised-cosine filter as

$$h(t) = \begin{cases} 1 - \alpha + 4\alpha/\pi, & t = 0 \\ \frac{\alpha}{\sqrt{2}}[(1+\frac{2}{\pi})\sin(\frac{\pi}{4\alpha}) + (1-\frac{2}{\pi})\cos(\frac{\pi}{4\alpha})], & t = \pm\frac{T}{4\alpha} \\ \frac{\sin[\pi(1-\alpha)\frac{t}{T}]+4\alpha\frac{t}{T}\cos[\pi(1+\alpha)\frac{t}{T}]}{\pi\frac{t}{T}[1-(4\alpha\frac{t}{T})^2]}, & \text{otherwise.} \end{cases}$$

Other published expressions of the continuous-time root raised-cosine impulse response (see page 283, equation (4.3.37) of [127]) have a factor of $\pi t/\sqrt{T}$ in the denominator instead of $\pi t/T$. This difference only scales the impulse response by a constant.

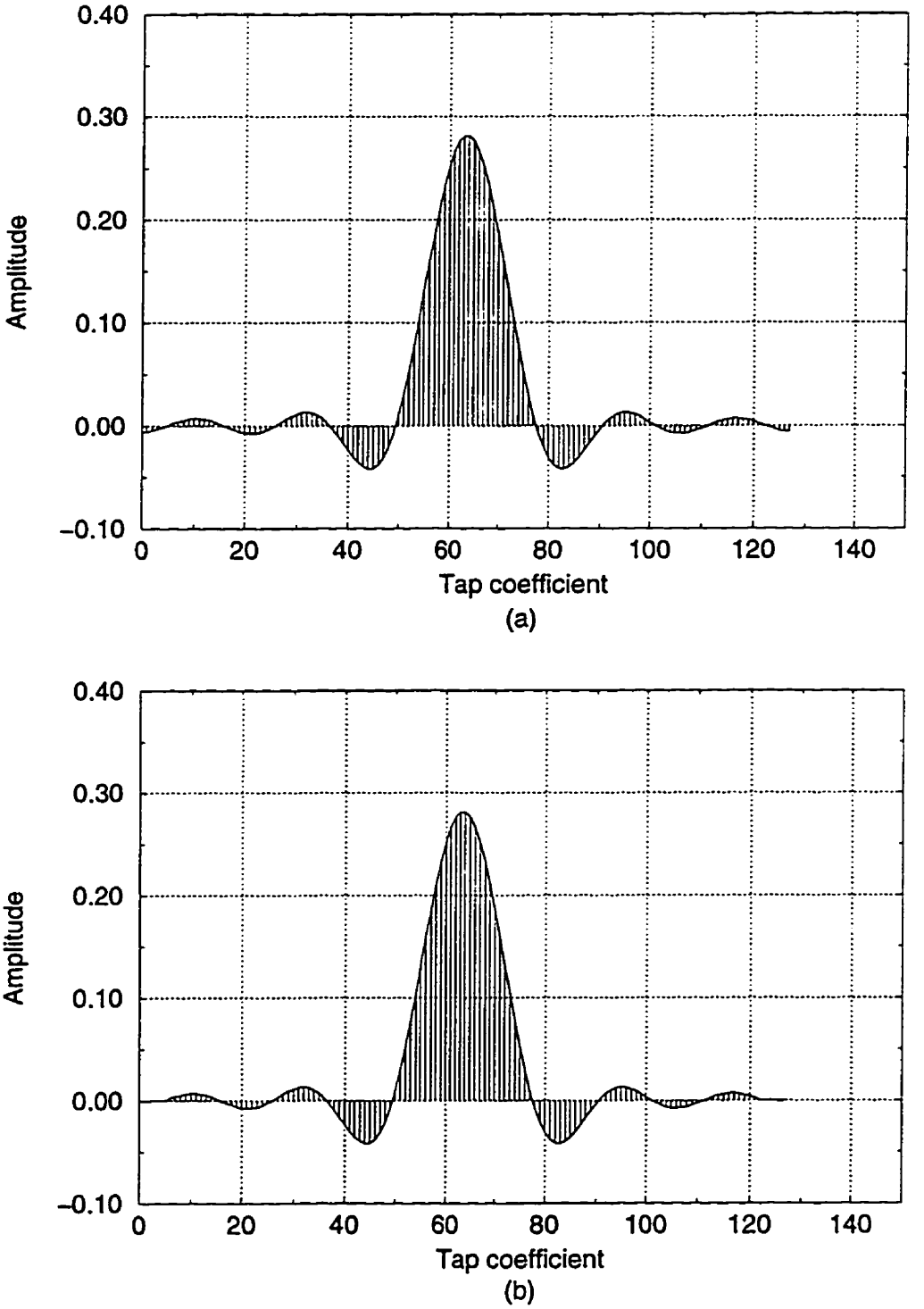

Figure B.2: Impulse response of RRC FIR filter with $N_{ss} = 16$, $L = 128$ taps, and $\alpha = 1$ (a) before application of rectangular window, (b) after application of rectangular window.

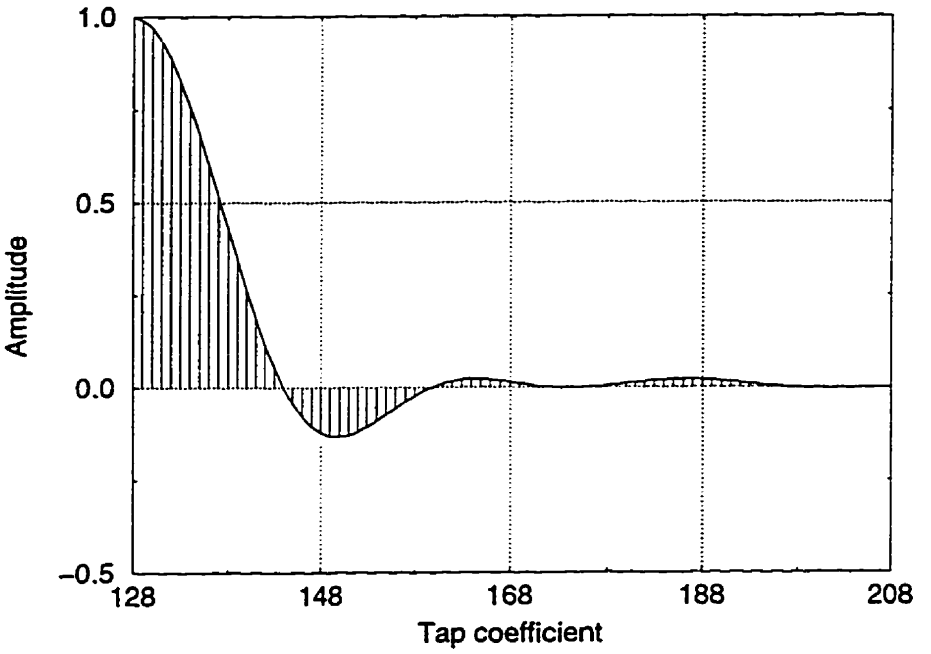

Figure B.3: Self-convolution of RRC impulse response to yield full raised-cosine impulse response.

Appendix C

The Cramér-Rao Bound (CRB)

The derivation of the CRB for biased estimators presented below is based on [128] and [47]. Errors in some intermediate steps of the derivation in [47] have been corrected.

Consider some observable random variable, x, whose value is dependent on a set of K fixed, unknown parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_K)$. Assume M such observations are available to form the observation vector $\mathbf{x} = (x_1, x_2, \ldots, x_M)$. The objective is to find the lower bound on the variance of any estimator, $\hat{\boldsymbol{\theta}}(\mathbf{x})$, which generates estimates of $\boldsymbol{\theta}$ based on the observations, $\mathbf{x}$.

Let the pdf of $\mathbf{x}$ conditioned on the unknown parameters, $\boldsymbol{\theta}$, be denoted by $p(\mathbf{x}|\boldsymbol{\theta})$ and let the likelihood function be expressed as

$$\Gamma(\boldsymbol{\theta}) = \ln(p(\mathbf{x}|\boldsymbol{\theta})),$$

emphasizing that $\boldsymbol{\theta}$ is the variable of the likelihood function.

Let the elements of the information matrix [128] be defined as

$$\begin{aligned} J_{ij} &\equiv E\left\{\frac{\partial}{\partial\theta_i}\Gamma(\boldsymbol{\theta})\frac{\partial}{\partial\theta_j}\Gamma(\boldsymbol{\theta})\right\} \\ &= -E\left\{\frac{\partial^2\Gamma(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\} \end{aligned} \tag{C.1}$$

where $i, j \in \{1, 2, \ldots, K\}$. Let

$$E\{\hat{\theta}_i\} = g_i(\boldsymbol{\theta}) = \theta_i + b_i(\boldsymbol{\theta})$$

where $b_i(\theta)$ represents the bias of $\hat{\theta}_i(\mathbf{x})$. Let

$$\Delta_{ij} = \frac{\partial g_i}{\partial \theta_j}$$

where, again, $i, j \in \{1, 2, \ldots, K\}$. Finally, let

$$V = V_{ij}, \; i, j \in \{1, 2, \ldots, K\}$$

represent the covariance matrix of θ_i and θ_j where the elements of V are given by

$$V_{ij} = E\{(\hat{\theta}_i - g_i)(\hat{\theta}_j - g_j)\}.$$

It can be shown that $V - \Delta J^{-1} \Delta^T$ is a non-negative definite matrix where J^{-1} is the inverse of J, and Δ^T is the transpose of Δ. Considering, then, the diagonal elements of $V - \Delta J^{-1} \Delta^T$, the diagonal elements of V may be written

$$V_{ii} \geq \sum_{m=1}^{K} \sum_{n=1}^{K} \frac{\partial g_i}{\partial \theta_m} (J^{-1})_{mn} \frac{\partial g_i}{\partial \theta_n}. \tag{C.2}$$

Equation (C.2) represents a lower bound, the CRB, on the variance of $\hat{\theta}_i$. Consider the special case of $i = 1$ and $K = 2$, and let the bias of θ_1 be independent of θ_2 so that

$$\frac{\partial g_1}{\partial \theta_2} = 0$$

and

$$V_{11} \geq \left(\frac{\partial g_1}{\partial \theta_1} \right)^2 (J^{-1})_{11}. \tag{C.3}$$

Equation (C.1), written out in matrix form, is

$$J = \begin{bmatrix} -E\left\{\frac{\partial^2 \Gamma}{\partial \theta_1^2}\right\} & -E\left\{\frac{\partial^2 \Gamma}{\partial \theta_1 \partial \theta_2}\right\} \\ -E\left\{\frac{\partial^2 \Gamma}{\partial \theta_1 \partial \theta_2}\right\} & -E\left\{\frac{\partial^2 \Gamma}{\partial \theta_2^2}\right\} \end{bmatrix}.$$

For a matrix of the form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

the inverse is given by

$$\frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

so that

$$(J^{-1})_{11} = \frac{-E\left\{\frac{\partial^2\Gamma}{\partial\theta_2^2}\right\}}{E\left\{\frac{\partial^2\Gamma}{\partial\theta_1^2}\right\}E\left\{\frac{\partial^2\Gamma}{\partial\theta_2^2}\right\} - E\left\{\frac{\partial^2\Gamma}{\partial\theta_1\partial\theta_2}\right\}^2}.$$

After substitution into (C.3), the CRB on the variance of the estimator of $\hat{\theta}_1$ may be expressed as

$$V_{11} \geq \frac{-\left(\frac{\partial g_1}{\partial\theta_1}\right)^2 E\left\{\frac{\partial^2\Gamma}{\partial\theta_2^2}\right\}}{E\left\{\frac{\partial^2\Gamma}{\partial\theta_1^2}\right\}E\left\{\frac{\partial^2\Gamma}{\partial\theta_2^2}\right\} - E\left\{\frac{\partial^2\Gamma}{\partial\theta_1\partial\theta_2}\right\}^2}. \tag{C.4}$$

Appendix D

Confidence Interval for the Variance of an Estimator

Consider a random variable, $\hat{\rho}$, which has an unknown mean, unknown variance, and unknown pdf. The problem is to estimate the variance of $\hat{\rho}$ to within a particular error with a certain degree of confidence.

A confidence interval cannot be stated without knowledge of the pdf, but it is possible to take advantage of the Central Limit Theorem (CLT) [65, 66] which states that the distribution of the sum of a sufficiently large number of statistically independent and identically distributed random variables approaches a Gaussian distribution [64]. An approximate confidence interval can be constructed by using standard techniques for a Gaussian random variable with unknown mean and unknown variance [129, 130]. Specifically, the Student-t distribution is used below to derive an approximate confidence interval for the variance of a random variable.

Consider a collection of N sets of estimates of the parameter, ρ, each set consisting of K individual estimates:

$$\begin{array}{cccc}
\hat{\rho}_{00} & \hat{\rho}_{10} & \cdots & \hat{\rho}_{N-1,0} \\
\hat{\rho}_{01} & \hat{\rho}_{11} & \cdots & \hat{\rho}_{N-1,1} \\
\vdots & \vdots & & \vdots \\
\underbrace{\hat{\rho}_{0,K-1}} & \underbrace{\hat{\rho}_{1,K-1}} & \cdots & \underbrace{\hat{\rho}_{N-1,K-1}} \\
\bar{\hat{\rho}}_0 = X_0 & \bar{\hat{\rho}}_1 = X_1 & \cdots & \bar{\hat{\rho}}_{N-1} = X_{N-1} \\
s^2_{\hat{\rho}_0} = Y_0 & s^2_{\hat{\rho}_1} = Y_1 & \cdots & s^2_{\hat{\rho}_{N-1}} = Y_{N-1}
\end{array}$$

where $\bar{\hat{\rho}}_i = X_i$ is the sample mean defined as

$$\bar{\hat{\rho}}_i = \frac{1}{K} \sum_{j=0}^{K-1} \hat{\rho}_{ij}, \; i \in \{0, 1, \ldots, N-1\}, \tag{D.1}$$

and $s^2_{\hat{\rho}_i} = Y_i$ is the unbiased sample variance defined as

$$s^2_{\hat{\rho}_i} = \frac{1}{K-1} \sum_{j=0}^{K-1} \hat{\rho}^2_{ij} - \frac{K}{K-1} (\bar{\hat{\rho}}_i)^2, \; i \in \{0, 1, \ldots, N-1\}. \tag{D.2}$$

If K is large enough ($K \geq 10$) the distribution of X_i is approximately normal due to the summation in (D.1). The distribution of Y_i is also approximately normal since the second term in (D.2) approaches a constant. The random variable, Y_i, has an unknown mean, unknown variance, and a pdf that is Gaussian. Further, it is assumed that each random variable, Y_i, $i \in \{0, 1, \ldots, N-1\}$, has identical statistics.

The sample mean of Y_i (or, equivalently, the sample mean of the sample variance of $\hat{\rho}_i$) is given by

$$\bar{Y} = \frac{1}{N} \sum_{i=0}^{N-1} Y_i, \tag{D.3}$$

computed using N trials of K estimates each. Further, let

$$s^2_Y = \frac{1}{N-1} \sum_{n=0}^{N-1} Y_i^2 - \frac{N}{N-1} \bar{Y}^2$$

where s^2_Y provides a measure of the variability of the sample variance of $\hat{\rho}$ from trial to trial. As N increases, $\bar{Y}$ and s^2_Y become more reliable in the sense that

$$\bar{Y} \rightarrow \mu_Y$$

and

$$s^2_Y \rightarrow \sigma^2_Y$$

where μ_Y is the true mean of Y (that is, the true variance of $\hat{\rho}$ which is sought), and σ^2_Y is the true variance of Y. If $N \geq 10$, $\bar{Y}$ is approximately normally distributed with unknown true mean, μ_Y, and unknown true variance, σ^2_Y, and it is possible to proceed with the derivation of a confidence interval based on standard techniques for a Gaussian random variable with unknown mean and unknown variance [129, 130].

Note that, given $N \times K$ samples of a random variable $\hat{\rho}_{ij}$, a direct expression for the unbiased sample variance is

$$s_{\hat{\rho}}^2 = \frac{1}{NK-1} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} \hat{\rho}_{ij}^2 - \frac{NK}{NK-1} (\bar{\hat{\rho}})^2 \tag{D.4}$$

where

$$\bar{\hat{\rho}} = \frac{1}{NK} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} \hat{\rho}_{ij}. \tag{D.5}$$

For the unbiased sample variance given by (D.4), $E\{s_{\hat{\rho}}^2\} = \sigma^2$, where σ^2 is the true variance of $\hat{\rho}_{ij}$. Following the example on p. 189 of [65], it is shown below that the expression given by (D.3) is also an unbiased estimator for σ^2.

Re-write (D.2) as

$$Y_i = \frac{1}{K-1} \sum_{j=0}^{K-1} (\hat{\rho}_{ij} - \bar{\hat{\rho}}_i)^2 \tag{D.6}$$

where $\bar{\hat{\rho}}_i$ is given by (D.1). Substituting this expression into (D.3), and taking expected values, one obtains

$$\begin{aligned} E\{\bar{Y}\} &= \frac{1}{N(K-1)} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} E\{(\hat{\rho}_{ij} - \bar{\hat{\rho}}_i)^2\} \\ &= \frac{1}{N(K-1)} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} E\left\{\left[(\hat{\rho}_{ij} - \mu) - (\bar{\hat{\rho}}_i - \mu)\right]^2\right\} \\ &= \frac{1}{N(K-1)} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} \Big[E\{(\hat{\rho}_{ij} - \mu)^2\} - 2E\{(\hat{\rho}_{ij} - \mu)(\bar{\hat{\rho}}_i - \mu)\} \\ &\qquad + E\{(\bar{\hat{\rho}}_i - \mu)^2\}\Big]. \end{aligned} \tag{D.7}$$

The first term in (D.7) is the true variance by definition; that is, $E\{(\hat{\rho}_{ij} - \mu)^2\} = \sigma^2$. The second term may be evaluated as

$$\begin{aligned} E\{(\hat{\rho}_{ij} - \mu)(\bar{\hat{\rho}}_i - \mu)\} &= E\left\{(\hat{\rho}_{ij} - \mu)\left(\frac{1}{K}\sum_{j=0}^{K-1} \hat{\rho}_{ij} - \mu\right)\right\} \\ &= \frac{1}{K} E\{(\hat{\rho}_{ij} - \mu)[(\hat{\rho}_{i0} - \mu) + (\hat{\rho}_{i1} - \mu) + \cdots + (\hat{\rho}_{i,K-1} - \mu)]\} \\ &= \frac{1}{K} E\{(\hat{\rho}_{ij} - \mu)(\hat{\rho}_{ij} - \mu)\} \\ &= \frac{\sigma^2}{K} \end{aligned}$$

because $E\{(\hat{\rho}_{ij}-\mu)(\hat{\rho}_{ik}-\mu)\}=0$ for $j\neq k$. The last term in (D.7) may be evaluated as

$$\begin{aligned}
E\{(\bar{\hat{\rho}}_i-\mu)^2\} &= E\left\{\left(\frac{1}{K}\sum_{j=0}^{K-1}\hat{\rho}_{ij}-\mu\right)\left(\frac{1}{K}\sum_{j=0}^{K-1}\hat{\rho}_{ij}-\mu\right)\right\} \\
&= \frac{1}{K^2}E\{[(\hat{\rho}_{i0}-\mu)+(\hat{\rho}_{i1}-\mu)+\cdots+(\hat{\rho}_{i,K-1}-\mu)] \\
&\qquad \times[(\hat{\rho}_{i0}-\mu)+(\hat{\rho}_{i1}-\mu)+\cdots+(\hat{\rho}_{i,K-1}-\mu)]\} \\
&= \frac{1}{K^2}E\{(\hat{\rho}_{i0}-\mu)^2+(\hat{\rho}_{i1}-\mu)^2+\ldots+(\hat{\rho}_{i,K-1}-\mu)^2\} \\
&= \frac{\sigma^2}{K}.
\end{aligned}$$

Combining terms, the expected value of $\bar{Y}$ is found as

$$\begin{aligned}
E\{\bar{Y}\} &= \frac{1}{N(K-1)}\sum_{i=0}^{N-1}\sum_{j=0}^{K-1}\left(\sigma^2-2\frac{\sigma^2}{K}+\frac{\sigma^2}{K}\right) \\
&= \frac{1}{N(K-1)}\sigma^2N(K-1) \\
&= \sigma^2
\end{aligned}$$

so that (D.3) is an unbiased estimator for the variance of $\hat{\rho}_{ij}$.

Assume $\bar{Y}$ is indeed normally distributed with unknown mean, μ_Y, and unknown variance, $\sigma_{\bar{Y}}^2$. The Student-t distribution is applicable in this case [129]. Define the normalized parameter, t, as

$$t=\frac{\bar{Y}-\mu_Y}{s_Y/\sqrt{N}}.$$

The confidence interval is expressed as

$$P\left\{-t_{\alpha/2}\leq\frac{\bar{Y}-\mu_Y}{s_Y/\sqrt{N}}\leq t_{\alpha/2}\right\}=1-\alpha$$

or, rewriting in terms of $\bar{Y}$,

$$P\left\{\bar{Y}-t_{\alpha/2}\frac{s_Y}{\sqrt{N}}\leq\mu_Y\leq\bar{Y}+t_{\alpha/2}\frac{s_Y}{\sqrt{N}}\right\}=1-\alpha.$$

In words, the probability that the true mean, μ_Y, is in the range $\bar{Y}\pm t_{\alpha/2}\,s_Y/\sqrt{N}$ is $(1-\alpha)\times100\%$. The parameter, $t_{\alpha/2}$, is the $\alpha/2$-percentile of the Student-t distribution

with $N-1$ degrees of freedom. The value, K, must be chosen large enough such that

$$\begin{aligned} t_{\alpha/2}\frac{s_Y}{\sqrt{N}} &< \epsilon\bar{Y} \\ \frac{s_Y^2}{\bar{Y}^2} &< \frac{N\epsilon^2}{t_{\alpha/2}^2}. \end{aligned} \tag{D.8}$$

That is, K must be chosen large enough such that the normalized sample variance of Y is less than the quantity specified in (D.8) where $\epsilon \times 100\%$ is the percent error.

As an example, say that it is desired to find K such that $\bar{Y}$ may be assumed to be within 20% error with 95% confidence using $N = 10$ trials (the minimum number required to assume Gaussian statistics). From a table of the Student-t distribution with 9 degrees of freedom (see [129] or [130]), $t_{0.05/2} = 2.262$ so that

$$\frac{s_Y^2}{\bar{Y}^2} < \frac{N\epsilon^2}{t_{\alpha/2}^2} = \frac{10 \times (0.2)^2}{(2.262)^2} = 0.078.$$

Note that this technique does not produce an expression that gives a value for K explicitly. It is necessary to choose K by trial and error until $s_Y^2/\bar{Y}^2$, computed from the collection of $N \times K$ estimates of ρ, yields a number less than $N\epsilon^2/t_{\alpha/2}^2$.

Appendix E

Some Basic Fading Terms and Concepts

A signal propagating through a mobile radio channel experiences amplitude fluctuations, and is "smeared" both in time and frequency. The underlying mechanism of the fading effects is due to the influence of physical topography on the propagation of the RF signal from a transmitter to a receiver. Since the physical topography changes from location to location, so do the fading characteristics. A fading channel can be labeled according to the distribution of the amplitude variations, and the characteristics of the time and frequency dispersion. For example, a channel may be labeled as a fast, frequency-selective Rayleigh fading channel. Comments on each of the three attributes of a fading channel are discussed below based on material found in [65, 24, 75, 95, 121, 105]. It is not the intention of this discussion to provide a comprehensive tutorial of fading theory, but to identify a few basic terms and concepts used in Chapter 3.

E.1 Amplitude Distribution

Due to the presence of many physical scatterers, the received signal is a superposition of echos forming a standing wave pattern[1] and, as a result, the received signal strength is spatially dependent. A moving receiver passes through the peaks and nulls of the standing wave pattern so that the received signal strength, over time, fluctuates in a manner exemplified by the plot shown in Figure E.1 where γ is the amplitude of the envelope of the received signal, T_S is the sampling period, and $\nu_{\max}$ is the *Doppler spread* (the maximum Doppler frequency). The usual assumption is that the received signal consists of a large number of echos, each of which is independent and randomly phased, so that samples, γ_k, of the signal envelope are distributed according to the Rayleigh distribution with cdf given by (see [95])

$$P(\gamma) = 1 - e^{-\gamma^2/\sigma^2}$$

where the rms value of γ is

$$\gamma_{\text{rms}} = \sqrt{E\{\gamma^2\}} = \sigma.$$

Samples of the phase of the composite received signal are uniformly distributed on $[0, 2\pi)$. Other distributions used to describe amplitude variations in particular environments are the Rice distribution [65] and the Nakagami-m distribution [131, 94].

E.2 Frequency Dispersion

Due to the motion of a receiver, the RF frequencies of the various echos making up the composite received signal are Doppler shifted causing frequency dispersion. As mentioned in Section E.1, the maximum Doppler frequency, $\nu_{\max}$, is known as the Doppler spread. The smearing of the received signal over frequency is evident in the *Doppler power spectrum* or *Doppler profile.* For a vertical monopole antenna

[1]For this discussion, the fluctuation of the signal envelope is placed in the context of a fixed base station transmitting to a mobile receiver (forward path), but a similar argument holds for the opposite direction (reverse path).

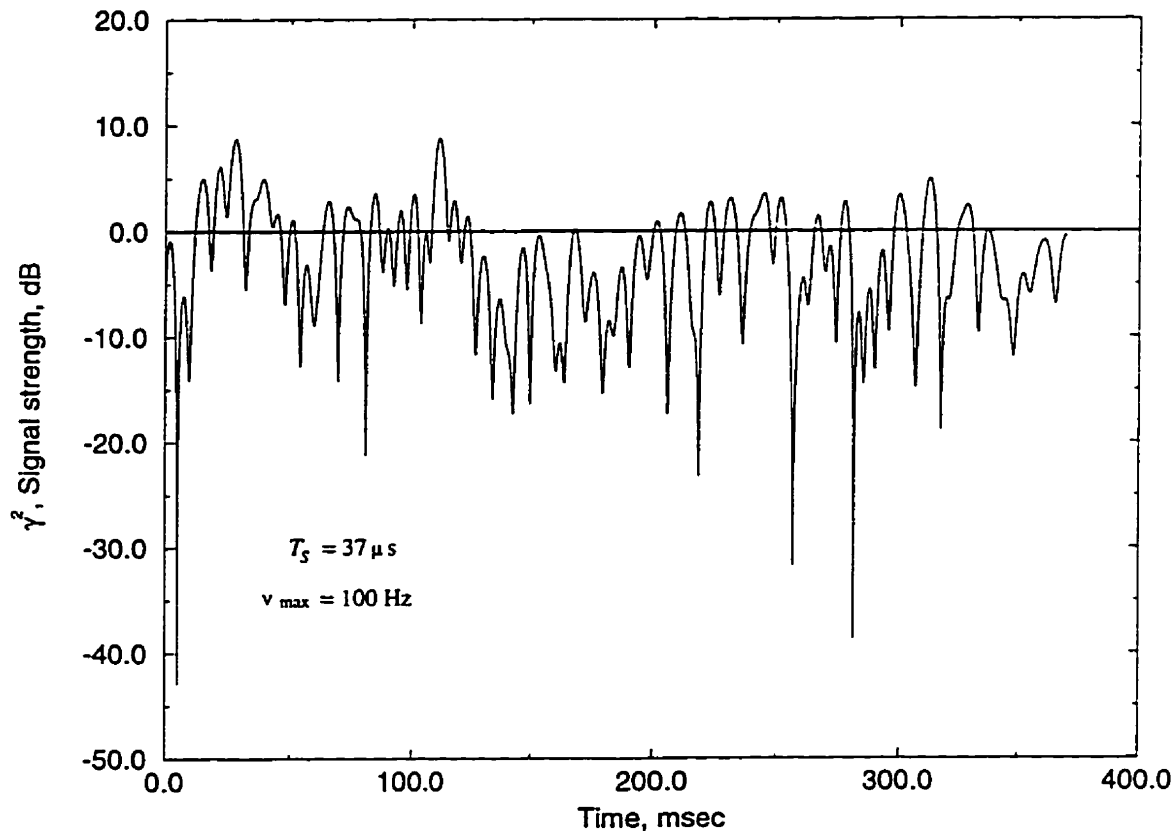

Figure E.1: Typical received signal strength variations due to motion of a receiver through a standing wave pattern. The rms signal level is 0 dB.

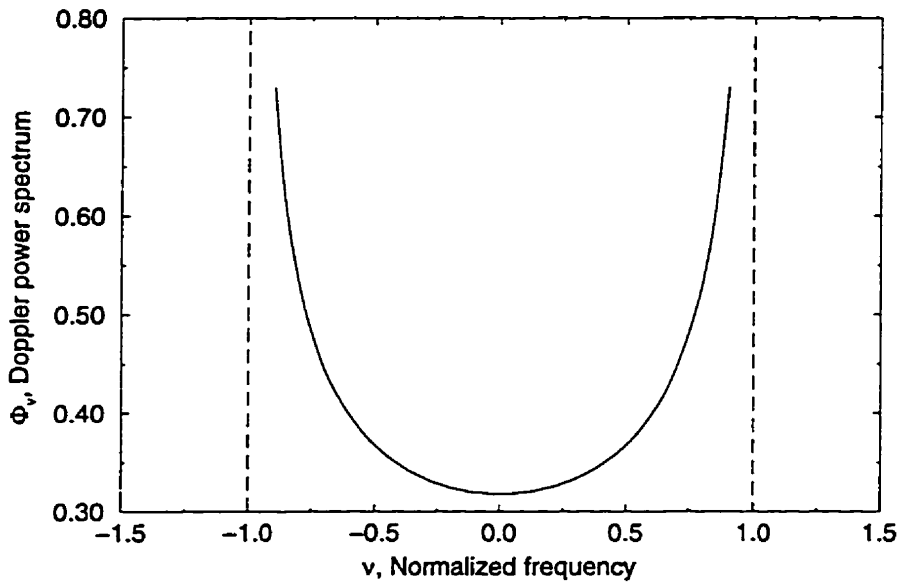

Figure E.2: Jakes Doppler power spectrum normalized to maximum Doppler frequency, $\nu_{\max}$.

in isotropic scattering, the Doppler profile can be modeled by the Jakes spectrum [25, 95]. Let the normalized Doppler power spectrum be denoted by $\Phi_\nu(\nu)$, then the Jakes spectrum may be expressed as

$$\Phi_\nu(\nu) = \frac{1}{\pi\nu_{\max}\sqrt{1-(\nu/\nu_{\max})^2}}, \quad |\nu| \leq \nu_{\max}$$

which has the classic "bathtub" shape shown in Figure E.2. As $\nu_{\max}$ increases, so does the rapidity of the amplitude and phase variations of the received signal. *Slow fading* refers to the case where $\nu_{\max}$ is so small that the amplitude and phase fluctuations of the received signal are negligible over a given observation period. *Fast fading* refers to the case where the amplitude and phase variations occur over a period of time comparable to, or smaller than, a symbol period. In the literature, this term is also applied to moderately fast fading channels as well; that is, channels where the

amplitude and phase variations are negligible over a symbol period, but not over a given observation period (spanning a number of symbols).

E.3 Time Dispersion

Echos arriving at the receiver with different delays contribute multipath-induced ISI to the received signal. The *delay power spectrum* or *delay profile* characterizes the time-smearing effect of a particular channel. A typical, idealized delay profile has a shape that decays exponentially with increasing delay [95]. Let the normalized delay power spectrum be denoted by $\Phi_\tau(\tau)$, then the exponentially-decaying delay profile may be expressed as

$$\Phi_\tau(\tau) = \frac{1}{\tau_{\text{rms}}} e^{-\tau/\tau_{\text{rms}}}, \quad \tau \gtrsim \tau_{\text{max}} \tag{E.1}$$

where τ_{rms} is the *rms delay spread* [104] defined as

$$\tau_{\text{rms}} = \sqrt{\bar{\tau^2} - \bar{\tau}^2}$$

where

$$\begin{aligned} \bar{\tau} &= \frac{\sum_k \gamma_k^2 \tau_k}{\sum_k \gamma_k^2} \\ \bar{\tau^2} &= \frac{\sum_k \gamma_k^2 \tau_k^2}{\sum_k \gamma_k^2}, \end{aligned}$$

and γ_k represents samples of the envelope at delay, τ_k. The quantity, $\bar{\tau}$, is known as the *mean excess delay* and is the first central moment of the delay profile [105]. The quantity, τ_{rms} is a measure of the spread of the delay profile about $\bar{\tau}$.

To illustrate, the normalized exponential delay profile with $\tau_{\text{rms}} = 3\,\mu$s is shown in Figure E.3, and τ_{rms} and $\bar{\tau}$ are identified (the normalization is such that the area under the delay profile curve is unity). If the rms delay spread is greater than 10–20% of the symbol period, then some sort of measures (adaptive equalization or

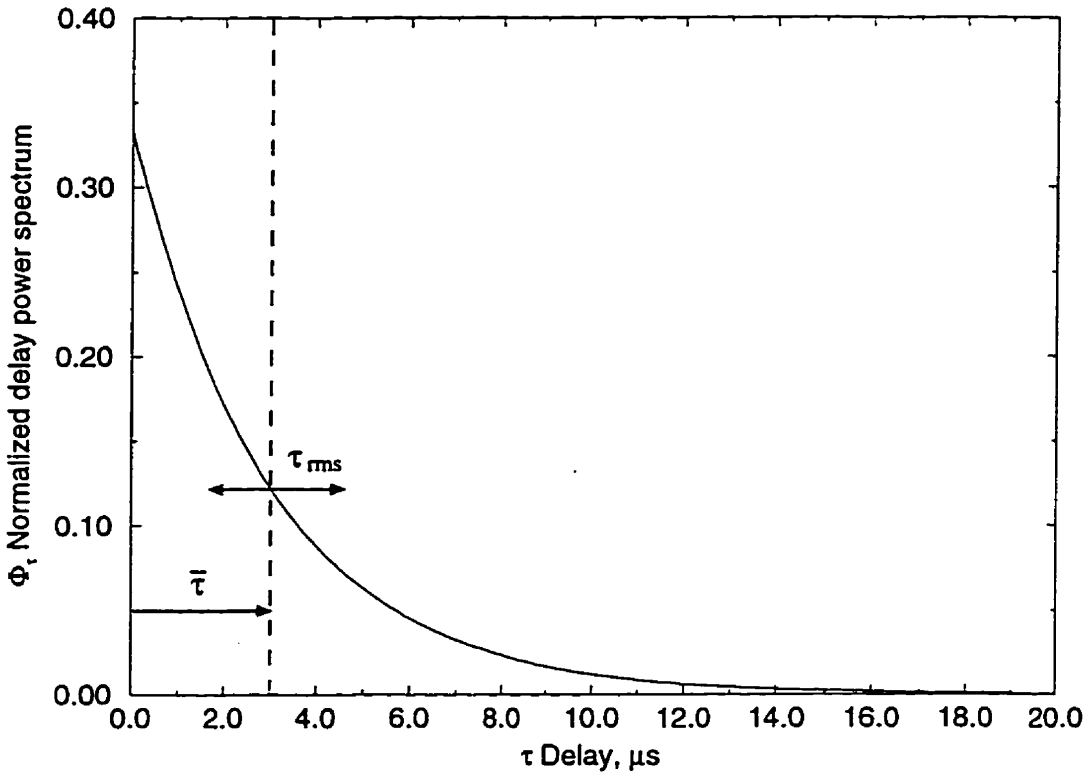

Figure E.3: Exponential delay power spectrum with $\tau_{rms} = 3\,\mu$s.

diversity combining or both) should be taken to counteract the multipath-induced ISI [103, 132]. In cases where the delay becomes comparable to or greater than the symbol period, the channel is said to be *frequency selective.*

Other measures of the delay spread are sometimes used. On page 707 of [65], a parameter called the *multipath spread of the channel* is used to quantify the range of values over which the delay profile is essentially nonzero. A more precise parameter is identified on page 44 of [105] called the *maximum excess delay* at x dB down from maximum which is the earliest value of delay at which the power delay profile first dips below a line x dB down from the maximum of the power delay profile. The rms delay spread, τ_{rms} is most commonly used to characterize multipath channels [105].

Consider two tones (unmodulated sinusoids) of different frequencies transmitted on the same fading channel. The fading experienced by any two tones is perfectly

correlated as long as τ_{rms} is zero or very small. As the rms delay spread gets larger, the fading experienced by the two tones becomes uncorrelated, and the frequency separation between the two tones for which the magnitude of the complex correlation is 0.5 is termed the *correlation* or *coherence bandwidth.*

There is an inverse relationship between the delay spread and the correlation bandwidth. As long as the correlation bandwidth is much larger than the modulation bandwidth (τ_{rms} small), the multipath-induced ISI is negligible and the channel is said to be a *flat fading* channel. When the correlation bandwidth approaches or exceeds the modulation bandwidth (τ_{rms} approaches or exceeds the symbol period), then the channel is said to be a frequency selective channel.

An expression for the magnitude of the correlation of two tones separated by a frequency, $\Delta\omega$, in a frequency-selective fading channel characterized by the exponential delay profile given by (E.1) is (see [95])

$$|\rho_{12}(\Delta\omega)| = \frac{1}{\sqrt{1 + (\Delta\omega\,\tau_{\text{rms}})^2}}.$$

This expression is plotted in Figure E.4 (where $\Delta f = \Delta\omega/2\pi$).

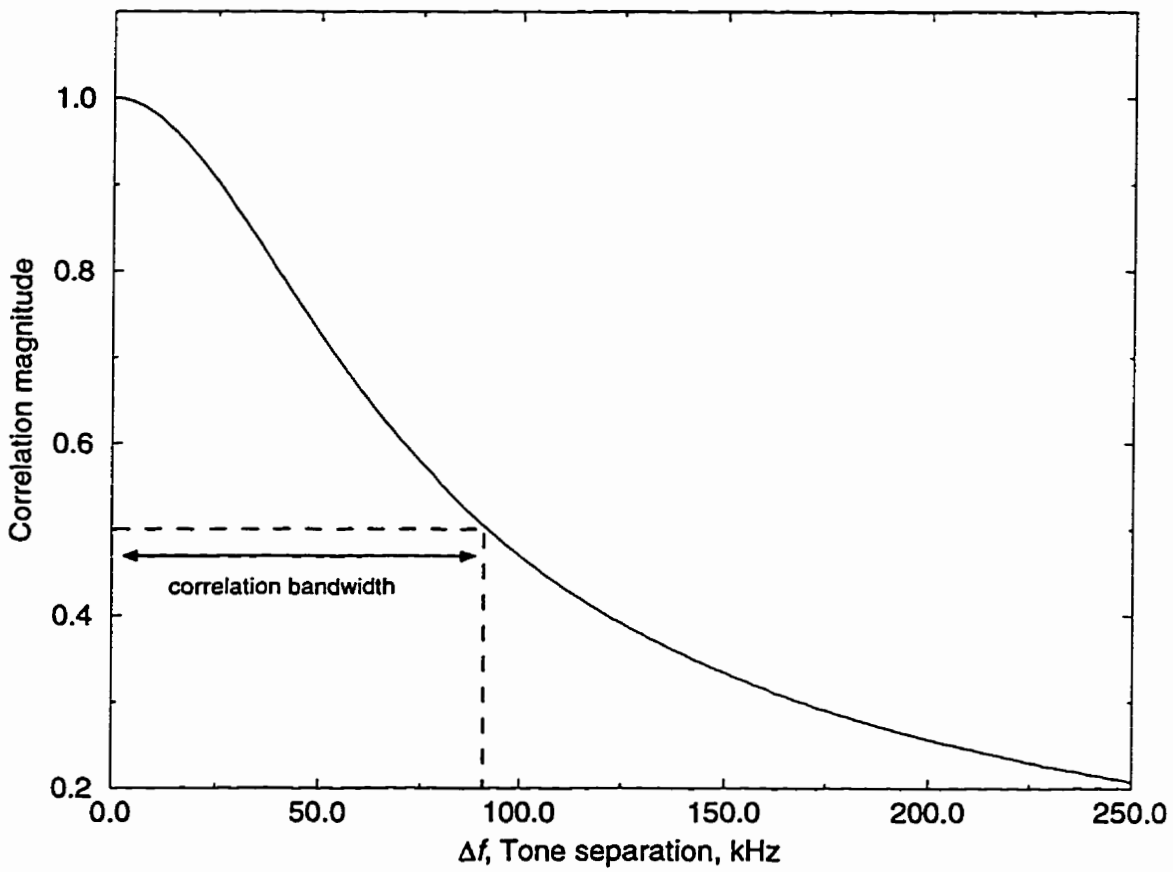

Figure E.4: Correlation of two tones separated by a frequency, Δf, in a channel characterized by an exponential delay profile with $\tau_{\text{rms}} = 3\,\mu\text{s}$.

Appendix F

A Brief Description of $\pi/4$-DQPSK Modulation

Conventional QPSK and DQPSK [65] are modulation techniques which encode information by the translation of binary pairs into symbols. Each symbol represents a phase-shift which may be used to modulate the phases of a complex sinusoid. At baseband, a symbol may be represented by one of four points in a signal constellation. The transitions from one point to another are dictated by a phase mapping rule. Table F.1 shows an example of the translation of binary bit-pairs to symbols represented by an absolute phase for coherent QPSK. Table F.2 shows the phase mapping for DQPSK where each bit-pair maps to a *differential* phase rather than an absolute phase. Assuming NRZ signals are transmitted in an infinite-bandwidth channel, the idealized signal constellation for coherent or differential QPSK appears as shown in Figure F.1. Phase changes of $\pm\pi$ cause transitions through the origin which create large fluctuations in the envelope of the modulated signal in a more re-

Information bits	Absolute phase
00	$\pi/4$
01	$3\pi/4$
10	$-\pi/4$
11	$-3\pi/4$

Table F.1: Mapping of information bit-pairs to absolute phases for coherent QPSK.

Information bits	Differential phase
00	0
01	$+\pi/2$
10	$-\pi/2$
11	$+\pi$

Table F.2: Mapping of information bit-pairs to differential phases for DQPSK.

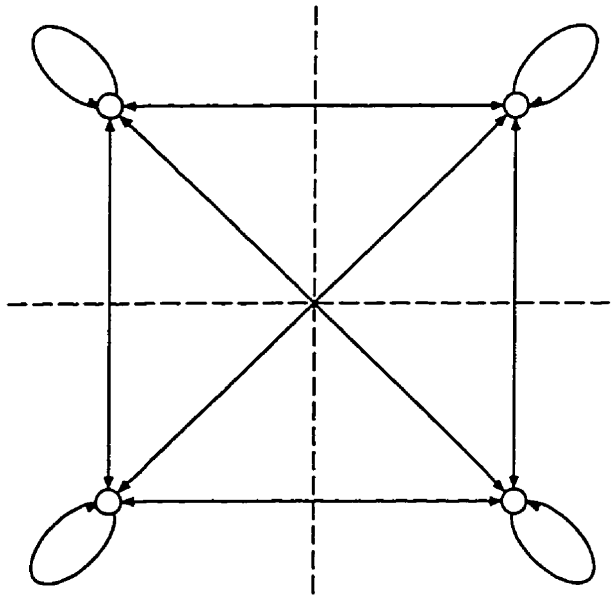

Figure F.1: Signal constellation for conventional QPSK or DQPSK using NRZ signals in an infinite-bandwidth channel.

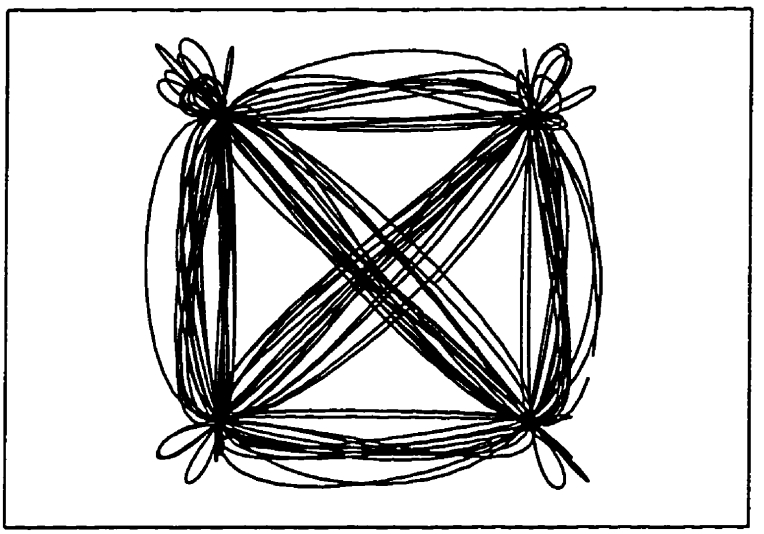

Figure F.2: Signal constellation for conventional QPSK or DQPSK using pseudo-random pulses shaped by raised-cosine filtering.

alistic bandlimited channel. As an example, consider a QPSK signal pulse-shaped by a raised-cosine filter with $\alpha = 0.5$. A sample signal constellation for pseudo-random source data appears in Figure F.2, and the in-phase and quadrature eye-diagrams are displayed in Figures F.3 and F.4, respectively. As a matter of practical concern, these large amplitude fluctuations dictate that a linear power amplifier (PA) be employed in the transmitter, instead of a less expensive non-linear PA, to avoid excessive adjacent-channel interference due to spectral spread [96].

In an attempt to prevent the penalty of spectral spread with the use of a cheaper non-linear PA, a scheme was developed (see [133, 134]) which avoids signal transitions through the origin, thus reducing envelope fluctuations and, theoretically, allowing the use of a non-linear PA. The phase mapping rule for this scheme is given in Table F.3 and an idealized signal constellation (NRZ signals in an infinite-bandwidth channel) is given in Figure F.5.

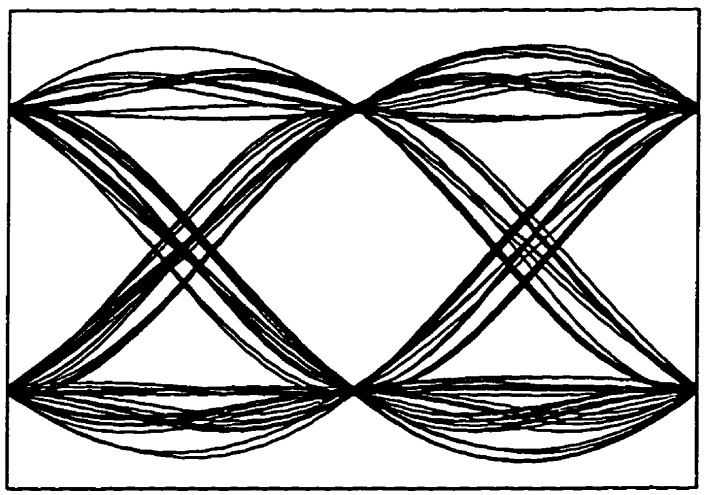

Figure F.3: Eye diagram of in-phase channel for conventional QPSK or DQPSK using pseudo-random pulses shaped by raised-cosine filtering.

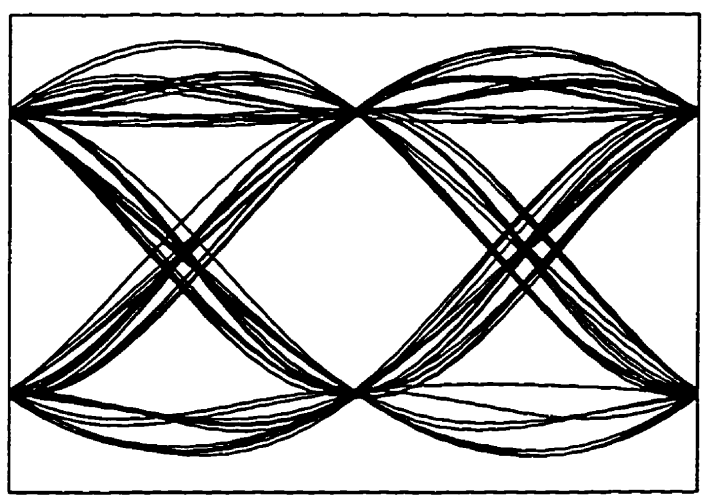

Figure F.4: Eye diagram of quadrature channel for conventional QPSK or DQPSK using pseudo-random pulses shaped by raised-cosine filtering.

Information bits	Differential phase
00	$+\pi/4$
01	$+3\pi/4$
10	$-3\pi/4$
11	$-\pi/4$

Table F.3: Mapping of information bit-pairs to differential phases for $\pi/4$-DQPSK.

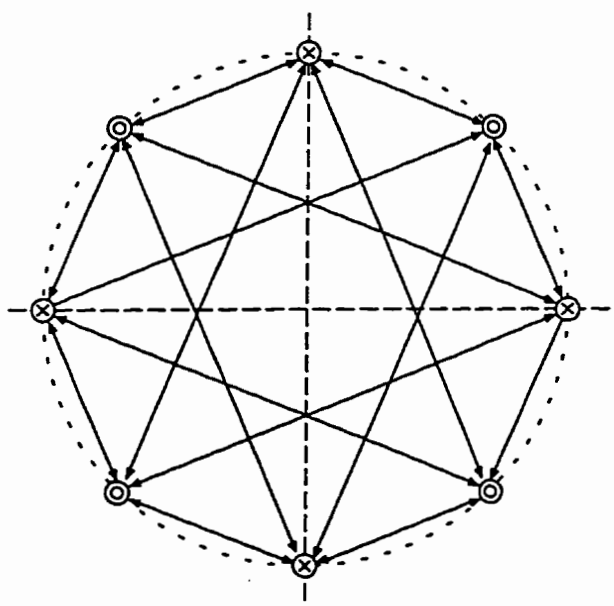

Figure F.5: Signal constellation for $\pi/4$-DQPSK using NRZ signals in an infinite bandwidth channel.

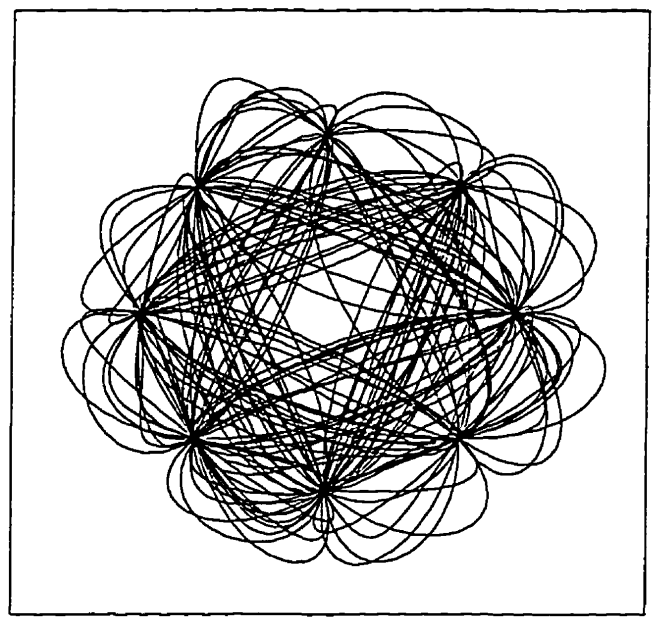

Figure F.6: Signal constellation for $\pi/4$-DQPSK before differential-detection using pseudo-random pulses shaped by raised-cosine filtering.

The signal constellation consists essentially of two conventional sets of QPSK signal points with one set rotated by $\pi/4$ radians with respect to the other. From one symbol interval to another, a constellation point transits from one set of QPSK points to the other; that is, referring to Figure F.5, only transitions from ◎ points to ⊗ points, or vice-versa, are permitted. Stated another way, the phase mapping rule allows successive phase changes of $\pm\pi/4$ and $\pm 3\pi/4$ only, eliminating transitions through the origin and reducing the amount of envelope fluctuation in a bandlimited channel. A sample signal constellation for pseudo-random source data pulse-shaped by a raised-cosine filter appears in Figure F.6, and the in-phase and quadrature eye-diagrams are displayed in Figures F.7 and F.8, respectively.

Comparing Figures F.2 and F.6, it can be seen that the $\pi/4$-DQPSK modulated

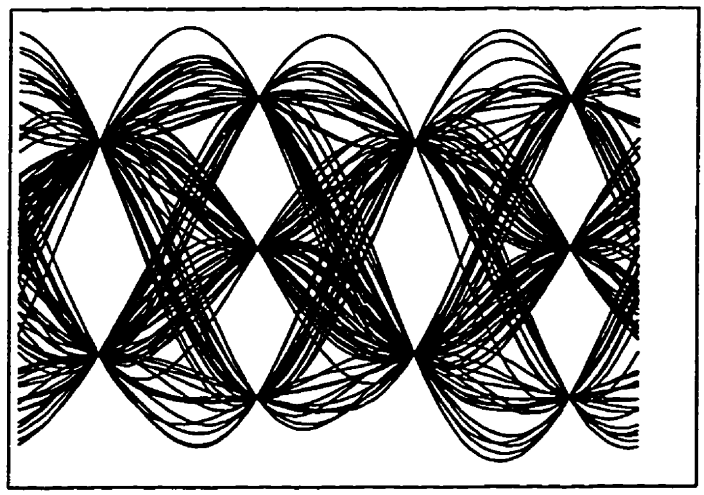

Figure F.7: Eye diagram of in-phase channel for $\pi/4$-DQPSK before differential-detection using pseudo-random pulses shaped by raised-cosine filtering.

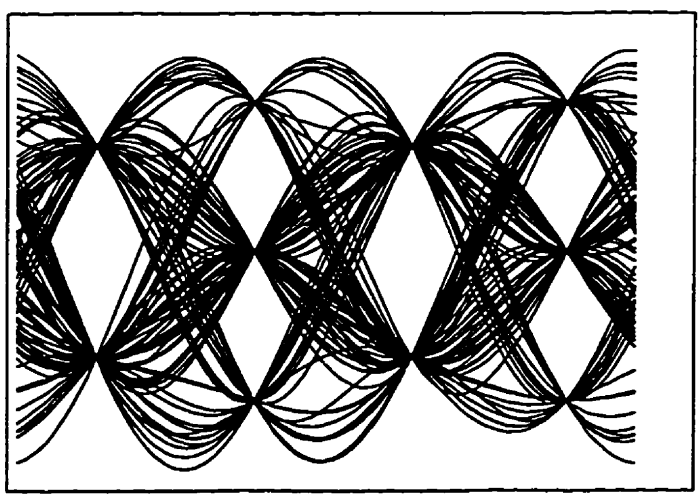

Figure F.8: Eye diagram of quadrature channel for $\pi/4$-DQPSK before differential-detection using pseudo-random pulses shaped by raised-cosine filtering.

signal passes near the origin far less frequently than the conventional QPSK modulated signal, but it is apparent from Figures F.6, F.7, and F.8 that there still is an appreciable amount of fluctuation in the signal envelope of the $\pi/4$-DQPSK signal. Consequently, the theoretical advantage of being able to use a non-linear PA in place of a linear PA is not easily realized in practice.

Note the "2-3-2" pattern of the eye diagrams in Figures F.7 and F.8. This is the eye-diagram corresponding to the $\pi/4$-DQPSK signal constellation phased as shown in Figure F.5. If that signal constellation were rotated by $\pi/8$ radians so that all signal points straddled the x-y axes, then the eye diagrams would display a "4-4" pattern instead.

It is interesting to observe the signal constellation and eye diagrams of the $\pi/4$-DQPSK modulated signal after differential detection. A sample signal constellation is displayed in Figure F.9, and the in-phase and quadrature eye diagrams are shown in Figures F.10 and F.11, respectively.

The signal constellation of Figure F.9 exhibits a curious asymmetry which is a result of the non-linear process of differential-detection applied to this particular modulation scheme. Note that the signal constellation of conventional DQPSK after differential detection doesn't look much different from the signal constellation before differential-detection—at least there is no glaring asymmetry as in the case of differentially-detected $\pi/4$-DQPSK.

The asymmetry of the $\pi/4$-DQPSK signal after differential detection is also visible in the eye diagrams of Figures F.10 and F.11. Note that the quadrature channel doesn't look much different from that of conventional QPSK (compare with Figure F.4), but the in-phase channel is affected on negative transitions. The effect is visible as a cusp between symbol centres.

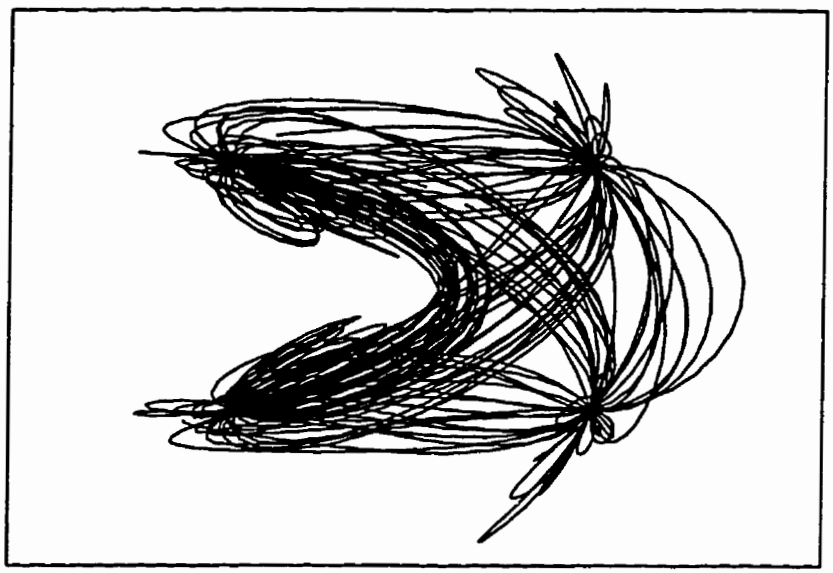

Figure F.9: Signal constellation for $\pi/4$-DQPSK after differential-detection using pseudo-random pulses shaped by raised-cosine filtering.

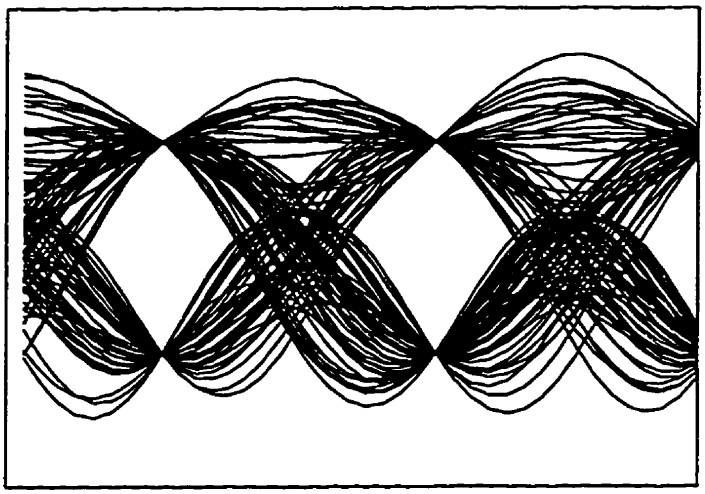

Figure F.10: Eye diagram of in-phase channel for $\pi/4$-DQPSK after differential-detection using pseudo-random pulses shaped by raised-cosine filtering.

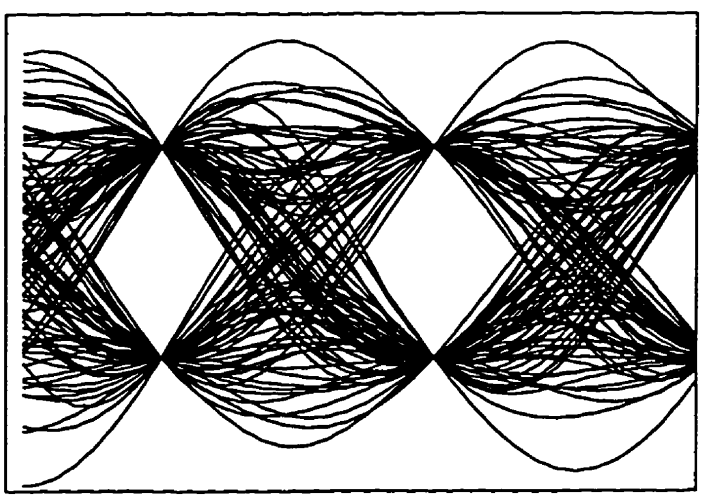

Figure F.11: Eye diagram of quadrature channel for $\pi/4$-DQPSK after differential-detection using pseudo-random pulses shaped by raised-cosine filtering.

References

[1] N. E. Nahi and R. M. Gagliardi, "Use of limiters for estimating signal to noise ratio," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 127–129, Jan. 1967.

[2] F. Adachi and K. Ohno, "BER performance of QDPSK with postdetection diversity reception in mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 237–249, Feb. 1991.

[3] A. J. Goldsmith, L. J. Greenstein, and G. J. Foschini, "Error statistics of real-time power measurements in cellular channels with multipath and shadowing," *IEEE Trans. Veh. Technol.*, vol. VT-43, pp. 439–446, Aug. 1994.

[4] J. M. Holtzman, "Adaptive measurement intervals for handoffs," in *Proc. IEEE Int. Conf. Commun.*, Chicago, IL, June 1992, vol. 2, pp. 1032–1036.

[5] N. Zhang and J. M. Holtzman, "Analysis of handoff algorithms using both absolute and relative measurements," in *Proc. IEEE Veh. Technol. Conf.*, Stockholm, Sweden, June 1994, vol. 1, pp. 82–86.

[6] R. Vijayan and J. M. Holtzman, "Analysis of handoff algorithm using nonstationary signal strength measurements," in *Proc. IEEE Global Commun. Conf.*, Orlando, FL, Dec. 1992.

[7] J. F. Whitehead, "Signal-level-based dynamic power control for cochannel interference management," in *Proc. IEEE Veh. Technol. Conf.*, Secaucas, NJ, May 1993, pp. 499–502.

[8] R. Beck and H. Panzer, "Strategies for handover and dynamic channel allocation in micro-cellular mobile radio systems," in *Proc. 39th IEEE Veh. Technol. Conf.*, San Francisco, CA, May 1989, vol. 1, pp. 178–185.

[9] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. VT-41, pp. 57–62, Feb. 1992.

[10] K. Raith and J. Uddenfeldt, "Capacity of digital cellular TDMA systems," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 323–332, May 1991.

[11] C.-L. Liu and K. Feher, "Bit error rate performance of $\pi/4$-DQPSK in a frequency-selective fast Rayleigh fading channel," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 558–568, Aug. 1991.

[12] K. Hamabe, T. Ueda, and T. Otsu, "Distributed adaptive channel allocation scheme with variable C/I threshold in cellular systems," in *Proc. IEEE Veh. Technol. Conf.*, Secaucas, NJ, May 1993, pp. 164–167.

[13] H. Andersson, H. Eriksson, A. Fallgren, and M. Madfors, "Adaptive channel allocation in a TIA IS-54 system," in *Proc. IEEE Veh. Technol. Conf.*, Denver, CO, May 1992, pp. 778–781.

[14] C. N. Chuah, R. D. Yates, and D. J. Goodman, "Integrated dynamic radio resource management," in *Proc. IEEE Veh. Technol. Conf.*, Chicago, IL, July 1995, pp. 584–588.

[15] E. A. Frech and C. L. Mesquida, "Cellular models and hand-off criteria," in *Proc. 39th IEEE Veh. Technol. Conf.*, San Francisco, CA, May 1989, pp. 128–135.

[16] T. Takenaka, T. Nakamura, and Y. Tajima, "All-channel concentric allocation in cellular systems," in *Proc. IEEE Int. Conf. Commun.*, Geneva, May 1993, vol. 2, pp. 920–924.

[17] R. W. Nettleton and G. R. Schloemer, "A high capacity assignment method for cellular mobile telephone systems," in *Proc. IEEE Veh. Technol. Conf.*, San Francisco, CA, May 1989, pp. 359–367.

[18] D. J. Goodman, S. A. Grandhi, and R. Vijayan, "Distributed dynamic channel assignment schemes," in *Proc. IEEE Veh. Technol. Conf.*, Secaucas, NJ, May 1993, pp. 532–535.

[19] A. Jalali and P. Mermelstein, "Effects of diversity, power control, and bandwidth on the capacity of microcellular CDMA systems," *IEEE J. Select. Areas Commun.*, vol. SAC-12, pp. 952–961, June 1994.

[20] S. Ariyavisitakul, "SIR-based power control in a CDMA system," in *IEEE Global Commun. Conf.*, Orlando, FL, Dec. 1992, pp. 868–873.

[21] S. Ariyavisitakul and L. F. Chang, "Signal and interference statistics of a CDMA system with feedback power control," *IEEE Trans. Commun.*, vol. COM-41, pp. 1626–1634, Nov. 1993.

[22] A. J. Viterbi, A. M. Viterbi, and E. Zehavi, "Performance of power-controlled wideband terrestrial digital communication," *IEEE Trans. Commun.*, vol. COM-41, pp. 559–569, Apr. 1993.

[23] D. G. Brennan, "Linear diversity combining techniques," *Proc. IRE*, vol. 47, pp. 1075–1102, June 1959.

[24] W. C. Y. Lee, *Mobile Communications Engineering*. New York: McGraw-Hill, 1982.

[25] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[26] F. Adachi, "Postdetection optimal diversity combiner for DPSK differential detection," *IEEE Trans. Veh. Technol.*, vol. VT-42, pp. 326–337, Aug. 1993.

[27] S. M. Hladik, S. Chennakeshu, and G. J. Saulnier, "Performance of differentially-detected $\pi/4$-shifted DQPSK with diversity," in *Proc. 42nd IEEE Veh. Technol. Conf.*, Denver, CO, May 1992, vol. 2, pp. 748–751.

[28] S. Chennakeshu and G. J. Saulnier, "Differential detection of $\pi/4$-shifted-DQPSK for digital cellular radio," *IEEE Trans. Veh. Technol.*, vol. VT-42, pp. 46–57, Feb. 1993.

[29] P. Balaban and H. P. Corrales, "Statistical performance estimation of digital radio over fading channels," in *Proc. IEEE Int. Conf. Commun.*, Denver, CO, June 1991, vol. 1, pp. 466–472.

[30] R. A. Ziegler and J. M. Cioffi, "Estimation of time-varying digital radio channels," *IEEE Trans. Veh. Technol.*, vol. VT-41, pp. 134–151, May 1992.

[31] A. P. Clark and S. Hariharan, "Efficient estimators for an HF radio link," *IEEE Trans. Commun.*, vol. COM-38, pp. 1173–1180, Aug. 1990.

[32] P. A. Wintz and E. J. Luecke, "Performance of optimum and suboptimum synchronizers," *IEEE Trans. Commun. Technol.*, vol. COM-17, pp. 380–389, June 1969.

[33] J. H. Winters, "Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading," *IEEE Trans. Veh. Technol.*, vol. VT-42, pp. 377–384, Nov. 1993.

[34] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE Trans. Veh. Technol.*, vol. VT-33, pp. 144–155, Aug. 1984.

[35] J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft-decision outputs and its applications," in *Proc. IEEE Global Commun. Conf.*, Dallas, TX, Nov. 1989, pp. 1680–1686.

[36] J. B. Scholz, "Error performance monitoring of digital communications systems," *Australian Telecommunications Research*, vol. 25, no. 2, pp. 1–25, 1991.

[37] E. A. Newcombe and S. Pasupathy, "Error rate monitoring for digital communications," *Proc. IEEE*, vol. 70, pp. 805–828, Aug. 1982.

[38] G. D. Hingorani and D. A. Chesler, "A performance monitoring technique for arbitrary noise statistics," *IEEE Trans. Commun. Technol.*, vol. COM-16, pp. 430–435, June 1968.

[39] D. J. Gooding, "Performance monitor techniques for digital receivers based on extrapolation of error rate," *IEEE Trans. Commun. Technol.*, vol. COM-16, pp. 380–387, June 1968.

[40] J. B. Scholz, S. C. Cook, and T. C. Giles, "A scheme for high performance real-time BER measurement," *IEEE Trans. Commun.*, vol. COM-40, pp. 1574–1576, Oct. 1992.

[41] N. E. Nahi and R. M. Gagliardi, "On the estimation of signal-to-noise ratio and applications to detection and tracking systems." University of Southern California, Los Angeles, CA, EE Report 114, July 1964.

[42] T. R. Benedict and T. T. Soong, "The joint estimation of signal and noise from the sum envelope," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 447–454, July 1967.

[43] D. Middleton, "The incoherent estimation of signal amplitudes in normal noise backgrounds," in *Time Series Analysis* (M. Rosenblatt, ed.), New York, Wiley, 1963, pp. 362–394. Proc. of symposium held at Brown University, Providence, RI, June 1962.

[44] R. B. Kerr, "On signal and noise level estimation in a coherent PCM channel," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-2, pp. 450–454, July 1966.

[45] R. M. Gagliardi and C. M. Thomas, "PCM data reliability monitoring through estimation of signal-to-noise ratio," *IEEE Trans. Commun. Technol.*, vol. COM-16, pp. 479–486, June 1968.

[46] C. E. Gilchriest, "Signal-to-noise monitoring," *JPL Space Programs Summary 37-27*, vol. IV, pp. 169–184, June 1966.

[47] C. M. Thomas, *Maximum Likelihood Estimation of Signal-to-Noise Ratio*. PhD thesis, University of Southern California, Los Angeles, 1967.

[48] J. W. Layland, "On S/N estimation," *JPL Space Programs Summary 37-48*, vol. III, pp. 209–212, Dec. 1967.

[49] F. Edbauer, "Signal-to-noise ratio estimation of binary PSK signals," in *National Telecommunications Conf. Record*, Los Angeles, CA, Dec. 1977, vol. 2, pp. 30:6–1–30:6–4.

[50] M. K. Simon and A. Mileant, "SNR estimation for the baseband assembly," *Telecommunications and Data Acquisition Progress Report 42-85*, Jet Propulsion Laboratory, Pasadena, CA, pp. 118–126, May 15, 1986.

[51] B. Shah and S. Hinedi, "The split symbol moments SNR estimator in narrowband channels," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-26, pp. 737–747, Sept. 1990.

[52] B. Shah and J. K. Holmes, "Improving the split-symbol moments estimator," Interoffice Memorandum 3338-90-223, Jet Propulsion Laboratory, Pasadena, CA, Dec. 19, 1990.

[53] R. Matzner, "An SNR estimation algorithm for complex baseband signals using higher order statistics," *Facta Universitatis (Nis)*, no. 6, pp. 41–52, 1993.

[54] R. Matzner and F. Engleberger, "An SNR estimation algorithm using fourth-order moments," in *Proc. IEEE Int. Symposium Inform. Theory*, Trondheim, Norway, June 1994, p. 119.

[55] R. Matzner, F. Engleberger, and R. Siewert, "Analysis and design of a blind statistical SNR estimator," in *AES 102nd Convention*, München, Germany, Mar. 1997.

[56] A. L. Brandão, L. B. Lopes, and D. C. McLernon, "In-service monitoring of multipath delay and cochannel interference for indoor mobile communication systems," in *Proc. IEEE Int. Conf. Commun.*, New Orleans, LA, May 1994, vol. 3, pp. 1458–1462.

[57] S. Yoshida, G. L. Tan, H. Zhou, A. Hirai, and T. Takeuchi, "Simple method of multipath delay difference detection for $\pi/4$-shift QPSK," *Electronics Letters*, vol. 27, pp. 1027–1028, June 1991.

[58] S. Yoshida, A. Hirai, G. L. Tan, H. Zhou, and T. Takeuchi, "In-service monitoring of multipath delay-spread and C/I for QPSK signal," in *Proc. 42nd IEEE Veh. Technol. Conf.*, Denver, CO, May 1992, vol. 2, pp. 592–595.

[59] M. Andersin, N. B. Mandayam, and R. D. Yates, "Subspace based estimation of the signal-to-interference ratio for TDMA cellular systems," in *Proc. IEEE Veh. Technol. Conf.*, Atlanta, GA, Apr. 1996, pp. 1155–1159.

[60] M. D. Austin and G. L. Stüber, "In-service signal quality estimation for TDMA cellular systems," in *PIMRC*, Toronto, Ontario, Canada, Sept. 1995, pp. 836–840.

[61] K. Feher, *Digital Communications: Satellite/Earth Station Engineering.* Englewood Cliffs, NJ: Prentice-Hall, 1983.

[62] R. E. Ziemer and W. H. Tranter, *Principles of Communications: Systems, Modulation, and Noise.* Boston: Houghton Mifflin, 3rd ed., 1990.

[63] W. H. Tranter and K. L. Kosbar, "Simulation of communication systems," *IEEE Commun. Mag.*, vol. 32, pp. 26–35, June 1994.

[64] M. Jeruchim, P. Balaban, and K. Shanmugan, *Simulation of Communication Systems.* New York: Plenum Press, 1992.

[65] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 2nd ed., 1989.

[66] A. Papoulis, *Probability, Random Variables and Stochastic Processes.* New York: McGraw-Hill, 3rd ed., 1991.

[67] H. L. Van Trees, *Detection, Estimation, and Modulation Theory.* New York: Wiley, 1968.

[68] J. C. Hancock and P. A. Wintz, *Signal Detection Theory.* New York: McGraw-Hill, 1966.

[69] German patent application no. DE 42 20 524 A1, June 1992.

[70] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. IT-40, pp. 340–349, Mar. 1994.

[71] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing.* New York: Cambridge University Press, 2nd ed., 1992.

[72] N. C. Beaulieu and C. Leung, "On the performance of three suboptimum detection schemes for binary signaling," *IEEE Trans. Commun.*, vol. COM-33, pp. 241–245, Mar. 1985.

[73] R. Matzner, F. Engleberger, and R. Siewert, "Analysis and design of a blind statistical SNR estimator," in *AES 102nd Convention*, München, Germany, Mar. 1997.

[74] M. V. Clark, L. J. Greenstein, W. K. Kennedy, and M. Shafi, "MMSE diversity combining for wide-band digital cellular radio," *IEEE Trans. Commun.*, vol. COM-40, pp. 1128–1135, June 1992.

[75] S. Stein, "Fading channel issues in system engineering," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 68–89, Feb. 1987.

[76] S. C. Gupta, R. Viswanathan, and R. Muammar, "Land mobile radio systems—a tutorial exposition," *IEEE Commun. Mag.*, vol. 23, pp. 34–45, June 1985.

[77] T. Kürner, D. J. Cichon, and T. C. Becker, "Degradation of digital communication systems in a multipath environment," in *Proc. 44th IEEE Veh. Technol. Conf.*, Stockholm, Sweden, June 1994, vol. 1, pp. 170–174.

[78] GSM Recommendation 05.05—DCS Version 3.1.0: European digital cellular telecommunications system (phase 1): Radio Transmission and Reception, ETSI/PT12, Feb. 1992.

[79] TIA/EIA IS-54, "Cellular system dual-mode mobile station-base station compatibility standard," Telecommunications Industry Association, Apr. 1992.

[80] M. R. L. Hodges, "The GSM radio interface," *Br. Telecom Tech. J.*, vol. 8, pp. 31–43, Jan. 1990.

[81] M. Mouly and M.-B. Pautet, *The GSM System for Mobile Communications.* Paris, 1992. ISBN 2-9507190-0-7.

[82] ETSI GSM Specifications, Series 01-12.

[83] K. Kinoshita, M. Kuramoto, and N. Nakajima, "Development of a TDMA digital cellular system based on Japanese standard," in *Proc. 41st IEEE Veh. Technol. Conf.*, 1991, pp. 642–645.

[84] Research and Development Center for Radio Systems (RCR), "Digital cellular telecommunication systems," RCR STD-27, Apr. 1991.

[85] D. J. Goodman, "Second generation wireless information networks," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 366–374, May 1991.

[86] J. E. Padgett, C. G. Günther, and T. Hattori, "Overview of wireless personal communications," *IEEE Commun. Mag.*, vol. 33, pp. 28–41, Jan. 1995.

[87] K. Feher, "MODEMS for emerging digital cellular-mobile radio system," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 355–365, May 1991.

[88] D. J. Goodman, "Trends in cellular and cordless communications," *IEEE Commun. Mag.*, vol. 29, pp. 31–40, June 1991.

[89] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst.*, vol. CS-11, pp. 360–393, Dec. 1963.

[90] A. Müller, "Simulation of multipath fading channels using the Monte-Carlo method," in *Proc. IEEE Int. Conf. Commun.*, New Orleans, LA, May 1994, vol. 3, pp. 1536–1540.

[91] D. C. Cox, "910 MHz urban mobile radio propagation: Multipath characteristics in New York City," *IEEE Trans. Commun.*, vol. COM-21, pp. 1188–1194, Nov. 1973.

[92] P. Hoeher, "A statistical discrete-time model for the WSSUS multipath channel," *IEEE Trans. Veh. Technol.*, vol. VT-41, pp. 461–468, Nov. 1992.

[93] H. Suzuki, "A statistical model for urban radio propagation," *IEEE Trans. Commun.*, vol. COM-25, pp. 673–680, July 1977.

[94] W. R. Braun and U. Dersch, "A physical mobile radio channel model," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 472–482, May 1991.

[95] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, vol. 47, pp. 957–1000, July–Aug. 1968.

[96] C.-L. Liu and K. Feher, "Noncoherent detection of $\pi/4$-QPSK systems in a CCI-AWGN combined interference environment," in *Proc. 39th IEEE Veh. Technol. Conf.*, San Francisco, CA, May 1989, vol. 1, pp. 83–94.

[97] C. S. Ng, T. T. Tjhung, F. Adachi, and K. M. Lye, "On the error rates of differentially detected narrowband $\pi/4$-DQPSK in Rayleigh fading and Gaussian noise," *IEEE Trans. Veh. Technol.*, vol. VT-42, pp. 259–265, Aug. 1993.

[98] S. H. Goode, H. L. Kazecki, and D. W. Dennis, "A comparison of limiter-discriminator, delay, and coherent detection for $\pi/4$ QPSK," in *Proc. 40th IEEE Veh. Technol. Conf.*, Orlando, FL, May 1990, pp. 687–694.

[99] V. Fung and T. S. Rappaport, "Bit-error simulation of $\pi/4$ DQPSK in flat and frequency-selective fading mobile radio channels with real time applications," in *Proc. IEEE Int. Conf. Commun.*, Denver, CO, 1991, vol. 2, pp. 553–557.

[100] V. Fung and T. S. Rappaport, "Bit error simulation of $\pi/4$ DQPSK mobile radio communications using two-ray and measurement-based impulse response models," *IEEE J. Select. Areas Commun.*, vol. SAC-11, pp. 393–405, Apr. 1993.

[101] U. Dersch and R. J. Rüegg, "Simulations of the time and frequency selective outdoor mobile radio channel," *IEEE Trans. Veh. Technol.*, vol. VT-42, pp. 338–344, Aug. 1993.

[102] B. Glance and L. J. Greenstein, "Frequency selective fading effects in digital mobile radio with diversity combining," *IEEE Trans. Commun.*, vol. COM-31, pp. 1085–1094, Sept. 1983.

[103] J. C.-I. Chuang, "The effects of time delay spread on portable radio communications channels with digital modulation," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 879–889, June 1987.

[104] T. S. Rappaport, S. Y. Seidel, and R. Singh, "900-MHz multipath propagation measurements for U.S. digital cellular radiotelephone," *IEEE Trans. Veh. Technol.*, vol. VT-39, pp. 132–139, May 1990.

[105] J. B. Andersen, T. S. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *IEEE Commun. Mag.*, vol. 33, pp. 42–49, Jan. 1995.

[106] B. D. Woerner, J. H. Reed, and T. S. Rappaport, "Simulation issues for future wireless modems," *IEEE Commun. Mag.*, vol. 32, pp. 42–53, July 1994.

[107] COST-207: "Digital land mobile radio communications," Final Report. (1989), EUR 12160 EN, Commission European Community, Luxembourg.

[108] I. Crohn and E. Bonek, "Modeling of intersymbol-interference in a Rayleigh fast fading channel with typical delay power profiles," *IEEE Trans. Veh. Technol.*, vol. VT-41, pp. 438–447, Nov. 1992.

[109] J. I. Smith, "A computer generated multipath fading simulation for mobile radio," *IEEE Trans. Veh. Technol.*, vol. VT-24, pp. 39–40, Aug. 1975.

[110] M. T. Le and A. U. Sheikh, "Performance of $\pi/4$-QDPSK in a Rayleigh/lognormal/delay spread/AWGN/co-channel interference environment," in *Proc. 42nd IEEE Veh. Technol. Conf.*, Denver, CO, May 1992, vol. 1, pp. 147–150.

[111] S. Gurunathan and K. Feher, "Multipath simulation models for mobile radio channels," in *Proc. 42nd IEEE Veh. Technol. Conf.*, Denver, CO, May 1992, vol. 1, pp. 131–134.

[112] S. A. Fechtel, "A novel approach to modeling and efficient simulation of frequency-selective fading radio channels," *IEEE J. Select. Areas Commun.*, vol. SAC-11, pp. 422–431, Apr. 1993.

[113] F. Goulam and A. J. Levy, "A statistical model for the simulation of multipath mobile propagation channel," in *Proc. 42nd IEEE Veh. Technol. Conf.*, Denver, CO, May 1992, vol. 1, pp. 135–138.

[114] T. S. Rappaport and V. Fung, "Simulation of bit error performance of FSK, BPSK, and $\pi/4$ DQPSK in flat fading indoor radio channels using a measurement-based channel model," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 731–740, Nov. 1991.

[115] H. Hashemi, "Simulation of the urban radio propagation channel," *IEEE Trans. Veh. Technol.*, vol. VT-28, pp. 213–225, Aug. 1979.

[116] G. D'Aria, F. Muratore, and V. Palestini, "Simulation and performance of the pan-European land mobile radio system," *IEEE Trans. Veh. Technol.*, vol. VT-41, pp. 177–189, May 1992.

[117] K.-W. Yip and T.-S. Ng, "Efficient simulation of digital transmission over WSSUS channels," *IEEE Trans. Commun.*, vol. COM-43, pp. 2907–2913, Dec. 1995.

[118] R. F. W. Coates, G. J. Janacek, and K. V. Lever, "Monte Carlo simulation and random number generation," *IEEE J. Select. Areas Commun.*, vol. SAC-6, pp. 58–66, Jan. 1988.

[119] S. K. Park and K. W. Miller, "Random number generators: Good ones are hard to find," *Communications of the ACM*, vol. 31, pp. 1192–1199, Oct. 1988.

[120] M. Fattouche and H. Zaghloul, "Equalization of $\pi/4$ offset DQPSK transmitted over flat fading channels," in *Proc. IEEE Int. Conf. Commun.*, Chicago, IL, 1992, pp. 296–298.

[121] G. A. Arredondo, W. H. Chriss, and E. H. Walker, "A multipath fading simulator for mobile radio," *IEEE Trans. Commun.*, vol. COM-21, pp. 1325–1328, Nov. 1973.

[122] P. Huber, *Robust Statistics.* New York: Wiley, 1981.

[123] N. Amitay and L. J. Greenstein, "Multipath outage performance of digital radio receivers using finite-tap adaptive equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 597–608, May 1984.

[124] K. Ohno and F. Adachi, "QDPSK signal transmission performance with post-detection selection diversity reception in land mobile radio," *IEEE Trans. Veh. Technol.*, vol. VT-40, pp. 798–804, Nov. 1991.

[125] F. Adachi, K. Ohno, and M. Ikura, "Postdetection selection diversity reception with correlated, unequal average power Rayleigh fading signals for $\pi/4$-shift QDPSK mobile radio," *IEEE Trans. Veh. Technol.*, vol. VT-41, pp. 199–210, May 1992.

[126] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1975.

[127] I. Korn, *Digital Communications.* New York: Van Nostrand Reinhold, 1985.

[128] C. R. Rao, *Linear Statistical Inference and its Applications.* New York: Wiley, 1965.

[129] I. W. Burr, *Applied Statistical Methods.* New York: Academic Press, 1974.

[130] H. G. Tucker, *An Introduction to Probability and Mathematical Statistics.* New York: Academic Press, 1962.

[131] M. Nakagami, "The *m*-distribution—a general formula of intensity distribution of rapid fading," in *Statistical Methods of Radio Wave Propagation* (W. C. Hoffman, ed.), Elmsford, NY: Pergamon, 1960.

[132] D. D. Falconer, F. Adachi, and B. Gudmundson, "Time division multiple access methods for wireless personal communications," *IEEE Commun. Mag.*, vol. 33, pp. 50–57, Jan. 1995.

[133] F. G. Jenks, P. D. Morgan, and C. S. Warren, "Use of four-level phase modulation for digital mobile radio," *IEEE Trans. Electromagn. Compat.*, vol. EMC-14, pp. 113–128, Nov. 1972.

[134] Y. Akaiwa and Y. Nagata, "Highly efficient digital mobile communications with a linear modulation method," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 890–895, June 1987.