TOWARDS INCREASING THE CAPACITY IN MIMO WIRELESS COMMUNICATION SYSTEMS

by

Seyed Alireza Banani

M.Sc., Shiraz University, 2007
B.Sc., Shiraz University, 2004

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the

School of Engineering Science

Faculty of Applied Sciences

SIMON FRASER UNIVERSITY

Fall 2011

APPROVAL

Name: **Seyed Alireza Banani**

Degree: **Doctor of Philosophy**

Title of Thesis: **Towards Increasing the Capacity in MIMO Wireless Communication Systems**

Examining Committee:

Chair: **Dr. Shawn Stapleton, P.Eng**
Professor – School of Engineering Science

_____

Dr. Rodney Vaughan
Senior Supervisor
Professor – School of Engineering Science

_____

Dr. James K. Cavers, P.Eng
Supervisor
Professor – School of Engineering Science

_____

Dr. Paul Ho, P.Eng
Supervisor
Professor – School of Engineering Science

_____

Dr. Sami (Hakam) Muhaidat
Internal Examiner
Assistant Professor – School of Engineering Science

_____

Dr. Lutz Lampe
External Examiner
Professor,
Department of Electrical and Computer Engineering,
University of British Columbia, BC, Canada

Date Defended/Approved: November 04, 2011

Partial Copyright Licence

ABSTRACT

The use of multiple antennas on both sides of a multipath channel can improve the capacity without additional transmit power. It is possible, in principle, for the capacity to increase linearly with the number of antennas even when the channel state information (CSI) is unknown at the transmitter. These known capacity results require many assumptions, including the need for the CSI to be known perfectly. In practice, perfect CSI is never known perfectly a priori, and its estimation, without using an ideal blind technique (none are available), requires bandwidth resource which reduces the capacity. Moreover, various factors such as digital modulation, finite block lengths, and imperfect power allocation degrade the capacity from the Shannon limit to the practicable possibilities of a digital link.

These practical impairments motivate new techniques for increasing the capacity in MIMO systems, and in this thesis, two sets of techniques are presented. The first set includes two new decision-directed techniques for estimating the channel matrix, and they are shown to have higher capacity compared to pilot symbol assisted modulation systems since no pilots are required. These techniques are applicable for various open-loop SISO/MIMO wireless communications systems including systems employing OFDM, nonlinear/linear equalization, MRC, Alamouti coding, and spatial multiplexing. In the second set, the eigen-MIMO capacity is maximized in the presence of different practical impairments. In particular, the joint influence of training-based

channel estimation and imperfect feedback on both the information-theoretic and the practicable water-filled eigen-MIMO capacities, are analyzed. Water-filled eigenchannels maximize the information theoretic capacity, but for implementation, the required adaptive modulation means high complexity. One simplification is to have fixed modulation over a fixed number of eigenchannels. However, the error rate deteriorates with the weakest eigenchannel and to counter this while maintaining high throughput, the information rate is maximized with an output SNR constraint. On the other hand, if higher complexity can be tolerated, adaptive modulation and coding can be deployed for high throughput. In this context, a high capacity eigen-MIMO system using Reed-Solomon coded *M*-QAM is developed. This includes an appropriate QAM, code rate, and power allocation for each eigenchannel.

Keywords: capacity; capacity efficiency; channel estimation; closed-loop wireless system; eigen-channel; feedback; multiple input multiple output; throughput

ACKNOWLEDGEMENTS

At first, I would like to thank my senior supervisor Dr. Rodney Vaughan for his great mentoring, and encouragement during the previous four years of my PhD studies. At many stages I benefited from his advice, particularly so when exploring new ideas. He gave me the freedom of exploring different aspects of research while giving me financial support when I needed it. His positive outlook and confidence in my research inspired me and gave me confidence. His careful editing contributed enormously to the production of this thesis. Special thanks go to my supervisory committee Dr. Jim Cavers and Dr. Paul Ho for their great ideas and commentary. Also I have to thank them for teaching me the fundamentals of wireless communications in their classes. Meanwhile, I want to express my gratitude to my examiners consisting of Dr. Sami Muhaidat and Dr. Lutz Lampe for reading this work and giving their comments.

I am grateful for the help and support of the faculty, staff and the graduate studies of the department of Engineering Science at Simon Fraser University. They willingly lent me their time, resources and expertise towards my research. I also would like to thank Natural Science and Engineering Research Council (NSERC) of Canada for providing the financial support of this work.

Many thanks to all my colleagues and friends for their warmth, understanding, back-ups, discussion, and being there for me, in particular Dr.

Jane Xyun, Dr. Sara Bavarian and Dr. Jinyun Ren. Special thanks also goes to Maral, Mehdi, Milad, Piraj, Moein, Mahsa and Parastoo for helping me keep things in perspective and providing a much needed form of escape from my studies.

I must express my warmest gratitude to my parents Dr. Abbas Banani and Dr. Mahnaz Faraghat for their continual support and encouragement. I was amazed by the patience of my mother and father who experienced all of the ups and downs of my research. Last but not the least, I do not have the words to thank my brother Javad and his wife Sholeh for their kind heart and believing in me in every step of the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ACM	adaptive coded modulation
AM	amplitude modulation
AR	auto-regressive
AWGN	additive white Gaussian noise
BER	bit-error rate
BLER	block error rate
BPSK	binary phase-shift keying
CMOS	complementary metal–oxide–semiconductor
CODEC	encoders/decoder
CP	cyclic prefix
CR	channel response
CSI	channel state information
DFT	discrete Fourier transform
DVB	digital video broadcasting
FEC	forward error correction
FFT	fast Fourier transform
FPGA	field-programmable gate array
GPS	global positioning system
HE	horizontal encoding

HPA	high power amplifier
ICA	independent component analysis
IEEE	institute of electrical and electronis engineers
IFFT	inverse fast Fourier transform
ISI	inter-symbol intereference
ISM	industrial, scientific and medical
ITU-R	international telecommunication union radio-communication
LAN	local area network
LDPC	low-density parity-check
LMMSE	linear minimum mean square error
LMS	least mean squares
LNA	low noise amplifier
LPD	linear precoder and decoder
LS	least squares
LUT	look-up table
MAP	maximum a posteriori
MIMO	multiple-input multiple-output
ML	maximum likelihood
MMSE	minimum mean square error
MRC	maximum ratio combining
MS	mobile station
NMSE	normalized mean square error

OFDM	orthogonal frequency-division multiplexing
OSUC	ordered successive cancellation
PDF	probability density function
PM	phase modulation
PSAM	pilot symbol assisted modulation
PSK	phase-shift keying
QAM	quadrature amplitude modulation
QPSK	quadrature phase-shift keying
RLS	recursive least squares
RMSE	root mean square error
RS	Reed-Solomon
RSR	residual systematic resampling
SER	symbol-error-rate
SIR	sampling importance resampling
SISO	single-input single-output
SM	spatial multiplexing
SNR	signal-to-noise ratio
SR	systematic resampling
SSPA	solid state power amplifier
SUC	successive cancellation
SVD	singular value decomposition
TDL	tapped delay line

UT	unscented transformation
VQ	vector quantizer
WiMAX	worldwide interoperability for microwave access
WLAN	Wireless local area network
WSS	wide-sense stationary
WSSUS	wide-sense stationary uncorrelated scattering
ZF	zero forcing

1: INTRODUCTION

Communication in many wireless channels is impaired predominantly by multipath fading. The random fluctuation in received signal level, known as fading, can severely affect the quality and reliability of wireless communications. In addition, the constraints posed by limited power and scarce frequency bandwidth make the task of designing high data rate, high reliability wireless communication extremely challenging. The use of multiple antennas at the transmitter and receiver in wireless systems, known as multiple-input multiple-output (MIMO) technology, has rapidly gained popularity over the past decade due to its powerful performance-enhancing capabilities. MIMO technology constitutes a breakthrough in system design. The technology offers benefits that help meet the challenges imposed by both the impairments in the wireless channel as well as the resource constraint. In addition to the time and frequency dimensions that are exploited in conventional single antenna wireless systems, MIMO uses the spatial dimension provided by the multiple antennas at the transmitter and receiver. In a line-of-sight situation, using the spatial dimension is the equivalent of adding antenna aperture, i.e., using antennas with high directional gain.

The benefits of MIMO systems can be expressed as array gain, diversity gain, spatial multiplexing gain, and interference reduction. *Array gain* refers to the average increase in the receive signal-to-noise ratio (SNR) that results from a

coherent combination of the wireless signals at the receiver or transmitter or both. *Diversity gain* mitigates fading and is realized by receiving multiple (ideally independent) copies of the transmitted signal in space (i.e., signals from different antenna patterns, having amplitude, phase and polarization differences), frequency, or time. *Spatial multiplexing* offers an increase in the capacity with no additional power expenditure. It is possible, in principle, for the capacity to increase linearly with the number of antennas even when the channel state information (CSI) is unknown at the transmitter. Interference in wireless networks results from multiple users sharing time and frequency resources. By exploiting the spatial dimension to increase the separation between users, *interference mitigation* also becomes possible.

The communications performance of time-varying wireless systems depends on the accuracy of the estimation of parameters such as the fading channel coefficients. The channel coefficients must be estimated by either pilot symbols (training sequences) or in a decision-directed manner. A note on the difference between “blind” and “decision-directed” is in order. In a strict interpretation, blind techniques imply that no pilots are used at all, *even for initialization of channel estimation*. Decision-directed means that pilots are needed, *even if just for initialization*. After initialization, decision-directed algorithms behave in a blind manner, in the sense that after the initialization, no pilots are required. Several papers in the literature refer to this case as blind, (see the references in Chapters 2 and 3). In this thesis, this case is referred to as

decision-directed, although the term “blind” is used in occasional passages where it may be helpful for clarity in relation to citations.

Even with perfect channel knowledge, it may not be possible to obtain simultaneously all the benefits described above due to their competing demands on the spatial degrees of freedom (number of antennas). However, using some combination of the benefits will result in improved capacity and/or reliability in a wireless network. Transceiver algorithms for MIMO systems may be broadly classified into two categories: those focussed on increasing the capacity; and those focused on increasing the reliability. Intuitively, for a fixed capacity, an increase in SNR will reduce the symbol-error-rate (SER) of the system. Similarly, at a fixed target SER, an increase in SNR may be leveraged to increase the transmission rate. In this way, a fundamental trade-off exists between the capacity and reliability. In the context of MIMO systems, this trade-off is often referred to as the diversity-multiplexing trade-off, with diversity signifying the reliability improvement and multiplexing signifying an increase in capacity.

In this thesis the increase in the capacity of practical MIMO wireless communication systems is a focus. As noted above, the capacity can increase with the number of antennas even when the CSI is unknown at the transmitter. A further increase in capacity is possible when CSI is known at the transmitter. But these capacity results require many assumptions, including the need for the CSI to be known perfectly. In practice, perfect CSI is never known perfectly a priori and its estimation requires bandwidth resource which in turn reduces the capacity. In general, fine differences in theoretical performance from different

architectures may be overwhelmed by the engineering compromises inherent in a practical design for a given channel type. Such compromises include the deployment of practical communications techniques including choices of modulation such as QAM signaling (instead of Gaussian signals), coding, pulse shaping, filter sizes, and finite block lengths (instead of infinitely long codes), and the associated guardbands which contribute to the required bandwidth, etc. These practical impairments motivate research for finding new capacity-efficient techniques for MIMO wireless systems.

The research reported in this thesis can be classified into two sets. The first set comprises two new techniques for estimating the channel matrix using decision-directed methods, aided by a necessary pilot-based initialization. Since no pilot or training sequence is used (after the initialization), these decision-directed systems turn out to be more capacity efficient compared to pilot symbol assisted modulation (PSAM) systems and so-called semi-blind techniques. This encourages the invention and development of further decision-directed channel estimation techniques for various SISO/MIMO wireless communications systems.

In the first technique, a decision algorithm obtains primary data estimates of the transmitted symbols based on the constrained linear minimum mean square error (LMMSE) criterion, and then these are applied to subsequent optimal minimum mean square error channel estimation. The second technique is based on non-stationary independent component analysis (ICA) for separating each source signal, and it uses particle filtering to track the time-varying channel. The advantages of these decision-directed techniques are:

1) With no pilot or training sequence used rather than initialization, the decision-directed systems turn out to be more spectrally efficient compared to PSAM systems;

2) They exploit the channel taps variances and only few values from the time correlation coefficient function, thus full knowledge of the second-order statistics is not required;

3) The algorithms do not require large sample sizes to converge as required in many higher-order statistics based approaches;

4) Unlike many previous decision-directed approaches that are based on block processing, the presented techniques are realized in symbol time in the sense that channel coefficients are estimated at the symbol rate; and

5) They have the potential to be formulated for various SISO/MIMO wireless communications systems including those employing orthogonal frequency-division multiplexing (OFDM) [1.1], nonlinear/linear equalizers [1.2]-[1.4], maximum-ratio combining (MRC) [1.5], Alamouti coding [1.6], spatial multiplexing [1.7], etc.

Further MIMO systems that may benefit from the proposed decision-directed techniques are those with nonlinearity in both transmitter and receiver amplifiers [1.8]-[1.9] and the ones with maneuvering terminals where the relative speed of the transmit/receive terminals may change in time [1.10]-[1.11].

In the second set of techniques, the aim is to maximize the MIMO information-theoretic capacity and the *practicable capacity.* (The practical

capacity is also referred to as *throughput* (and many other terms, see Chapter 6) in which this term is specifically interpreted as *the throughput of correctly detected bits that would arise if an idealized coder was used*, in the presence of different practical impairments. Throughput is used mainly in Chapter 6, and practicable capacity is mainly used in Chapers 5 and 7, following the references' terminology.) In particular, an analysis is developed for the joint influence of training-based channel estimation and imperfect feedback (mean feedback) on both the information-theoretic and the practicable water-filled eigen-MIMO capacity [1.12]-[1.13]. The channel is estimated via training, then a vector quantizer (VQ) codes the CSI which is fed back to the transmitter via a CSI feedback link with a throughput constraint. Both of these capacities are formulated as a function of various parameters. Then the optimal number of training symbols and the optimal power allocation for training and data transfer is obtained with the criterion of the associated capacity. It turns out that by jointly optimizing over the number of training symbols and the feedback delay time (by varying the VQ code size), a further increase in both of the capacities is possible.

Even with perfect CSI, there are other practical challenges with each type of optimized linear precoder-decoder design. For offering insight into MIMO configurations that are practical but can still maintain high performance, new formulations are required. For the (capacity-maximizing) case of water filling, the different eigenchannel gains (and the varying numbers of eigenchannels, in general) contribute to the difficulty of allocating different digital modulations to different eigenchannels. This adaptive modulation has high complexity and is not

easy to implement. Sophisticated hardware is needed (and usually not available in legacy systems) at the terminals to support the variable modulation, and a sophisticated protocol is required to support the adaptation via data exchange and handshakes. Also, adaptive modulation is not effective in fast-changing channels. This is because of the prolonged protocol required: the transmitter needs to know the SNR on each eigenchannel in order to adapt the modulations and inform the receiver of the adaptation. Even if the protocol, including the ongoing channel sounding, can keep up with the changing channel, the resulting capacity overhead may still be significant. One simplification is to have fixed common modulation with a common signaling rate, for a fixed number of eigenchannels, also referred to as non-adaptive modulation. With this fixed modulation scheme, the reduced complexity extends to the protocol and to the hardware of all the terminals in a multi-user system.

A feature of water-filling with fixed (or adaptive) modulation, is that the uncoded SNR/SER performance of the overall system deteriorates with the weakest eigenchannel. For such a system, the uncoded SNR/SER performance trades off with the throughput which depends on the receiver SNR and the power allocation on the eigenchannels. Since there is no exact analytical formulation for throughput, the trade-off is quantified based on maximizing the theoretic capacity under the constraint of a maximum allowable SNR loss, relative to the receiver SNR, on the weakest eigenchannel. The SNR constraint guarantees a desired error performance. [1.14]. It is shown via simulation that the throughput (recall this is the same as practicable capacity) using the new power allocation

algorithm shows similar behavior to the information-theoretic capacity. In this sense, the information theoretic capacity for uncoded and fixed modulation is being maximized in order to establish limits of practical, reduced complexity systems.

On the other hand, if high complexity can be tolerated, then adaptive modulation together with forward-error correction coding will produce a better throughput as long as the channel is slowly fading. In this context, a capacity efficient technique and its analysis is presented where adaptive QAM and Reed-Solomon (RS) coding are combined for capacity realization in eigen-MIMO [1.15]-[1.16]. RS coding has the advantages of algorithmic simplicity, low memory requirements, and decoder complexity. Its closed-form error probability is useful for obtaining an optimal power allocation, signal constellation size(s), and code rate(s) on the eigenchannels, for maximum throughput. The adaptation is applied to two different architectures of the encoders/decoders (CODECs) for eigen-MIMO, and their performances are compared with the uncoded case. The *outer coding* architecture refers to a single CODEC working on the overall serial data, and *inner coding* refers to separate CODECs for different eigenchannels. The adaptive system with optimum power allocation reveals new capacity behavior which is different to that of water-filling; and similar behaviour holds for practicable capacity maximization with uncoded non-adaptive modulation [1.17]. Finally, a selection procedure is presented between different system configurations, including the adaptive RS coded modulation applied to dominant

eigenmode transmission, in order to obtain the highest practicable capacity subject to either average or instantaneous output BER.

A note on complexity is in order since the thesis is in a theoretical approach to increasing capacity rather than a treatment of complexity. It is conceded that theory is not readily separated from complexity since feasible systems are the focus of engineering research, but in order to make progress here, the focus is taken off complexity considerations. In some chapters, a reduced complexity system is proposed, and in others, the increased complexity (generally required for improved capacity) is set aside, although calculations are offered for estimating aspects of the algorithmic complexity.

The rest of the thesis is organized as follows: Chapter 2 describes the first proposed decision-directed channel estimation technique with the application to uncoded OFDM system. In Chapter 3, the first new decision-directed technique is applied to MIMO systems with nonlinear amplifiers. Here, signal processing is used also to compensate for the non-linearity. The ICA-based decision-directed channel estimation is presented in Chapter 4. Its advantages are demonstrated for MIMO systems with maneuvering terminals where the algorithm makes it possible to track the relative speed of the terminals. In Chapter 5, an analysis is presented for the influence of training-based channel estimation and imperfect feedback on both the information-theoretic and the practicable capacities. The problem of information-rate maximization with output SNR constraint with non-adaptive modulation is treated in Chapter 6. Chapter 7 addresses adaptive RS

coded modulation in eigen-MIMO. The thesis concludes with proposed future directions for this research in Chapter 8.

The notation of this thesis is conventional, as follows. Symbols for matrices (in capital letters) and vectors are in boldface. The notations $(.)^H$, $(.)^T$ and $(.)^*$ stand for conjugate transpose, transpose, and complex conjugate, respectively. $\mathbf{I}$ is the identity matrix, $\|.\|_F$ denotes the Frobenius norm, $\mathrm{tr}(.)$ is the trace of a matrix and $E(.)$ denotes expectation. A matrix $\mathbf{X}$ containing entries of i.i.d. zero-mean complex Gaussian entries with variance 1, is written $\mathbf{X} \sim CN(0,1)$.

REFERENCE LIST

[1.1] S. A. Banani, and R. G. Vaughan, “OFDM with Iterative Blind Channel Estimation,” *IEEE Trans. Veh. Technol,* vol. 59, no. 9, pp. 4298-4308, Nov. 2010.

[1.2] S. A. Banani, and R. G. Vaughan, “Joint Blind Channel Estimation and Equalization in Dispersive Fading Channel,” *IET Communications,* Vol. 5, No. 11, pp. 1577 – 1586, Jul. 2011.

[1.3] S. A. Banani, and R. G. Vaughan, “Iterative blind linear equalization in time-varying dispersive channels,” in *Proc. IEEE Canadian Conf. and Elect. Comput. Eng.,* pp. 1-6, May 2010.

[1.4] S. A. Banani, and R. G. Vaughan, “Blind correlated TDL equalization in fast Rician fading channels,” in *Proc. IEEE Canadian Conf. Elect. and Comput. Eng.,* pp. 1-5 , May 2011.

[1.5] S. A. Banani, and R. G. Vaughan, “Blind channel estimation for systems with maximum-ratio receiver combining,” in *Proc. IEEE Veh. Technol. Conf.*, pp. 1-5, May 2010.

[1.6] S. A. Banani, and R. G. Vaughan, “Blind channel estimations for systems employing Alamouti coding,” in *Proc. IEEE Wireless Commun. and Network. Conf.* pp. 1-6, Apr. 2010.

[1.7] S. A. Banani, and R. G. Vaughan, “ICA with Particle filtering for blind channel estimation in high data-rate MIMO systems,” in *Proc. IEEE Wireless Commun. and Network. Conf.,* pp. 1-6, Apr. 2010.

[1.8] S. A. Banani, and R. G. Vaughan, "Compensating for non-linear amplifiers in MIMO communications systems," accepted in *IEEE Trans. Antennas and Propagation*, Manuscript ID: AP1006-0617R2, Jul. 2011.

[1.9] S. A. Banani, and R. G. Vaughan, “Blind channel estimation for MIMO systems with nonlinearities at the receiver,” in *Proc. IEEE Veh. Technol. Conf.* pp. 1-5, May 2010.

[1.10] S. A. Banani, and R. G. Vaughan, “Blind Channel Estimation and Speed Tracking in Wireless Systems using ICA with Particle Filtering,” accepted in *IET Communications*, Manuscript ID: COM-2010-0836R1, Oct. 2011.

[1.11] S. A. Banani, and R. G. Vaughan, “Blind channel estimation for MRC systems with maneuvering transmit/receive terminals,” in *Proc. IEEE Canadian Conf. Elect. and Comput. Eng.,* pp. 1-6, May 2010.

[1.12] S. A. Banani, and R. G. Vaughan, “Practicable capacity in eigen-MIMO with channel estimation and mean feedback,” submitted in *IET Communications,* Jan. 2011, Manuscript ID: COM-2011-0045.

[1.13] S. A. Banani, and R. G. Vaughan, “The effect of training-based channel estimation on the capacity of closed-loop MIMO systems with imperfect CSI feedback,” in *Proc. IEEE Veh. Technol. Conf.,* pp. 1-5, 2010.

[1.14] S. A. Banani, and R. G. Vaughan, “Throughput and SER trade-off in eigen-MIMO with fixed modulation,” submitted to, *IEEE Trans. Vehicular Technology,* Manuscript ID: VT-2011-01658, Nov. 2011.

[1.15] S. A. Banani, and R. G. Vaughan, “Capacity Maximization in Eigen-MIMO from Adaptive Modulation and Reed-Solomon Coding,” submitted in *IET Communications.,* Sep. 2011, Manuscript ID: COM-2011-0730.

[1.16] S. A. Banani, and R. G. Vaughan, “Adaptive Reed-Solomon coding in eigen-MIMO with non-adaptive modulation,” *IEEE Veh. Technol. Conf.*, San Francisco, USA, Sep. 2011.

[1.17] S. A. Banani, and R. G. Vaughan, “Power allocation for practicable capacity maximization in eigen-MIMO,” in *Proc. IEEE Veh. Technol. Conf.*, pp. 1-5, May 2011.

2: OFDM WITH DECISION-DIRECTED ITERATIVE CHANNEL ESTIMATION

OFDM has broad uptake in wireless communications. Although standards are in place, the structure of OFDM transmission is still evolving as new technology and applications develop. OFDM in its current form is best for dispersive channels which do not vary quickly. Because of OFDM's ubiquitous uptake, including its increasing extensions to time-varying mobile channels with widely different rates of change, and higher radio frequencies, its structure and capacity efficiency should continue to be researched.

OFDM signals have been optimized over various components such as partitioning [2.1], adaptive loading [2.2] and coding [2.1]. All optimization depends on the channel, thus the modeling and the estimation of the channel is critical. Furthermore, for OFDM systems with multiple antennas, accurate channel estimation is required for the combining, interference suppression, and the ensuing signal detection [2.1].

With accurate channel estimation and tracking, OFDM systems can use coherent detection (taken here to mean a receiver which has CSI, and optimal coherent detection is taken to mean having perfect CSI) for a ~3dB SNR gain over differential detection. The system with differential detection has simpler implementation and uses differential modulation (also called differential encoding) at the transmitter and conventional differential detection at the receiver

in order to avoid the need for channel estimation. Its development is discussed in documents such as in the Eureka 147 project from the 1980s in which DPSK modulation is employed in each subcarrier. The simplicity of the receiver design of a differential system comes at the price of a performance degradation relative to that of coherent combining. The degradation can be attributed to several different mechanisms. Firstly, if the reference symbol is a noisy estimate of the actual symbol, a conventional differential detector suffers from error propagation (sometimes referred to as noise boosting) in the decision metric, compared to coherent detection. Secondly, in differential detection, two symbol times are needed for detection, so the noise power is essentially doubled relative to coherent detection. Thirdly, for time-varying channels, the channel variation between adjacent symbols will introduce significant performance degradation for differential detection. For the case of OFDM, the subchannel bandwidth is small, resulting in a long OFDM symbol interval, which may allow significant channel variation between adjacent OFDM symbols. In addition, inter-carrier interference (ICI) may be introduced by OFDM due to the loss of orthogonality caused by the Doppler spread of the time-varying channels.

In practice, the channel can be estimated either by the use of pilot symbols or in a decision-directed manner. Adding training sequence pilots or carrier pilots in the time or frequency domain enables estimation of the channel response (CR) at the pilot positions. The CRs at the other positions can be estimated by interpolation. Pilot schemes for channel sounding not only require a reduction in the available payload capacity, but they may also add complexity –

both the receiver and transmitter must have in-built *a priori* knowledge of the pilot scheme, and both must deploy the extra processing to implement the pilot scheme. (For compatibility, the Standards must also enumerate the complex details of the timing and its tolerances of the pilot scheme.) Nevertheless, this is the status quo, because the link performance cannot be realized currently by other means. No formal studies appear to have been undertaken on the complexity of pilot systems versus non-pilot systems; the point here is that some complexity must be associated with the inclusion of a pilot sceme in a system for channel sounding.

Many pilot-based methods for estimating the channel have been proposed; e.g., [2.3]-[2.10], and more recently others too numerous to list. Particular attention has been paid to optimizing and analyzing the effects of the location of the pilots within the frequency band for a given pilot density, the number of pilot symbols per packet, and the power dedicated to the pilots relative to the data, e.g., [2.6]-[2.8]. As the channel varies more rapidly (in time or frequency), denser pilots are required. But increasing the number of pilots also acts to reduce the available payload capacity, so the pilot-based scheme can only be optimized for a given rate of change of the channel. There have been attempts to get round this while maintaining constant pilot density. In [2.9] a pseudo pilot algorithm based on a regressional model-based least-squares-fitting is proposed for fast-varying channels without increasing the pilot density. In [2.10], pseudo pilots are created using diversity detected data symbols. However, a fundamental problem remains with such systems - pilots are still required.

Existing blind channel estimation methods for OFDM systems can be classified as statistical or deterministic. (They are referred to as “blind” in this section, following the terminology of the references, and thereafter in this thesis, “decision directed” is used.) Examples of statistical blind channel estimation include using correlation methods [2.11] and cumulant fitting [2.12],[2.13]. These statistical approaches exhibit slow convergence, making them unsuitable for mobile radio channels, or any burst transmission. Another blind equalization criterion has been introduced in [2.14]; it does not apply to traditional OFDM systems since it relies on a transmitter without cyclic prefix (CP). Also, correlation-matching methods based on the transmitted signal cyclostationarity have been proposed in [2.15]-[2.16]. However, their possible implementation in existing systems is complicated in practice by the presence of null side carriers which affect the proposed identification results. Some algorithms based on subspace decomposition are also proposed [2.17]-[2.19] that essentially take advantage of the inherent redundancy introduced by the CP to blindly estimate the channel. In order to avoid the convergence period of the blind subspace algorithms during which the estimation is unreliable, known pilots are transmitted at the beginning of each OFDM frame. These methods are sometimes referred to as “semi-blind” despite the need for channel-sounding pilots. More recently, a new blind system was proposed in [2.20] that comprises simple linear transformation applied to blocks of symbols before they enter the OFDM system. The transform imposes a correlation structure on the transmitted blocks, which is used at the receiver to recover the channel via simple cross-correlation

operations. Channel estimation/tracking based on Kalman filtering is also well established, e.g., [2.21]-[2.22] based on the tenet that the recursive least squares (RLS) algorithm and the Kalman filtering algorithm are both better than the LMS algorithm in convergence rate and tracking capability.

The blind deterministic methods process the post DFT received blocks and exploit the finite-alphabet property of the information-bearing symbols. A blind deterministic approach based on the maximum likelihood (ML) principle has been proposed for OFDM systems in [2.23]. This method has the advantage of producing a channel estimate from a single OFDM symbol, but high complexity is needed to execute the maximization in the algorithm and there is an ambiguity in the recovered phase. To recover the phase completely, additional pilots can be inserted. Although only a few such pilots are needed, the scheme is no longer blind. To address this problem and reduce complexity, a modification of the ML-method has been introduced in [2.24] which combines different modulation schemes on adjacent subcarriers to resolve the phase ambiguity. A limitation of the ML-methods ([2.23],[2.24]) is the restriction to constant modulus modulations (PSK), otherwise they have high complexity owing to the vector maximization algorithm [2.23]. Iterative Bayesian methods that alternate between channel estimation and symbol detection/decoding have also been proposed in [2.25]-[2.27] for the case of coded OFDM systems. Furthermore, taking into account specific properties of *M*-PSK or QAM signals while utilizing an exhaustive search, a blind method has been presented in [2.28]. In comparison to the statistical methods, the deterministic ones converge much faster; however, they involve

high complexity, which becomes even higher as the constellation order increases. Finally, two deterministic blind channel estimation methods that take advantage of receive diversity and do not require an exhaustive search, are proposed in [2.29] and [2.30].

In this work [2.31], a new statistical decision-directed channel estimation technique is presented for uncoded OFDM, open loop systems. The open loop form means that no channel information is sent back to the transmitter, so there is no adaptive power control at the transmitter. It works as follows. When an OFDM symbol is received, a decision algorithm first makes primary symbol estimates of the data on each subcarrier based on constrained linear minimum mean square error criterion. These primary data estimates are then used for the MMSE channel estimation. Once the channel estimates are obtained, standard data detection is applied.

The main thrust of this work is to introduce and describe the new decision-directed channel estimation technique and compare its error performance with: optimal coherent/differential detection as the benchmark; a known decision-directed technique based on decision-directed Kalman-filtering [2.21]; and two pilot-aided OFDM schemes (block pilots and comb pilots). Here, improved performance is observed. Clearly, the presented system requires more computational complexity. Complete evaluation of the computational complexity is quite cumbersome; however; some insight is possible into computational complexity in one iteration (symbol time), through an approximate operations count of the implementation equations. Finally, the impact of assumptions in the

channel modeling is quantified using simulation, offering a feel for the performance with mismatch between the channel model and the receiver assumptions. These are: channel taps cross-correlation mismatch; and channel taps time-correlation coefficient mismatch.

The rest of the chapter is organized as follows. Section 2.1 describes the channel and OFDM signal models. The decision-directed estimation is formulated in section 2.2 followed by complexity evaluation in section 2.3. The simulations are presented in section 2.4, with the conclusions in section 2.5.

2.1 Channel and OFDM Signal Model

In general, at the *n*th OFDM symbol time, N data symbols $\{s[n,k]\}; k = 1,...,N$ are converted to the time-domain using the IFFT. The cyclic prefix of length P, although itself consuming potentially available payload capacity, is needed to preserve the subcarrier orthogonally and to eliminate the inter-OFDM symbol interference. The OFDM symbol time is $T_{\text{OFDM}} = (N+P)T_{\text{s}}$, with T_{s} the duration of the input data symbol, $s[n,k]$.

The baseband representation of the channel impulse response can be modeled by a tapped delay line (TDL), here with *M* taps [2.32]

$$h(t,\tau) = \sum_{m=1}^{M} h_m(t)\delta(\tau - \tau_m) \tag{2-1}$$

where τ_m is the delay of the *m*th tap, and the $h_m(t); m = 1,...,M$ are the complex taps. When the symbol duration is much greater than the delay spread, then assuming wide-sense stationary uncorrelated scattering (WSSUS), the TDL taps

should be correlated [2.33],[2.34]. According to [2.33], the tap correlations often lead to analytical difficulties. It is common to simply treat the taps as uncorrelated [2.32]. Clearly, this changes the channel from what was intended in the modeling. However, for the derivation of the proposed algorithm here, the $h_m(t)$'s are assumed in the usual way to be independent, wide-sense stationary (WSS), zero mean, complex Gaussian processes. The impact of finite channel taps' cross-correlation on an OFDM system is quantified in the simulations below.

In addition, it is assumed that the time-varying channel taps each have the same autocorrelation function

$$R_{h_m}(\Delta t) = \sigma_{h_m}^2 \rho(\Delta t) \tag{2-2}$$

with $\sigma_{h_m}^2; m = 1,..., M$ the corresponding channel branch coefficients' variances and $\rho(\Delta t)$ the normalized time-correlation function. For simulation, it is taken as $J_0(2\pi f_{\mathrm{D}} \Delta t)$ where f_{D} is the maximum Doppler frequency (this may correspond to a vehicle moving with the speed of $\mathrm{v} = f_{\mathrm{D}} \lambda_{\mathrm{c}}$ m/s with λ_{c} the wavelength of the arriving plane wave in meters). This form encompasses a commonly used, but major, assumption - a 2D-omnidirectional incident uncorrelated power distribution and 2D-omnidirectional antennas. However, the exact form of the correlation function is not important here in the sense that the algorithm below uses only one sample value of it. In practice, if this correlation sample value changes with time, it will do so slowly, and this allows time to track its estimate for its application in this algorithm. In fact sensitivity analysis (shown below) indicates that the

proposed decision-directed techniques are robust for a moderate mismatch of the channel's second-order statistics, i.e., the correlation function values, relative to the *a priori* information assumed by the receiver.

Using (2-1) and its assumptions, the frequency response of the time-varying dispersive channel is denoted

$$H(t,f)=\int_{-\infty}^{+\infty}h(t,\tau)e^{-j2\pi f\tau}d\tau=\sum_{m=1}^{M}h_m(t)\,e^{-j2\pi f\tau_m} \tag{2-3}$$

and its time-frequency correlation function has the well-known structure, e.g., [2.3], of

$$\begin{aligned}R_H(\Delta t,\Delta f)&\overset{\Delta}{=}E\left\{H(t+\Delta t,f+\Delta f)H^*(t,f)\right\}\\&=\sum_{m=1}^{M}R_{h_m}(\Delta t)e^{-j.2\pi\Delta f\,\tau_m}=\rho(\Delta t)\sum_{m=1}^{M}\sigma_{h_m}^2e^{-j2\pi\Delta f\,\tau_m}\\&=\sigma_H^2\rho(\Delta t)\rho_H(\Delta f)\end{aligned} \tag{2-4}$$

where σ_H^2 is the total average power gain of the channel, i.e.,

$$\sigma_H^2=\sum_m\sigma_{h_m}^2 \tag{2-5}$$

and

$$\rho_H(\Delta f)=\sum_m\frac{\sigma_{h_m}^2}{\sigma_H^2}e^{-j2\pi\Delta f\,\tau_m} \tag{2-6}$$

is the frequency correlation coefficient function.

The received waveform is sampled at $t=kT_s$ and the CP is removed before the FFT. It is assumed that there is no intercarrier interference. For deriving the algorithm, the channel $h(t,\tau)$ is approximated as remaining constant

for the OFDM symbol duration. In the simulations, the channel changes within an OFDM symbol according to eqn. (2.4). By assuming perfect synchronization (in practice, "good enough" synchronization is achieved using the preambles in standard OFDM), the received symbol of the *k*th subcarrier at the *n*th OFDM symbol time (after the FFT block) can be represented by

$$y[n,k] \approx H[n,k]s[n,k] + v[n,k] \tag{2-7}$$

where $H[n,k]$ is the frequency response of the channel at the *k*th subcarrier and the *n*th OFDM symbol time, and $v[n,k]$ is the zero mean white Gaussian noise with variance σ_v^2.

2.2 Decision-directed Channel Estimation and Data Recovery

Equation (2-7) can be written in matrix form as

$$\mathbf{y}_n \approx \mathbf{\Lambda}_n \mathbf{s}_n + \mathbf{v}_n \tag{2-8}$$

where

$$\mathbf{y}_n = \begin{bmatrix} y[n,1] & y[n,2] & \cdots & y[n,N] \end{bmatrix}^T$$

$$\mathbf{s}_n = \begin{bmatrix} s[n,1] & s[n,2] & \cdots & s[n,N] \end{bmatrix}^T$$

$$\mathbf{v}_n = \begin{bmatrix} v[n,1] & v[n,2] & \cdots & v[n,N] \end{bmatrix}^T$$

and $\mathbf{\Lambda}_n$ is a diagonal matrix with elements $H[n,k]$. With knowledge of the $H[n,k]$, the MMSE estimator of $\mathbf{s}_n$ given $\mathbf{y}_n$, is known, e.g., [2.35], to be

$$\hat{\tilde{\mathbf{s}}}_n = \mathbf{R}_s \Lambda_n^H \left[\mathbf{R}_{\mathbf{v}} + \Lambda_n \mathbf{R}_s \Lambda_n^H \right]^{-1} \mathbf{y}_n \tag{2-9}$$

where $\mathbf{R}_s = E\{\mathbf{s}\mathbf{s}^H\}$ and $\mathbf{R}_{\mathbf{v}} = E\{\mathbf{v}\mathbf{v}^H\}$. The statistics are assumed known, so here these quantities are not estimated on the fly. The matrix inversion in (2-9) is averted if the data, $\{s[n,k]\}$, is i.i.d. with variance σ_s^2, and similarly for the noise, $\{v[n,k]\}$. In this case, (2-9) simplifies to

$$\hat{\tilde{s}}[n,k] = \left(\frac{H^*[n,k]}{\frac{\sigma_v^2}{\sigma_s^2} + |H[n,k]|^2} \right) y[n,k] \; ; \; k = 1,...,N \tag{2-10}$$

In (2-10), the channels are needed at each OFDM symbol interval. Here, the idea is to estimate each subchannel in a decision-directed manner for obtaining the data estimates from (2-10). Upon reception of the *n*th OFDM symbol, the decision algorithm first obtains a primary data estimate of $s[n,k]$, denoted $\hat{s}_{pri}[n,k]$, for the *k*th subcarrier, as follows.

We assume that the unbiased MMSE estimate of the channel frequency response at *k*th subcarriers, $\hat{H}[n-1,k]$, obtained from previous OFDM symbol interval and the corresponding variance associated with the estimation process, $\sigma_{\hat{H}_k}^2[n-1]$, are both available. This is a standard type of assumption in deriving iterative algorithms, see, e.g., [2.35]. Then for the *k*th subcarrier at the end of (*n*-1)th OFDM symbol interval, the channel estimation is

$$\hat{H}[n-1,k] \approx H[n-1,k] + \sigma_{\hat{H}_k}[n-1] w_{\hat{H}_k}[n-1] \tag{2-11}$$

where the last term is the estimation noise with $w_{\hat{H}_k}[n-1]$ denoting white Gaussian noise with unity variance.

Furthermore, from the channel frequency response correlation function [see (2-4)] at $\Delta t = (N+P)T_s$ and $\Delta f = 0$, i.e., $R_H(\Delta t = T_{\text{OFDM}}, \Delta f = 0) = \sigma_H^2 \rho(T_{\text{OFDM}})$, the relation between $H[n,k]$ and $H[n-1,k]$ can be approximated by the first order autoregressive (AR) process

$$H[n,k] \approx \rho(T_{\text{OFDM}}) H[n-1,k] + \sigma_H \sqrt{1-\left|\rho(T_{\text{OFDM}})\right|^2}\, w[n,k] \tag{2-12}$$

where $w[n,k]$ is the AR white Gaussian noise process with variance 1.

It is noted here that the approximation models of (2-11) and (2-12) (and also in Chapter 3, e.g., eqn. (3-11), and in Chapter 5, eqn. (5-4)) for are only used for deriving the proposed decision algorithm. In the simulations, the correct structure of MMSE estimator (i.e., without approximation) is used for the channel estimation in (2-22) below.

Considering (2-11) and (2-12) along with (2-7), the new measurements $y[n,k]\ ; k = 1,...,N$ can be approximated as

$$y[n,k] \approx \rho(T_{\text{OFDM}}) \hat{H}[n-1,k] s_{pri}[n,k] + \tilde{v}[n,k]\ ; k = 1,...,N \tag{2-13}$$

where

$$\begin{aligned} \tilde{v}[n,k] = v[n,k] &+ \sigma_H \sqrt{1-\left|\rho(T_{\text{OFDM}})\right|^2}\, w[n,k] s_{pri}[n,k] \\ &- \rho(T_{\text{OFDM}})\, \sigma_{\hat{H}_k}[n-1] w_{\hat{H}_k}[n-1] s_{pri}[n,k] \end{aligned} \tag{2-14}$$

is the zero mean random noise with variance

$$\sigma_{\tilde{v}_k}^2 = \sigma_v^2 + \left(\sigma_H^2 - \left|\rho(T_{\text{OFDM}})\right|^2 \left(\sigma_H^2 - \sigma_{\hat{H}_k}^2[n-1]\right)\right)\sigma_s^2 \quad (2\text{-}15)$$

The equations of (2-13) can be re-written as

$$y[n,k] \approx \left(\rho(T_{\text{OFDM}})\,\hat{H}[n-1,k]s_{pri}[n,k]\right)g + \tilde{v}[n,k] \quad (2\text{-}16)$$

where $g = 1$. In the arrangement of (2-16), the key idea, also used in later chapters, of introducing g is to treat it as an unknown deterministic variable and attempt to estimate it. Now for each subchannel, there is a search over the transmitted signal constellation set to see which candidate of $s_{pri}[n,k]$ yields the MMSE of g according to the observations $y[n,k]$ constrained by $E\{\hat{g}\} = 1 + j0$.

Denoting $C = \{s_1, s_2, ..., s_L\}$ to be the set of transmitted signal constellation points with set size $|C| = L$, then the primary estimate of symbol $s[n,k]$ is

$$\hat{s}_{pri}[n,k] = \arg\min_{s_j \in C}\left(\left|g - \hat{g}_j\right|^2\right); k = 1,...,N \quad (2\text{-}17)$$

where $\hat{g}_j$; $j = 1,...,L$, is the unbiased linear minimum mean square estimator of g obtained by

$$\hat{g}_j = \left(f_j^*(\sigma_{\tilde{v}_k}^2)^{-1} f_j\right)^{-1} f_j^*(\sigma_{\tilde{v}_k}^2)^{-1} y[n,k] = y[n,k] \Big/ f_j \;;\; j = 1,...,L \quad (2\text{-}18)$$

where

$$f_j = \rho(T_{\text{OFDM}})\hat{H}[n-1,k].s_j$$

and the variance in estimation g_j is

$$E\left\{\left|g-\hat{g}_j\right|^2\right\}=\left(f_j^{*}(\sigma_{\tilde{v}_k}^2)^{-1}f_j\right)^{-1} \tag{2-19}$$

As noted in chapter 1, the proposed decision algorithms in this thesis are presented in a general format in the sense that they can be formulated for SISO or MIMO systems. For SISO systems, or for the special cases where the MIMO system is decoupled into an equivalent system of SISO channels, the estimator $\hat{g}_j$ beomes independent of the variance of noise $\tilde{v}$. (The variances in the numerator and denominator of eqn. (2-18) cancel.) As a result, the structure of the proposed decision algorithm (2-17) becomes a scaled version of the classical minimum Euclidean distance estimator which is independent of the variance of noise. This situation holds for OFDM where a wideband frequency-selective channel is treated as many SISO narrowband subchannels. Consequently, for OFDM, $\hat{g}_j$ in (2-18) becomes independent from the noise variance $\sigma_{\tilde{v}_k}^2$ as noted above and the proposed estimator $\hat{s}_{pri}[n,k]$ is also independent of $\sigma_{\tilde{v}_k}^2$ as

$$\hat{s}_{pri}[n,k]=\arg\min_{s_j\in C}\left(\left|y[n,k]-\rho(T_{OFDM})\hat{H}[n,k]s_j\right|^2\Big/\left|s_j\right|^2\right);k=1,...,N\ .$$

As an example, for L-PSK, the conventional minimum Euclidean distance detector would produce the same decisions, $\hat{s}_{pri}[n,k]$, as the proposed decision algorithm.

The inclusion of the variable g, is specifically for MIMO systems, the subject of the thesis. For MIMO, the estimators $\hat{s}_{pri}[n,k]$ and $\hat{g}_j$ turn out to be functions of the covariance of the associated noise which in turn depend on the

channel estimates at the previous symbol time. (*cf.*, eqs. (3-20) – (3-23) in chapter 3 for MIMO system, ref. [1.6] for Alamouti coding, and ref [1.5] for MRC receiver with correlated antennas.) Thus, for MIMO systems, the decision algorithm (2-17) is different to that of ML, MMSE and ZF estimators. Intuitively (i.e., from the principles of estimation), (2-17) should work better for MIMO, because of the extra statistical information about the channel estimates used in the algorithm.

Specifically, consideration of the channel estimates at the previous symbol time in the proposed decision algorithm should result in more accurate primary symbol estimates $\hat{s}_{pri}[n,k];k=1,...,N$, than those obtained from mismatched receivers. Mismatched receivers means receivers that effectively assume the channel estimate is perfect, and these are known to be suboptimal simply because the channel estimate is never perfect. For ML, MMSE and ZF estimators, one factor contributing to the imperfection is that some statistical information regarding the channel estimate at the previous time is discarded and is not used for data symbol estimation. This is not the case for the presented decision algorithm using *g*, for MIMO. This is the basis for expecting (2-17) to work better than ML, etc., in MIMO. However, this difference has not been tested; it is emphasized that it is an assumption in this thesis.

Now by having the primary data estimates, $\left\{\hat{s}_{pri}[n,k]\right\}\ ;k=1,...,N$ at hand, the TDL channel impulse response coefficients, $h_m[n];m=1,...,M$, in the duration of the *n*th OFDM symbol, can be estimated. Equation (2-8) can be modified to

$$\mathbf{y}_n \approx \hat{\mathbf{S}}_{pri}\mathbf{Q}\mathbf{h}_n + \mathbf{v}_n \tag{2-20}$$

where

$$\hat{\mathbf{S}}_{pri} = diag\left\{\left[\hat{s}_{pri}[n,1] \quad \hat{s}_{pri}[n,2] \quad \cdots \quad \hat{s}_{pri}[n,N]\right]\right\}$$

$$\mathbf{Q} = \left[\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_M\right]$$

$$\mathbf{q}_m = \frac{1}{\sqrt{N}}\left[1 \quad e^{-jm\Delta\omega T_s} \quad \cdots \quad e^{-jm\Delta\omega T_s(N-1)}\right]^T$$

with $\Delta\omega = 2\pi/NT_s$, and $\mathbf{h}_n$ is length-M vector of channel coefficients $h_m[n]; m = 1,...,M$. The $N \times M$ matrix $\mathbf{Q}$ is composed of the first M columns of the $N \times N$ unitary DFT matrix $\mathbf{F}$ defined as

$$[\mathbf{F}]_{i,k} \overset{\Delta}{=} \frac{1}{\sqrt{N}} e^{-j\frac{2\pi ik}{N}} \; ; \; i,k = 0,...,N-1 \tag{2-21}$$

Thus, the MMSE estimator of $\hat{\mathbf{h}}_n$ given $\mathbf{y}_n$, is obtained as

$$\hat{\mathbf{h}}_n = \left(\mathbf{A}^{-1} + \frac{1}{\sigma_v^2}\mathbf{Q}^H\hat{\mathbf{S}}_{pri}^H\hat{\mathbf{S}}_{pri}\mathbf{Q}\right)^{-1}(\sigma_v^2)^{-1}\mathbf{Q}^H\hat{\mathbf{S}}_{pri}^H\mathbf{y}_n \tag{2-22}$$

with $\mathbf{A} = diag\left\{\left[\ \sigma_{h_1}^2 \quad \sigma_{h_2}^2 \quad \cdots \quad \sigma_{h_M}^2 \ \right]\right\}$ and the channel estimation covariance matrix

$$\mathbf{C}_{\hat{\mathbf{h}}_n} = \left(\mathbf{A}^{-1} + \frac{1}{\sigma_v^2}\mathbf{Q}^H\hat{\mathbf{S}}_{pri}^H\hat{\mathbf{S}}_{pri}\mathbf{Q}\right)^{-1} \tag{2-23}$$

Note that the diagonal elements of $\mathbf{C}_{\hat{\mathbf{h}}_n}$, are the associated channel taps' estimation variances, i.e. $\sigma^2_{\hat{h}_m[n]} = \left[\mathbf{C}_{\hat{\mathbf{h}}_n}\right]_{m,m}; m = 1,...,M$. Finally, the MMSE vector estimate of the channel frequency response is

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{H}[n,1] \\ \hat{H}[n,2] \\ \vdots \\ \hat{H}[n,N] \end{bmatrix} = \mathbf{F} \begin{bmatrix} \hat{\mathbf{h}}_n \\ \mathbf{O}_{(N-M)\times 1} \end{bmatrix} \tag{2-24}$$

with associated error covariance matrix

$$\mathbf{C}_{\hat{\mathbf{H}}} = \mathbf{F} \begin{bmatrix} \mathbf{B}_{M\times M} & \mathbf{O}_{M\times(N-M)} \\ \mathbf{O}_{(N-M)\times M} & \mathbf{O}_{(N-M)\times(N-M)} \end{bmatrix} \mathbf{F}^H - \hat{\mathbf{H}}\hat{\mathbf{H}}^H \tag{2-25}$$

where $\mathbf{B}$ is a diagonal matrix with elements $[\mathbf{B}]_{m,m} = \sigma^2_{\hat{h}_m} + \left|\hat{h}_m\right|^2$, and accordingly, $\sigma^2_{\hat{H}_k}[n] = \left[\mathbf{C}_{\hat{\mathbf{H}}}\right]_{k,k}; k = 1,...,N$.

The matrix inversion in (2-22) can be omitted if all the signal constellation points have the same energy E_s. In this case, since $\mathbf{Q}^*\mathbf{Q} = \mathbf{I}$, (2-22) simplifies to

$$\hat{\mathbf{h}}_n = \overline{\mathbf{B}}\mathbf{Q}^H \hat{\mathbf{S}}^H_{pri} \mathbf{y}_n \tag{2-26}$$

with $$\overline{\mathbf{B}} = diag\left\{ \left[\frac{\sigma^2_{h_1}}{E_s\sigma^2_{h_1} + \sigma^2_v} \quad \frac{\sigma^2_{h_2}}{E_s\sigma^2_{h_2} + \sigma^2_v} \quad \cdots \quad \frac{\sigma^2_{h_M}}{E_s\sigma^2_{h_M} + \sigma^2_v} \right] \right\}.$$

Finally, by substituting $\hat{H}[n,k]$ into (2-10), the estimates of the transmitted symbols, $\hat{\hat{s}}[n,k]$, are obtained. The hard decision block then yields the detected symbols, $\hat{s}[n,k]$ by setting the optimal thresholds in the constellation regions of

the transmitted signal [2.32]. Figure 2-1 illustrates the block diagram of the proposed receiver after the DFT operator.

The channel estimates obtained from (2-24) and the variances $\sigma^2_{\hat{H}_k}[n]; k = 1,..., N$ also help the decision algorithm in the subsequent, (n+1)th iteration (OFDM symbol interval) to obtain the next set of primary data estimates $\hat{s}_{pri}[n+1, k]; k = 1,..., N$. For initialization of the proposed iterative algorithm in typical application (i.e. a slow channel fading rate, see below) of OFDM systems, any the algorithms developed in [2.36] can be used only once at the beginning of the whole data sequence transmission. So for the case of slow channel fading, the initialization is once only for the lifetime of the transmission. This means that, for the first OFDM symbol, any kind of training-based scheme, e.g., using a block-type or comb-type pilot-aided structure, can be used to form an estimate of the channel $\hat{H}[n = 1, k]$. From the second OFDM symbol, $n = 2$, the algorithm switches to the proposed iterative algorithm and uses the channel estimate $\hat{H}[1, k]$ that is at hand. The initial channel estimates obtained this way, are accurate enough for the algorithm to have fast convergence for the typical application of OFDM systems where there is a slow time-varying channel (using IEEE.802 standards in the 2.4 GHz band, for a maximum walking speed of 2 m/s, the maximum Doppler frequency is $f_m = 2/\lambda_c = 16$ Hz. With this Doppler shift and the OFDM time duration of $4\mu s$ for IEEE.802 standards, the fade rate normalized to the OFDM symbol duration is $f_D T_{OFDM} = 0.00064$ which corresponds to a very slow changing channel). The initialization training

requirement of the proposed algorithm for different fade rates is analyzed via simulations below, see Fig. 2-4. For the faster fade rates than “typical”, multiple initializations may be required.

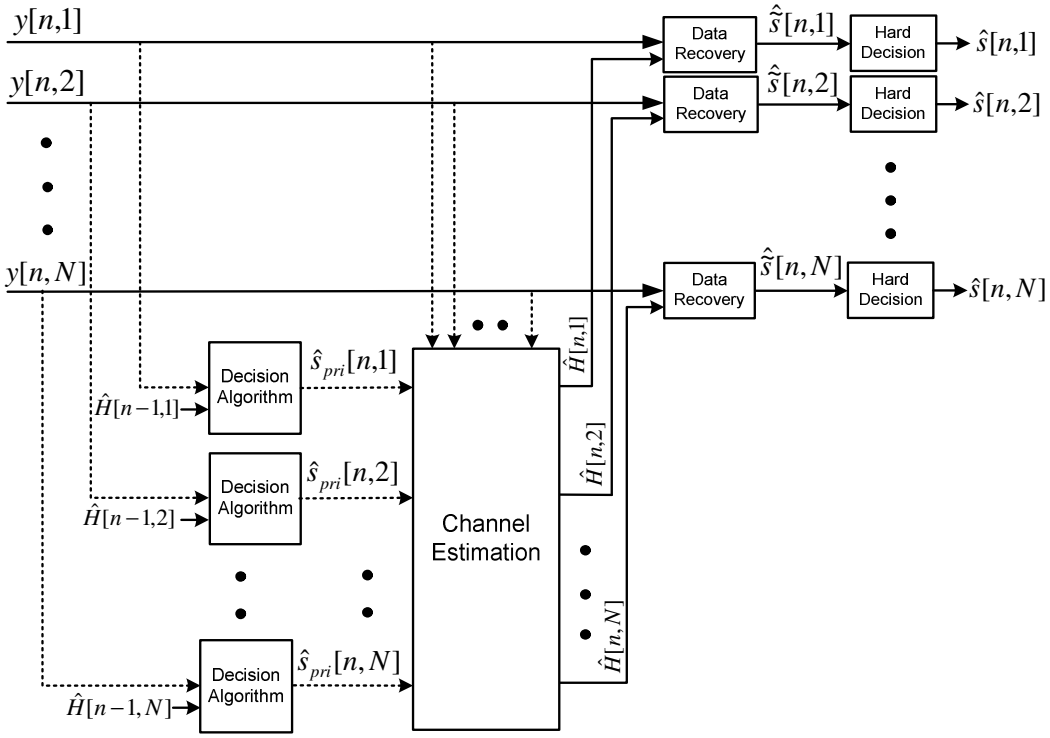

Figure 2-1 Block diagram of the proposed receiver - after the FFT operator (the CP is already removed)

2.3 Complexity Evaluation

Complete evaluation of the computational complexity is major exercise requiring the development of a working FPGA prototype. However; some insight is possible by considering the computational complexity per iteration (symbol time) using an approximate operations count of those implementation equations. An “operation” is roughly equivalent to a multiply-and-accumulate. This would be

implemented in fixed point, unlike the floating point used in the simulated performance evaluation. Fixed point operations should not exceed 32 bit accuracy for FPGA implementations, but for the range of SNRs of interest, this is not a biting constraint. The presented decision-directed channel estimation comprises four steps: the decision algorithm; the channel impulse response estimation; the channel frequency response estimation; and data recovery.

Each iteration begins with the decision algorithm. For the *k*th subcarrier, the primary data symbol estimate $\hat{s}_{pri}[n,k]$ is obtained using (2-17)-(2-18) with complexity of $O(L)$. Since there are *N* subcarriers, the overall complexity of the decision algorithm is $O(NL)$.

Next, the channel impulse response is estimated via (2-22) and (2-23). A breakdown of operations counts in implementing (2-22) is as follows. Here, the formulas in angular brackets (e.g., $\langle MN \rangle$) above and below the matrix formulas, give approximate operation counts for implementing the equation. The channel estimation is

$$\hat{\mathbf{h}}_n = \underbrace{(\sigma_v^2)^{-1}}_{\langle M \rangle} \times \underbrace{\left(\mathbf{A}^{-1} + \underbrace{\frac{1}{\sigma_v^2}}_{\langle M^2 \rangle} \times \overbrace{\mathbf{Q}^H \times \overbrace{\hat{\mathbf{S}}_{pri}^H \times \hat{\mathbf{S}}_{pri}}^{\langle N \rangle} \times \underbrace{\mathbf{Q}}_{\langle M^2 N \rangle}}^{\langle MN \rangle} \right)^{-1}}_{\langle M^3 \rangle} \times \overbrace{\underbrace{\mathbf{Q}^H}_{\langle MN \rangle} \times \overbrace{\hat{\mathbf{S}}_{pri}^H \times \mathbf{y}_n}^{\langle N \rangle}}^{\langle M^2 \rangle} \tag{2-27}$$

So this requires a total of $M^2(M+N+2)+2N(M+1)+M$ operations. The complexity in estimating the channel impulse response follows as $O(M^2(M+N))$. However, the complexity at this step is reduced significantly if all

the signal constellation points have the same energy (e.g, PSK). In this case, (2-26) is used as an alternative to (2-22). We have

$$\hat{\mathbf{h}}_n = \overbrace{\overline{\mathbf{B}}}^{\langle M \rangle} \times \underbrace{\mathbf{Q}^H}_{\langle MN \rangle} \times \overbrace{\hat{\mathbf{S}}_{pri}^H \times \mathbf{y}_n}^{\langle N \rangle} \tag{2-28}$$

which yields the overall complexity of $O(MN)$.

Now, the MMSE vector estimate of the channel frequency response, along with its associated error covariance matrix, are obtained through (2-24) and (2-25) with the breakdown of operation counts as follows

$$\hat{\mathbf{H}} = \underbrace{\mathbf{F} \times \begin{bmatrix} \hat{\mathbf{h}}_n \\ \mathbf{O}_{(N-M)\times 1} \end{bmatrix}}_{\langle NM \rangle} \tag{2-29}$$

$$\mathbf{C}_{\hat{\mathbf{H}}} = \overbrace{\mathbf{F} \times}^{\langle N^2 M \rangle} \underbrace{\begin{bmatrix} \mathbf{B}_{M\times M} & \mathbf{O}_{M\times(N-M)} \\ \mathbf{O}_{(N-M)\times M} & \mathbf{O}_{(N-M)\times(N-M)} \end{bmatrix} \times \mathbf{F}^H}_{\langle NM \rangle} - \underbrace{\hat{\mathbf{H}} \times \hat{\mathbf{H}}^H}_{\langle N^2 \rangle} \tag{2-30}$$

This requires a total of $N^2(M+1)+2NM$ operations. Thus the complexity is $O(N^2M)$. Finally, by substituting $\hat{H}[n,k]$ into (2-10), the data symbols are recovered with the complexity of $O(N)$. As a result, the presented technique has the overall complexity of $O(N(NM+L))$.

2.4 Simulation Results

We consider an uncoded OFDM system which essentially follows current IEEE 802.11a and 802.11g standards configurations [2.37],[2.38]. The IEEE 802.11g wireless LAN standard is virtually identical to the 802.11a Standard in its link layer design. The bandwidth of 300 MHz is divided into 20 MHz channels that

can be assigned to different users. In the simulations, the number of subcarriers is N = 64, and all the subcarriers are used for data transmission. (In the 802-11 Standards, subcarriers at the band edges are not used.) The length of the cyclic prefix is set as $P = 16$, so the total number of samples associated with each OFDM symbol, including both the data samples and the cyclic prefix, is $N + P = 80$.

Since $P = 16$ and $1/T_s = 20$ MHz ($T_s = 0.05\,\mu s$), the maximum delay spread for which ISI is removed is $\tau_{max} < PT_s = 0.8\,\mu s$ which corresponds to the maximum excess delay for a typical indoor environment. The long-range situation is of most interest in wireless, and this is typically non-line-of sight with Rayleigh fading as formulated above. The channel is frequency selective modeled by a TDL with 8 independent taps with an average power gain of unity. The power delay profile is exponential with a mean path delay of $0.17\,\mu s$ and an RMS delay spread of $0.2\,\mu s$. Including both the OFDM symbol and cyclic prefix, there are 80 samples per OFDM symbol, so the OFDM symbol time is $T_{OFDM} = (N + P)T_s = 4\mu s$. For the transmitted symbols, i.i.d. data is modulated with each subcarrier using QPSK, with variance 1; but the technique is readily applicable for any linearly modulated signaling.

Here, the value of $\rho(T_{OFDM}) = J_0(2\pi f_D T_{OFDM})$ and the channel tap variances are assumed to be known, i.e., estimation of these statistics is not included as part of the algorithm. Including these estimates is a relatively straightforward extension, but the goal here is to quantify the behavior of the decision-directed

estimation with known channel statistics. The impact of any mismatch in the time-correlation coefficient mismatch (perhaps caused by an incorrect estimate of f_{D}) on the proposed system's error performance is analyzed by simulation below. It is shown that the proposed technique is robust for a mild mismatch of the channel's second-order statistics, i.e., the correlation function values which relate to f_{D}, relative to the a priori information assumed by the receiver. However, a large degradation in performance is demonstrated as the mismatch increases, as expected, and this is demonstrated below.

The BER performance, from simulation, is compared with the optimum coherent/differential detection curves [2.39], decision-directed Kalman-based estimation in [2.21] and two pilot-aided OFDM schemes: a block-type pilot; and a comb-type pilot [2.40]. Throughout the simulations, the proposed scheme and the Kalman filtering both are initialized using the comb-type pilot structure.

In the block-type pilot system, OFDM channel estimation symbols are transmitted periodically. Here, each block consists of a fixed number of OFDM symbols, which is 8 in the simulations. Pilots are sent in all the sub-carriers of the first OFDM symbol of each block and the channel estimated at the beginning of the block is used for all the following OFDM symbols of the block. In the comb pilot system, N_{p} pilot tones are uniformly inserted, i.e., the pilots and the data are transmitted simultaneously on all OFDM symbols. For the latter scheme, 8 tones are used corresponding to carrier indices $\{8L+4\}; L=0,...,7$ for training and the second-order interpolation is used to estimate the channel at the data subcarrier positions at each OFDM symbol. For a fair comparison between the two pilot-

aided systems, in the block-type pilot scheme the channel is estimated only after receiving 8 consecutive OFDM symbols. This is to ensure that both pilot schemes have the same training overhead.

Figure 2-2 illustrates the BER results for *fade rate* $f_D T_{OFDM} = 0.015$ or $f_D \approx 3750\,\text{Hz}$ (for the $2.4\,\text{GHz}$ band, this corresponds to a speed of $v = f_D \lambda_c \approx 468\,\text{m/s}$; or better interpreted, for the $60\,\text{GHz}$ band, a speed of $19\,\text{m/s}$) and an SNR range of 0-25 dB. The error performance is within 1 to 3 dB of the optimum, with the gap increasing with increasing SNR. The performance is always better than optimal differential detection and this difference is almost 2 dB for high SNR, at least for the fade rate of 0.015.

The proposed technique also outperforms the decision-directed Kalman-based channel estimation. This is because in the conventional decision-directed

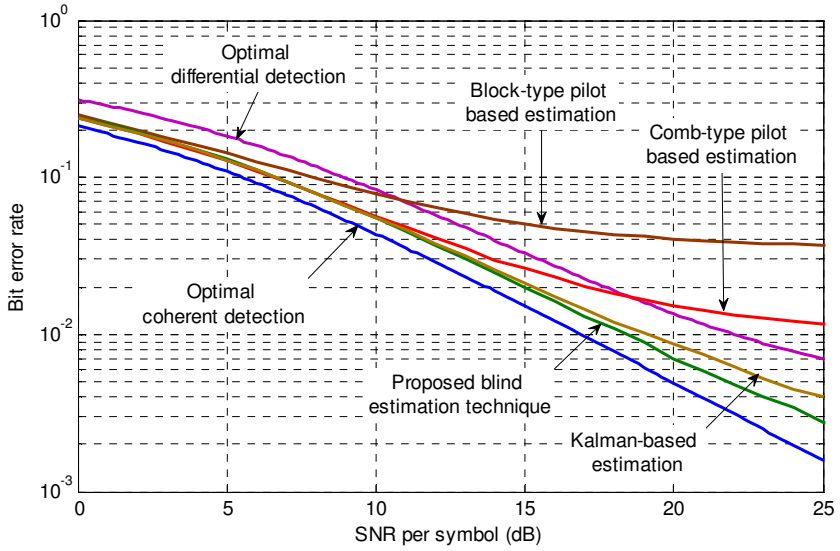

Figure 2-2 Comparison of BER performance of OFDM systems employing different channel estimation techniques for fade rate $f_D T_{OFDM} = 0.015$ for $T_{OFDM} = 4\mu s$

Kalman-filter algorithm, once the channel is predicted in the prediction stage, it is used – as if it is perfect - to obtain the coarse symbol estimates needed to formulate the measurement equations in the Kalman filtering. The decision-directed Kalman filtering uses the zero forcing criterion for obtaining the coarse symbol estimates before the update stage and uses the MMSE criterion for final data symbol recovery once the channel estimates are obtained in the update stage. Since the proposed decision algorithm yields the same result as ML estimator (for the system under consideration), more accurate primary data symbol estimates are obtained compared to the coarse symbol estimates of the decision-directed Kalman-filtering with zero forcing.

With the two pilot-aided schemes, the comb pilot system has the better error performance with the difference increasing with increasing SNR. The proposed decision-directed system performs better than the comb pilot system for SNR larger than about 10 dB, at least for the specific fade rate. For increasing SNRs, the improvement increases. This is because, in the proposed technique, the channel is estimated at all subcarrier positions, while in the comb pilot system, the pilots enable estimation of the channel response only at the pilot positions, and the CRs at the other positions can be estimated by interpolation. In addition, since $\sigma^2_{\hat{H}_k}[n-1]$ (more information from previous iteration) is also contributed in the channel estimation process, more accurate channel estimation is obtained at the $\{8L+4\}; L=0,...,7$ indices compared to the pilot sub-carriers position in comb-type pilot system which uses MMSE estimation relying only on the information from current received symbols. Furthermore, in the block-type

pilot-aided system the channel is estimated only after 8 consecutive OFDM symbols whereas in the proposed decision-directed system the channel is estimated at each OFDM symbol which results in more accurate channel estimates in a time-varying environment.

We can further compare the performance of different channel estimation techniques by setting the normalized mean square error (NMSE) criterion as a measure in estimating the channel frequency response. Here, $M = 100$ Monte Carlo runs are carried out and the NMSE at *n*th iteration (OFDM symbol time) is defined as

$$\mathrm{NMSE}(n) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{k=1}^{N} \left| \hat{H}_m[n,k] - H[n,k] \right|^2 \qquad (2\text{-}31)$$

where $\hat{H}_m[n,k]$ denotes the estimate of $H[n,k]$ at *m*th Monte Carlo run. We note that for the block-type pilot scheme, since the channel is estimated at the beginning of the block and is used for all the following OFDM symbols of the block, the definition (2-31) is not suitable and the NMSE has to be defined as the NMSE averaged over a block.

Figure 2-3 illustrates the resulted NMSE versus iteration index *n* for SNR = 20 dB and fade rate 0.015. The proposed scheme converges in three iterations (very fast) and has the least NMSE among others. This was expected beforehand from Figure 2-2, since the proposed scheme outperforms the other techniques. Furthermore, as the proposed scheme and the decision-directed Kalman filtering both are initialized using the comb-type pilot structure, the NMSEs at the first iteration are the same as the one within the comb-type pilot

system. The NMSE associated with the comb-type does not improve with iteration index *n*, since the channel estimates are obtained in the independent manner from previous iterations. The same holds for the block-type scheme where the channel estimates are obtained independently from block to block.

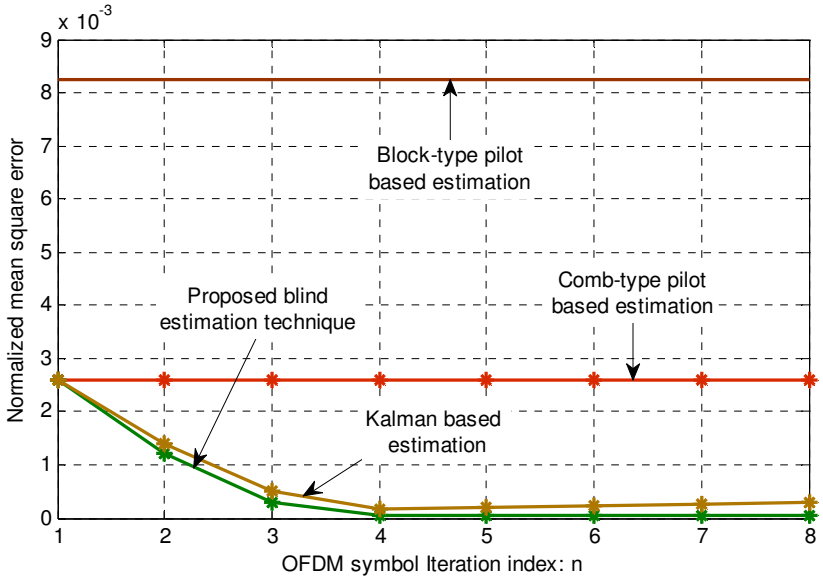

Figure 2-3 Comparison of normalized mean square error (NMSE) in estimating the channel frequency response using different estimation techniques.

SNR=20 dB and fade rate is $f_D T_{OFDM} = 0.015$. For the block-type pilot system, the NMSE is defined as the NMSE averaged over a block.

NMSE measure can also be used for analyzing the initialization training requirement of the proposed algorithm. In general, re-initialization is needed when a decision-directed channel estimator fails to track the time-variation of the channel. The re-initialization criteria is usually of the form $\mathrm{NMSE}(n) > \varepsilon$, where ε is small and positive. Figure 2-4 illustrates the initialization period for various values of maximum Doppler shift over sending 100 OFDM symbols with $\varepsilon = 5\times10^{-3}$. While re-initialization is needed after each 45 OFDM symbols at

maximum Doppler shift of $f_D = 3750\,\text{Hz}$ (i.e. $f_D T_{OFDM} = 0.015$ for $T_{OFDM} = 4\,\mu\text{s}$), no re-initialization is required over 100 OFDM symbols for values of fade rate smaller than f_D <~ $50\,\text{Hz}$ (i.e. $f_D T_{OFDM}$ <~ 0.0002 for $T_{OFDM} = 4\,\mu\text{s}$; this corresponds to the very slowly changing channel of typical applications of OFDM systems).

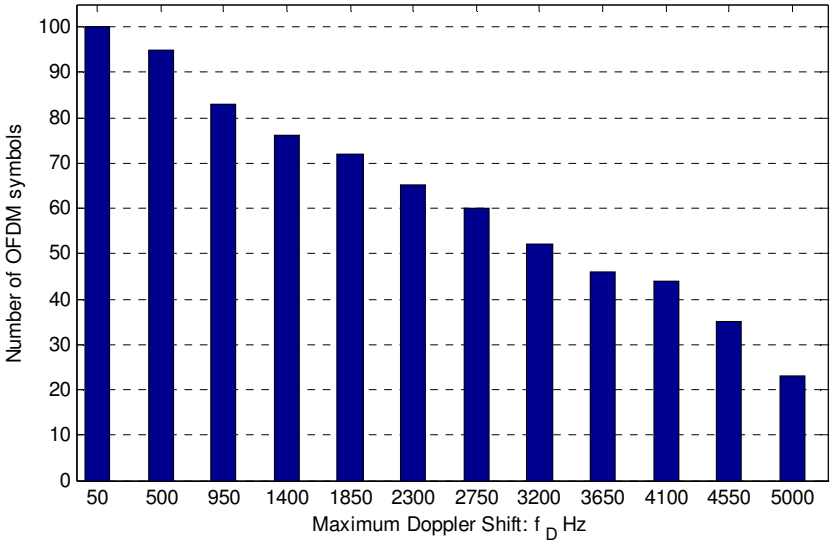

Figure 2-4 The initialization period requirement of the proposed decision-directed algorithm over sending 100 OFDM symbols and $\varepsilon = 5\times10^{-3}$

The robustness of the proposed technique to the time-variations of the channel can also be quantified by simulation. The effect of fade rate (governed by the maximum Doppler frequency) on error performance is shown in Figure 2-5 at $\text{SNR} = 20\,\text{dB}$ and the range to demonstrate this is $50\,\text{Hz} < f_D < 5000\,\text{Hz}$ (i.e. $0.0002 < f_D T_{OFDM} < 0.02$ for $T_{OFDM} = 4\,\mu\text{s}$). The algorithm is presented in a general format using the baseband representative of channel and signal model. It can be applied to an OFDM system whose terminal can be moving at practical speeds operated in any carrier frequency (e.g any ISM bands defined by the ITU-R

(International Telecommunication Union Radio-communication) such as 2.4 GHz, 5.8 GHz, 24 GHz, 60 GHz, etc.). The range of maximum Doppler frequency $50\,\mathrm{Hz} < f_{\mathrm{D}} < 5000\,\mathrm{Hz}$ given in Figure 2-5 provides us useful information regarding the performance of OFDM systems for a desired speed in any ISM band. The main reason that the *x*-axis in Figure 2-5 is labeled as maximum Doppler frequency is that for any speed at any carrier frequency, the f_{D} can be calculated according to $f_{\mathrm{D}} = \mathrm{v}/\lambda_c$ and the corresponding system performance can be read from the figure.

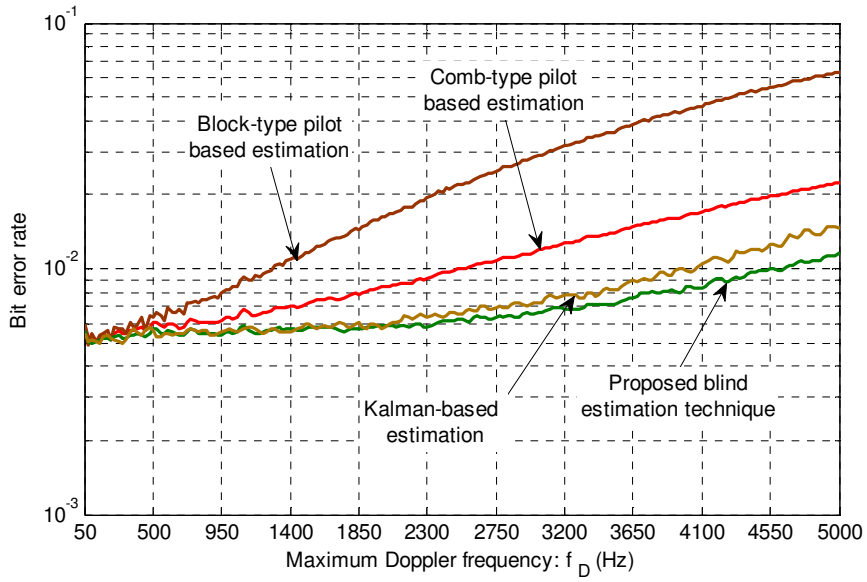

Figure 2-5 The impact of maximum Doppler frequency on system BER performance

SNR = 20 dB. For a static fading channel, the curves coincide. The range of Doppler frequency shown above corresponds to fade rate range $0.0002 \le f_{\mathrm{D}} T_{\mathrm{OFDM}} \le 0.02$ **for** $T_{\mathrm{OFDM}} = 4\,\mu\mathrm{s}$

For a given mobile speed, as the carrier frequency increases, the corresponding f_{D} increases and the advantage of the presented algorithm compared to the conventional techniques becomes more pronounced. The

values from Figure 2-2 at $\mathrm{SNR} = 20\,\mathrm{dB}$ correspond to the values in Figure 2-5 for a maximum Doppler frequency of $f_{\mathrm{D}} \approx 3750\,\mathrm{Hz}$ ($f_{\mathrm{D}} T_{\mathrm{OFDM}} = 0.015$). In this example, there is a maximum performance degradation of 16% in the BER (compared to 100% for comb pilot system) for a fading rate $f_{\mathrm{D}} T_{\mathrm{OFDM}} \leq\sim 0.01$, relative to the case of block fading in a quasi-static channel. It is clear that the pilot-aided systems are more sensitive to time-variations of the channel.

However, in general, the system performances depend on many factors such as channel fade rate, SNR and pilot scheme, and the carrier frequency used. For example, the block pilot system is always worse than the comb pilot system on this measure, and at low fade rates (approaching static fading) which can correspond to a practical speed in WLAN systems operating at 2.4 GHz band (e.g., the speed of 120 km/h in 2.4 GHz band corresponds to the maximum Doppler frequency of $267\,\mathrm{Hz}$ or fade rate $f_{\mathrm{D}} T_{\mathrm{OFDM}} = 0.0011$ for $T_{\mathrm{OFDM}} = 4\,\mu\mathrm{s}$), the four algorithms (two-pilot aided systems, Kalman-based estimation, and the proposed decision-directed system) yield nearly the same performance at all values of SNR. However, as the fade rate increases, the difference between the BER curves increases. Some other examples for the practical speed of $120\,\mathrm{km/h}$ $(= 33.33\,\mathrm{m/s})$ is presented in for the examples of the ISM bands: In 5.8 GHz band, the maximum Doppler shift is $f_{\mathrm{D}} \approx 643\,\mathrm{Hz}$, and at 24 GHz, $f_{\mathrm{D}} \approx 2667\,\mathrm{Hz}$, and finally at 60 GHz, $f_{\mathrm{D}} \approx 6667\,\mathrm{Hz}$. The benefit of the proposed algorithm over the conventional techniques is evident for carrier frequencies above about 5 GHz.

The results show that it is beneficial to use the presented decision-directed system for practical vehicular speeds at ISM bands such as 5.8 GHz, 24 GHz and 60 GHz where the use of current training-based OFDM systems becomes prohibitive (including the IEEE 802.11 configurations at the 2.4 GHz ISM band). There is only a small degradation in performance compared to quasi-static fading case as in WLAN systems. The degradation in performance occurs since the primary data symbol estimation (2-17) becomes less accurate for increasing fade rates because of extra noise, as shown in (2-16), with its higher variance given explicitly in (2-15). It should be noted here that there may well emerge a new Standard for the higher ISM bands (e.g., 24 GHz or 60 GHz), although there is no action on this as yet, and so in this sense the comparisons here may not be fair. But currently, the de facto usage is for 802.11 type standards, and many systems already use these at high ISM bands (e.g., at 60 GHz), and for this situation, the comparison is fair.

It is not clear if decision-directed channel estimation systems (igonring the initialization) or pilot-based systems are more capacity efficient, so this is investigated as well. For a benchmark, the upper bound for the approximate aggregate number of payload bits per second per Hertz (uncoded throughput) for a general OFDM system is [2.1]

$$R = \left(\frac{1}{N}\sum_{k=1}^{N} b_k\right)\frac{N}{N+P} = \frac{1}{N+P}\sum_{k=1}^{N}\log_2\left(1+\frac{\mathrm{SNR}_k}{\Gamma}\right) = \frac{1}{N+P}\log_2\left(\prod_{k=1}^{N}\left(1+\frac{\mathrm{SNR}_k}{\Gamma}\right)\right) \tag{2-32}$$

where N is the number of subchannels and P is the CP length. b_k and SNR_k denote the data rate in bits/sec/Hz, and the SNR per symbol assigned to the *k*th subchannel, respectively. Γ is called the *SNR gap* which for any coding scheme and a given target probability of error, is defined as [2.1]

$$\Gamma \stackrel{\Delta}{=} \frac{2^{2C}-1}{2^{2R}-1} = \frac{\text{SNR}}{2^{2R}-1} \tag{2-33}$$

Here, R is the rate in bits/sec/Hz and C is the maximum rate, or capacity, of an AWGN channel, i.e., C is the upper bound for the capacity of a fading channel with the same average SNR [2.41]. The use of codes, say trellis coding and/or forward error correction, reduces the gap, for example, to 1 dB for $P_e \leq 10^{-6}$. For uncoded, square QAM, and $P_e = 10^{-6}$, the SNR gap Γ is about 8.8 dB.

In Figure 2-6, the upper bound for the uncoded rate of the decision-directed OFDM system, is compared with the block and comb pilot systems. The same channel and system parameters are used as above. Both pilot systems have the same rate because they are configured (see above) to have the same transmission overhead in order to get a fair comparison. The pilot systems and decision-directed system behave similarly. In fact, the decision-directed system outperforms the pilot systems by up to a few dB in SNR, with the difference increasing with increasing SNR.

This demonstrates that the proposed system has better capacity, as defined above, than the pilot systems considered here. It appears that decision-directed detection can replace pilot systems.

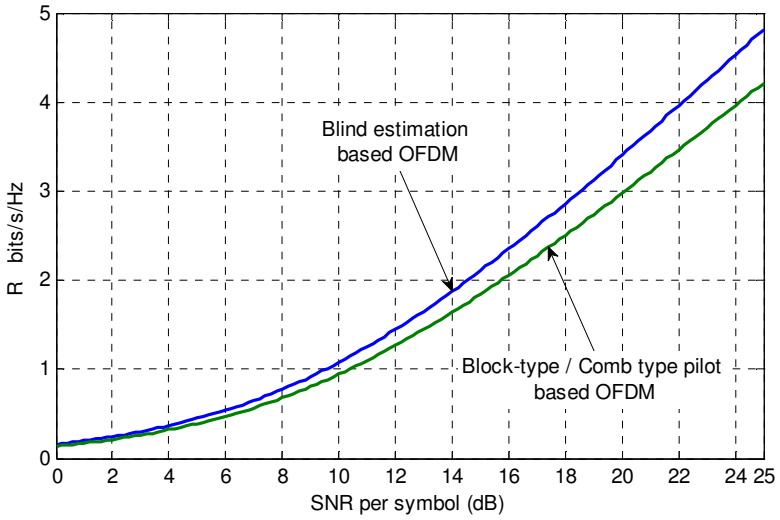

Figure 2-6 Comparison of the upper bound for aggregate data transmission rate in bits/s/Hz for decision-directed estimation and the pilot-based systems, with uncoded square QAM modulation on each subcarrier

This is a significant result although it is found for one value of pilot density only. Systems using a different density of pilots may behave differently. From the promising results above, further research seems justified on this subject.

The comparison of simulation results with benchmark and other known results puts a focus on the importance of checking the channel modeling. Thus, the rest of the simulations concerns with analyzing the impact of mismatches between the channel model and the receiver assumption.

2.4.1 Channel Taps Cross-Correlation Mismatch

In the above simulations, the frequency selective channel was generated using a TDL with independent taps. We used a receiver with the same TDL model. This is a common approach in communications modeling. But such matched models for the channel and receiver can lead to optimistic performance.

As noted above, in a T-spaced TDL channel, the taps are usually correlated in real-world channels [2.33],[2.34]. To study the impact of mismatches between the channel model (taps with finite cross-correlation) and the receiver assumption (assumed to have zero cross-correlation), a correlated TDL channel is applied and the receiver with the same structure and fixed fading rate as above is employed. Here, the taps' normalized cross-correlation matrix is modeled with a tri-diagonal correlation matrix in which

$$\rho_{i,j} = \begin{cases} 1 & ; \ i = j \\ \rho_0 & ; \ i - j = 1 \\ \rho_0^* & ; \ j - i = 1 \\ 0 & ; \ elsewhere \end{cases} \tag{2-34}$$

where $\rho_{i,j} = E\{h_i(n)h_j^*(n)\}/\sigma_{h_i}\sigma_{h_j}$; $i, j = 1,...,8$ denotes the finite complex correlation coefficient between channel taps $h_i(n)$ and $h_j(n)$. Accurate analysis of the impact of taps cross-correlation on system performance involves considering the whole region $0 \le |\rho_0| \le 1$ in the two-dimensional complex correlation domain which is rather complicated. Thus, in order to further simplify the analysis, only real values of $\rho_0 \in \Re^+$ are considered here. Figure 2-7 illustrates the resultant BER for different real values of $\rho_0 \in \Re^+$, at SNR = 20 dB.

Even with the perfect channel state information (CSI) at the receiver (lowest curve), the error performance degrades sharply for higher correlation, as $\rho_0 \in \Re^+$ increases. The proposed decision-directed scheme and the Kalman-based estimation system tracks (runs in parallel to) the perfect CSI performance curve.

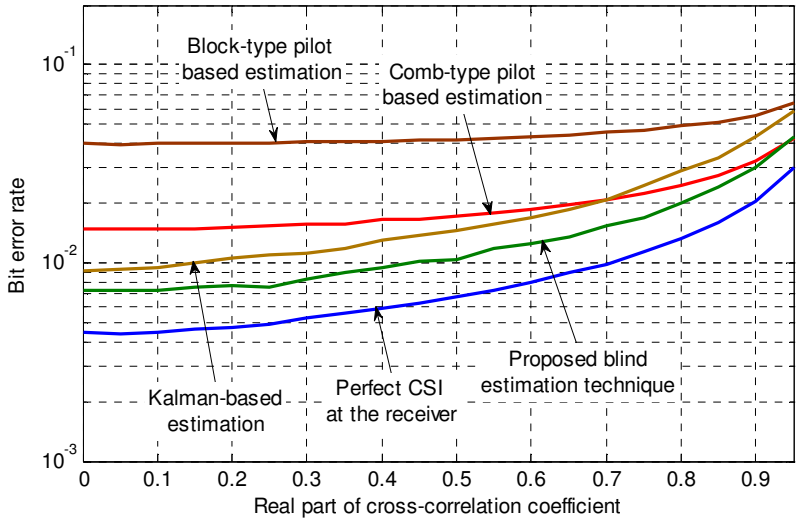

Figure 2-7 The impact of channel taps cross-correlation coefficient on BER performance at SNR = 20 dB and fade rate $f_{\mathrm{D}}T_{\mathrm{OFDM}} = 0.015$

The pilot schemes are less susceptible in error performance to the increasing taps' cross-correlations, although they start from worse performance position. For the systems considered here, and the given fading rate, the proposed system still outperforms the comb pilot system for moderately correlated ($\rho_0 <\sim 0.8 ; \rho_0 \in \Re^+$) taps. However, for high correlations, the gap closes and the comb pilot system is better than the decision-directed technique for $\rho_0 >\sim 0.92$; $\rho_0 \in \Re^+$. The decision-directed scheme outperforms the block-pilot scheme and the Kalman-based estimation for all correlations.

2.4.2 Channel Taps Time-Correlation Coefficient Mismatch

In the previous sections, the normalized time-correlation function of the channel coefficients is assumed to be known beforehand. In the channel estimation process use is made of one value of the channel normalized

correlation coefficient function, i.e. $\rho(T_{\mathrm{OFDM}}) = J_0(2\pi f_{\mathrm{D}} T_{\mathrm{OFDM}})$. However, a real-world, local autocorrelation is seldom exactly J_0 even in the main lobe region, and almost never J_0 away from the main lobe. Furthermore, estimating a correlation coefficient from finite samples introduces uncertainty depending on the number of samples used. In order to get a reliable estimate where the function becomes small, a very large number of samples is required [2.42].

Here, in order to test the performance sensitivity to the second-order statistics, a simple percentage correlation function sample mismatch is defined as

$$\begin{aligned} \eta_{miss}(T_{\mathrm{OFDM}}) &= \left| \frac{\rho(T_{\mathrm{OFDM}}) - \hat{\rho}(T_{\mathrm{OFDM}})}{\rho(T_{\mathrm{OFDM}})} \times 100 \right| \\ &= \left| \frac{J_0(2\pi f_{\mathrm{D}} T_{\mathrm{OFDM}}) - \hat{\rho}_{a_0}(T_{\mathrm{OFDM}})}{J_0(2\pi f_{\mathrm{D}} T_{\mathrm{OFDM}})} \times 100 \right| \end{aligned} \tag{2-35}$$

where the normalized time-correlation function used at the receiver takes on different values from $\rho(T_{\mathrm{OFDM}}) = J_0(2\pi f_{\mathrm{D}} T_{\mathrm{OFDM}})$. Figure 2-8 illustrates the effect of the mismatch parameter $\eta_{miss} = \eta_0$ on BER performance, given as the relative change in BER, $\left(BER_{\eta_{miss}=0} - BER_{\eta_{miss}=\eta_0}\right) / BER_{\eta_{miss}=0}$, using the example of BPSK with SNR = 20 dB. The system remains moderately insensitive to the mismatch in (2-35) up to 20% for fate rate 0.015, but for higher mismatch, there will be a larger degradation in the BER performance. Also, there is more sensitivity to the mismatch for higher fade rates. The reason is that for a given mismatch, the higher fade rate results in the addition of an error term to (2-16) which in turn causes more signal estimation inaccuracies in the decision algorithm.

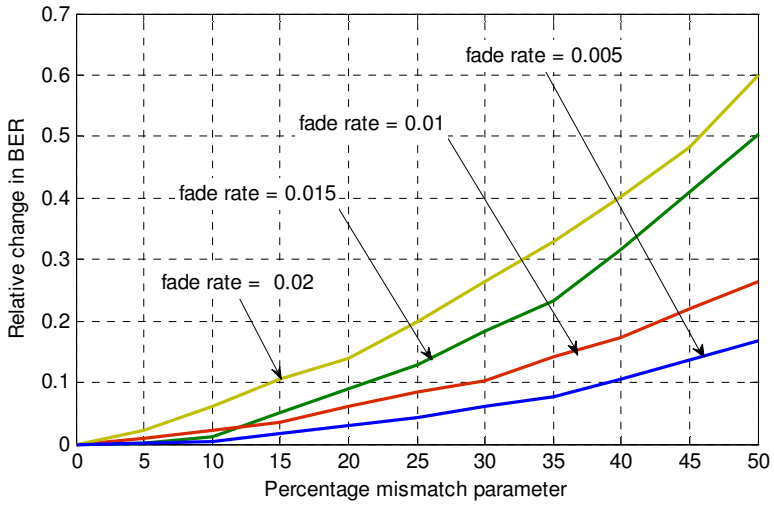

Figure 2-8 The effect of mismatch parameter on relative BER performance of system employing the proposed decision-directed channel estimation at SNR = 20 dB

Thus, for the typical applications of OFDM system, the proposed decision-directed technique is robust to the mild mismatches of the channel's second-order statistics since the interest is in very slow changing channels (e.g. for the pedestrian speed of 2 m/s in 2.4 GHz band using IEEE 802 Standards with OFDM time duration of $4\mu s$, the fade rate would be 0.00064 which corresponds to very slow changing channel). The figure helps to quantify the sensitivity of the proposed receiver to the modelling of the temporal second order statistics of the channel.

2.5 Summary and Conclusions

In this work, a new decision-directed channel estimation technique for OFDM is formulated and its performance evaluated by simulation. The simulated OFDM system essentially follows current IEEE 802.11 standards configurations, but the formulation is for a general structure. The performance is very promising.

The simulations show that the technique compares favorably with known decision-directed Kalman-based estimation, and with the training-based schemes from established research which have a fixed pilot density. Specific tie points are as follows. The bit error-rate of the decision-directed system stays within about 3 dB SNR of optimum coherent detection, with this gap decreasing with decreasing SNR. In particular, it is very close - within 1.5 dB - of the optimum, for SNR less than 15 dB at fade rate of $f_{\mathrm{D}}T_{\mathrm{OFDM}} = 0.015$. The decision-directed system always outperforms optimum differential detection and known decision-directed Kalman-based channel tracking. It also outperforms the standard pilot-based systems for SNR values greater than 10 dB (Figure 2-2). In addition, the proposed system is more spectrally efficient than pilot-aided systems for the range of simulated channels, and is more robust to the time-variations of the channel than pilot-aided systems. The results show that the decision-directed channel estimation and data detection can be employed for high Doppler frequencies. Here, the use of current OFDM systems with IEEE 802.11 configurations suffers from prohibitive training requirements. These high Doppler frequencies are of particular interest for the increasingly-used higher ISM bands (e.g., 24 GHz, or 60 GHz ISM band).

The comparison of simulation results with benchmark and other known results puts a focus on the importance of checking the channel modeling. For the channel taps' cross-correlation mismatch, the error performance of the decision-directed technique tracks that of the optimum behavior, i.e., both of these have similar sensitivity; however, the pilot techniques are less sensitive, but they are

also “less tuned” in the sense that they operate from a worse error performance datum. For high correlations between all taps (correlation coefficient more than about 0.9), the comb pilot system gains an edge on the decision-directed system. The decision-directed system outperforms the comb pilot system for lower correlations, and always outperforms the block pilot system and decision-directed Kalman-based tracking technique, under this error measure. Furthermore, it is shown that the proposed technique is robust for a moderate mismatch of the channel’s second-order statistics, i.e., the correlation function values, relative to the *a priori* information assumed by the receiver. However, a large degradation in performance is demonstrated as the mismatch increases, as expected.

REFERENCE LIST

[2.1] Y. Li, and G. L. Stuber, *Orthogonal Frequency Division Multiplexing for Wireless Communications*, Atlanta, GA: Georgia Inst. Technol*,* 2006.

[2.2] S. Thoen, L. Van der Perre, M. Engels, H. De Man, "Adaptive loading for OFDM/SDMA-based wireless networks," *IEEE Trans. Commun.,* vol. 50, no. 11, pp. 1798-1810, Nov. 2002.

[2.3] Y. Li, L.J. Cimini Jr., and N.R. Sollenberger "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.

[2.4] B. Song, L. Gui, and W. Zhang, "Comb type pilot aided channel estimation in OFDM systems with transmit diversity," *IEEE Trans. Broadcast.*, vol. 52, no. 1, pp. 50–57, Mar. 2006.

[2.5] M.-X. Chang, "A new derivation of least-squares-fitting principle for OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 726–731, Apr. 2006.

[2.6] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.,* vol. 44, no. 3, pp. 1122-1128, Aug. 1998.

[2.7] A. Yazdan-Panah, R. G. Vaughan, and R. W. Heath, "Optimizing pilot locations using feedback in OFDM systems," *IEEE Trans. Veh. Technol.* vol. 58, no. 6, pp. 2803-2814, May 2009.

[2.8] J. W. Choi and Y. H. Lee, "Optimal pilot patterns for channel estimation in OFDM systems," *IEEE Trans. Wireless Commun.,* vol. 4, no. 5, pp. 2083-2088, 2005.

[2.9] M. X. Chang, T.D. Hsieh, "Detection of OFDM signals in fast-varying channels with low-density pilot symbols," *IEEE Trans. Veh. Technol.* vol. 57, no.2, pp. 859-872, Mar. 2008.

[2.10] A. Yazdan-Panah, B. Nosrat-Makouei, and R. G. Vaughan, "OFDM with cyclic-pilot time diversity," in *Proc. IEEE Veh. Technol. Conf.* pp. 1371-1375, Sep./Oct. 2007.

[2.11] B. Muquet and M. de Courville, "Blind and semi-blind channel identification methods using second order statistics for OFDM systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, pp. 2745–2748, Mar. 1999.

[2.12] S. Chen and S. McLaughlin, “Blind channel identification based on higher-order cumulant fitting using genetic algorithms,” in *Proc. IEEE Signal Process. Workshop on Higher-Order Stat.*, pp. 184–188. Jul. 1997.

[2.13] F.-C. Zheng, S. McLaughlin, and B. Mulgrew, “Blind equalization of nonminimum phase channels: higher order cumulant based algorithm,” *IEEE Trans. Signal Process.*, vol. 41, no.2, pp. 681–691, Feb. 1993.

[2.14] M. de Courville, P. Duhamel, P. Madec, and J. Palicot, “Blind equalization of OFDM systems based on the minimization of a quadratic criterion,” in *Proc. Int. Conf. Commun.*, vol. 3, pp. 1318–1321, Jun. 1996.

[2.15] G. B. Giannakis, “Filterbanks for blind channel identification and equalization,” *IEEE Signal Process. Lett.*, vol. 4, no. 6, pp. 184–187, Jun. 1997.

[2.16] R. Heath and G. B. Giannakis, “Exploiting input cyclostationarity for blind channel identification in OFDM systems,” *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 848–856, Mar. 1999.

[2.17] B. Muquet, M. de Courville, and P. Duhamel, “Subspace-based blind and semi-blind channel estimation for OFDM systems,” *IEEE Trans. Signal Process.,* vol. 50, no. 7, pp. 1699-1712, Jul. 2002.

[2.18] Scaglione, G. B. Giannakis, and S. Barbarossa, “Redundant filterbank precoders and equalizers - Part II: blind channel estimation, synchronization and direct equalization,” *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 2007–2022, Jul. 1999.

[2.19] C. Li, S Roy, “A subspace blind channel estimation method for OFDM systems withoutcyclic prefix,” in *Proc. IEEE Veh. Technol. Conf.,* Vol. 4, pp. 2148-2152, Oct. 2001.

[2.20] A. Petropulu, R. Zhang, and R. Lin, “Blind OFDM channel estimation through simple linear precoding,” *IEEE Trans. Wireless Commun.,* vol. 3, no. 2, pp. 647-655, Mar. 2004.

[2.21] W. Chen, R Zhang, “Kalman-filter channel estimator for OFDM systems in time and frequency-selective fading environment”, *Proc. ICASSP,* Vol. 4, pp. 377-380, May 2004.

[2.22] K.Y. Han, S.W. Lee, J.S. Lim, and K.M. Sung, “Channel estimation for OFDM with fast fading channels by modified Kalman filter”, *IEEE Trans. Consum. Electron.,* vol. 50, no. 2, pp. 443-449, May 2004.

[2.23] N. Chotikakamthorn and H. Suzuki, “On identifiability of OFDM blind channel estimation,” in *Proc. Veh. Technol. Conf.*, vol. 4, pp. 2358–2361, Sep. 1999.

[2.24] M. C. Necker, and G. L. Stuber, “Totally blind channel estimation for OFDM on fast varying mobile radio channels,” *IEEE Trans. Wireless Commun.,* vol. 3, no. 5, pp. 1514-1525, Sep. 2004.

[2.25] L. Lin, J. C.-I. Chuang, and Y. Li, “Near optimal joint channel estimation and data detection for COFDM systems,” in *Proc. GLOBECOM*, vol. 2, pp. 726-730, Nov. 2000.

[2.26] B. Lu, X. Wang, and K. R. Narayanan, “LDPC-based spacetime coded OFDM systems over correlated fading channels: Analysis and receiver design,” *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 74–88, Jan. 2002.

[2.27] Z. Yang and X. Wang, “Blind detection of OFDM signals in multipath fading channels via sequential Monte Carlo,” *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 255–270, Feb. 2002.

[2.28] S. Zhou and G. B. Giannakis, “Finite-Alphabet based channel estimation for OFDM and related multicarrier systems,” *IEEE Trans. Commun.*, vol. 49, No. 8, pp. 1402–1414, Aug. 2001.

[2.29] H. Wang, Y. Lin, and B. Chen, “Blind OFDM channel estimation using receiver diversity,” in *Proc. Conf. Inf. Sci. Sys.*, Princeton, NJ, Mar. 2002.

[2.30] H. Wang, Y. Lin, and B. Chen “Data-efficient blind OFDM channel estimation using receiver diversity,” *IEEE Trans. Signal Process.,* vol. 51, no. 3, pp. 2613-2623, Oct. 2003.

[2.31] S. A. Banani, and R. G. Vaughan, “OFDM with iterative blind channel estimation,” *IEEE Trans. Veh. Technol.,* vol. 59, no. 9, pp. 4298-4308, Nov. 2010.

[2.32] J. G. Proakis, *Digital Communications,* 5th ed. New York: McGraw-Hill, 2008.

[2.33] G. L. Stuber, *Principle of Mobile Communication*, Norwel, MA: Kluwer, 1996.

[2.34] J. Sykora, “Tapped delay line model of linear randomly time-variant WSSUS channel”, *IEEE Electron. Lett.*, vol. 36, no. 19, pp. 1656 – 1657, Sep. 2000.

[2.35] H. Sayed, *Fundamentals of Adaptive Filtering,* Hoboken, NJ: Wiley, 2003.

[2.36] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, “Estimation and equalization of fading channels with random coefficients,” *Signal Process.*, vol. 53, no. 2–3, pp. 211–229, Sep. 1996.

[2.37] *High-speed physical layer in the 5 GHz, band,* IEEE std. 802.11a-1999, 1999.

[2.38] *Further higher-speed physical layer extension in the 2.4 GHz band,* IEEE std. 802.11g-2003, 2003.

[2.39] R. Prasad, *OFDM for Wireless Communications systems,* Norwood, MA: Artech House, 2004.

[2.40] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.,* vol. 48, no. 3, pp.223-229, Sep. 2002.

[2.41] Goldsmith, *Wireless Communications,* Cambridge, U.K: Cambridge Univ. Press, 2005.

[2.42] R. Vaughan and J. Bach Andersen, *Channels, Propagation and Antennas for Mobile Communications*, London, U.K.: IEE, 2003.

3: JOINT DECISION-DIRECTED CHANNEL ESTIMATION AND NONLINEARITY COMPENSATION IN MIMO SYSTEMS WITH NONLINEAR AMPLIFIERS

The use of multiple antennas at both the transmitter and receiver in wireless communications provides increased spectral efficiency compared to single antenna systems [3.1]. Nevertheless, there is not a significant commercial presence for full-MIMO systems that use large numbers of elements, which is where the potential increase of spectral efficiency becomes dramatic. One reason for this lack of commercial uptake is that the extra hardware for many-element systems is too expensive for commercial viability. Being able to use low cost amplifiers is particularly important.

The performance of MIMO systems has been studied extensively, and simulation is the basic tool for estimating the communications performance. Most studies assume that both the transmitter and receiver amplifiers are operated in the linear region and so that the MIMO channel matrix is linear. A microwave amplifier can indeed operate as a quasi-linear device under small signal conditions, and low signal distortion is possible with low power efficiency and higher cost (because they are higher power) amplifiers. It follows that there is a tradeoff between power efficiency and the resulting signal distortion. In communications systems, this tradeoff is governed by the need to limit the out-of-band interference in the radio spectrum. Nonlinear distortion at the transmitter causes interference both inside and outside the signal bandwidth. The in-band

component determines a degradation of the communications performance, often expressed as bit-error rate (BER) [3.2]-[3.3], whereas the out-of-band component affects users in adjacent frequency bands.

For more powerful MIMO systems (i.e., those with more antenna elements) to emerge as economically viable, there is a need for highly efficient power amplifiers, especially for battery-powered terminals. In other communications systems, such as satellite, the payload weight, including the on-board high power amplifiers (HPAs) is critical, and the amplifiers must run at high efficiency. In these applications, distortion compensation is also possible using signal processing.

At the receiver, the low noise amplifiers (LNAs) are key components because their gain and noise tend to dominate the sensitivity. In fact, the LNA design involves many trade-offs. These are between noise figure, gain, linearity, impedance matching, power dissipation and cost. With large-dimensioned MIMO systems where the capacity efficiency, with sufficient assumptions, becomes proportional to the number of antennas, a large number of high performance LNAs are required. This can make the cost prohibitive, and lower cost (with greater distortion) LNAs are always a pragmatic solution. However, the impact of their non-linearities will need to be compensated. The cost for the compensation using signal-processing is relatively low in the sense that the digital processors are in place anyway and the extra processing is relatively modest.

Nonlinear LNAs and their problems appear in many applications. For example, designers strive to minimize the intermodulation distortion at the

receiver by a minimum of nonlinear LNA cascades [3.4]-[3.5]. Also, to enable consumer products (e.g., GPS, etc.), an integrated receiver should minimize the number of off-chip components, particularly the number of passive filters which are relatively expensive. These considerations motivate research into highly integrated CMOS solutions which typically feature nonlinearity in the LNAs [3.6]-[3.7]. Another example is satellite diversity (a form of MIMO, using many PAs), used for fading channels in low earth orbit systems [3.8]. These issues motivate research into the impact of nonlinear amplifiers in MIMO systems.

A related problem in communications is estimating the channel in order to deploy signal processing for reliable communications. The time-varying MIMO channel has to be estimated and this is undertaken either by using pilot symbols or in a decision-directed manner. Many channel estimation techniques for linear MIMO systems have been reported. Examples of pilot-aided systems are [3.9]-[3.11]. The cost of using training-based approaches includes: a reduced payload bandwidth because the pilots bite into the communications bandwidth; and the added complexity of the implementation. Also, the training symbols can produce inaccurate channel estimates owing to the limited duration and number of training intervals available for time-varying channels [3.12].

Decision-directed techniques have become topical in research, because of the unique advantage they offer, *viz*., the channel is effectively sounded without biting into the communications bandwidth. In decision-directed techniques, the channel estimation is developed based on second or higher-order statistics of the fading process, see, for example, [3.13]-[3.32], and other references too

numerous to list. However, most blind methods that employ higher order statistics typically require a large number of data symbols, and a shortfall of symbols causes poor convergence which means poor channel estimates and decreased channel efficiency. The convergence problems associated with blind techniques can be avoided by using a so-called semiblind technique [3.20]-[3.25], i.e., employing a reduced number of training symbols together with blind statistical information.

One semi-blind approach for identification of the main eigenmode, without estimating the channel matrix itself, is presented in [3.23]. Two whitening-rotation-based algorithms for semi-blind estimation of the flat MIMO channel are presented in [3.24] and [3.25]. Such estimation procedures arise naturally in the ICA-based source separation [3.26]. The use of higher-order statistics based ICA is widespread in multiuser detection, e.g., [3.27]-[3.28]. The main advantage of ICA techniques is that, under mild mathematical conditions (independence of the sources), signal recovery is guaranteed regardless of the source constellation and spectral characteristics [3.29]. But this guarantee is for static or slow-changing channels only. Channel estimation/tracking based on Kalman/particle filtering is also well established, e.g., [3.30]-[3.31]. These are based on the tenet that the recursive least squares (RLS) algorithm and the Kalman filtering algorithm are both better than the LMS algorithm, in convergence rate and tracking capability. Kalman filtering has long been used to extend forms of the recursive least-squares (RLS) algorithm which tracks better than the standard RLS and LMS forms [3.32].

In this work [3.33], signal processing is used to compensate for the non-linearity. This is done in conjunction with formulating the proposed decision-directed channel estimation technique for MIMO systems with nonlinearity in both transmitter and receiver amplifiers. The mobile MIMO channel considered is time-selective and has Rayleigh flat fading, i.e., it is narrowband - which is where the power of MIMO dominates other communications techniques for gaining spectral efficiency. Bussgang's theorem [3.34]-[3.35] is used to model the amplifier-induced nonlinearity in the received signal. The model has time-varying coefficients which depend on both the data and channel. Using this model, a primary data vector estimate of the transmitted signal is obtained based on the constrained linear minimum mean square error (LMMSE) criterion. Then, the channel matrix is estimated/tracked using two alternative, recursive methods: statistical linearization via unscented transformation; or analytical linearization which results in a nested iterative scheme for updating the channel matrix estimate. Finally, the transmitted data vector is recovered using the channel matrix estimates.

This extends the work in [3.36] where the nonlinearity is introduced only at the receiver amplifier. Here, we formulate and describe the first proposed decision-directed channel estimation technique in the form of two alternative linearization approaches (statistical and analytical linearizations) and evaluate the performance of the systems with nonlinear amplifiers at both the transmitter and the receiver. For a nonlinear system, simulations are used to compare the results with the benchmark performance of coherent detection with perfect

channel state information (CSI) at the receiver. We also compare the results of the two linearization techniques with a look-up table (LUT) approach to the nonlinear problem. The statistical/analytical linearization approaches are marginally better for low SNR (the usual region for wireless) but the LUT is better for high SNR. In addition, for the linear system, the performance comparison is made with the result of the conventional decision-directed Kalman filtering and two pilot-aided systems already known in the signal-processing community. Here, improved performance is observed over the known techniques. The robustness of the proposed technique to time variations of the channel is also quantified and compared with that of conventional channel estimation techniques. Finally, the impact of assumptions in the channel modeling is quantified using simulation, offering a feel for the performance with time-correlation coefficient mismatch between the channel model and the assumed model at the receiver.

The rest of the chapter is organized as follows. Section 3.1 describes the system model, and this, and the other sections, are couched in terms of communications signal processing notation for MIMO. The decision-directed estimation technique is formulated in section 3.2 and the simulation results are presented in section 3.3. Section 3.4 concludes the chapter. The notation is conventional, as follows. In order to distinguish the amplifiers at the transmitter and the receiver, the sub-(super-) script “T” refers to the amplifiers at the transmitter side and the sub-(super-) script “R” refers to the amplifiers at the receiver.

3.1 Nonlinear MIMO System Model

The baseband equivalent representation of a non-linear MIMO system with M transmit and N receive antennas employing a spatial multiplexing scheme is displayed in Figure 3-1. At the transmitter, data is picked up from a constellation set with set size *L,* and the uncoded data stream is demultiplexed to M branches at each symbol time with symbol time duration T_s. Then data symbols are passed through transmit amplifiers (each with nonlinear function $f_{\mathrm{T}}(.)$) before launching from transmit antennas. It is assumed that the transmitted pulse shape is full Nyquist. For a Rayleigh flat fading MIMO channel, the received signal from the *n*th antenna is modelled by the equality

$$y_k^{(n)} = f_{\mathrm{R}}\left(\sqrt{\frac{1}{M}}\mathbf{h}_k^{(n)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k) + n_k^{(n)}\right) \;\; ; \;\; n = 1,...,N \qquad (3\text{-}1)$$

where $\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k) = \left[f_{\mathrm{T}}(s_1) \quad f_{\mathrm{T}}(s_2) \quad \cdots \quad f_{\mathrm{T}}(s_M)\right]^T$ and $f_{\mathrm{R}}(.)$ is the nonlinear function introduced by the receiver amplifier. The vector $\mathbf{s}_k$ is the $M \times 1$ transmit signal vector with i.i.d. data symbol entries each with variance σ_s^2, and n_k is the zero mean additive white Gaussian noise with variance σ_n^2. (In practice, for a terrestrial system, this noise is dominated by interference from other users,

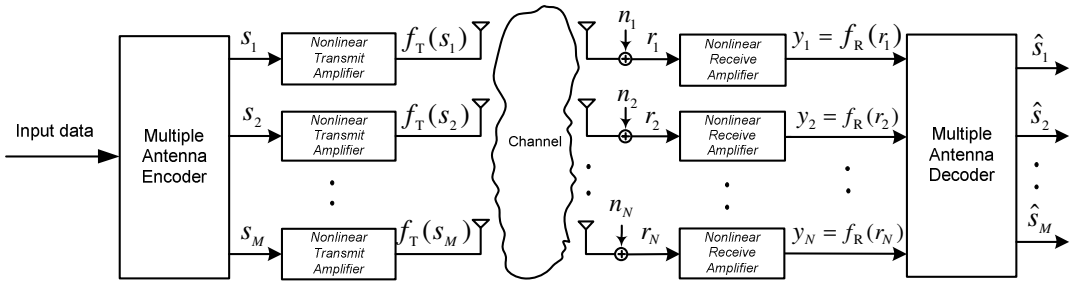

Figure 3-1 Baseband representative of nonlinear MIMO system

and other systems, i.e., the interference is assumed here to be Gaussian.) $\mathbf{h}_k^{(n)}$ is the *n*th row of channel matrix $\mathbf{H}_k = \left[\mathbf{h}_k^{(1)} \quad \mathbf{h}_k^{(2)} \quad \cdots \quad \mathbf{h}_k^{(N)}\right]^T$, comprising i.i.d. complex Gaussian entries, $h_{n,m}; m = 1,...,M$, each with variance 1; index *k* is the time index showing that at each iteration (symbol time) the channel matrix is changing.

For both amplifiers' nonlinearity ($f_{\mathrm{T}}(.)$ and $f_{\mathrm{R}}(.)$), the general memoryless AM/AM and AM/PM characteristics [3.37]-[3.38] are considered. In particular, denoting a transmitter/receiver amplifier complex input as $r(t) = x(t) + j\,y(t) = \rho(t)e^{j\varphi(t)}$, the signal at the output of a nonlinear block, becomes

$$\begin{aligned} o(t) &= F_A^{\mathrm{T/R}}\left(\rho(t)\right) e^{\,j\,F_P^{\mathrm{T/R}}(\rho(t))}.e^{j\phi(t)} \\ &= f_{\mathrm{T/R}}\left(x(t), y(t)\right) = f_{\mathrm{T/R}}\left(r(t)\right) \end{aligned} \qquad (3\text{-}2)$$

where $F_A^{\mathrm{T/R}}(.)$ and $F_P^{\mathrm{T/R}}(.)$ are the real functions for AM-to-AM and AM-to-PM conversions respectively. The super-script T/R in $F_A^{\mathrm{T/R}}(.)$ and $F_P^{\mathrm{T/R}}(.)$ means that (3-2) is written in the general format and holds for any amplifier at the transmitter or receiver with AM/AM and AM/PM characteristics.

The application of the nonlinear functions $f_{\mathrm{T}}(.)$ and $f_{\mathrm{R}}(.)$ to symbol-spaced (i.e., non-oversampled), discrete-time signals constitutes an approximation. This is because the nonlinearity can cause bandwidth expansion, and yet the sampling rate remains unchanged. The accuracy of this sampling (with aliasing) depends on the detail of the fluctuations of the continuous-time signal. This accuracy has not been checked here, however, the simulation results include the effect of this approximation.

Based on Bussgang's theorem [3.34]-[3.35], the output of a nonlinear memoryless amplifier excited by a Gaussian distributed signal can be represented by the scaled version of the original signal plus an additive noise term. Here, use is made of of Bussgang's theorem to approximate the nonlinearity introduced at the receiver side, i.e., $f_{\mathrm{R}}(.)$, with a linearized model having time-varying coefficients. By using the constant amplitude signaling with $\sigma_s^2 = E\{|s|^2\} = |s|^2$, e.g. *L*-PSK, however, the Gaussianity of the inputs of the receive amplifiers is guaranteed. (It is noted, therefore, that the formulation presented here is not applicable to 16-QAM, etc.) Under such a condition, recalluing the input of a receive amplifier as $r(t)$, the associated receive amplifier output can be represented by [3.34]

$$o(t) = K r(t) + d(t) \tag{3-3}$$

where K is a deterministic complex coefficient (expressed as a function, below), and $d(t)$ is an additive noise term. It is desirable to have zero-mean noise, uncorrelated to the input process $r(t)$. This can be achieved by setting K as [3.3]

$$K = \frac{1}{2} E\left\{ \frac{\partial f_{\mathrm{R}}(x, y)}{\partial x} - j \frac{\partial f_{\mathrm{R}}(x, y)}{\partial y} \right\} \tag{3-4}$$

where the expectation is over the ensemble of the channels and signals. But calculating this expectation likely to be very difficult, so an approximation is made using an instantaneous value, see below.

Denoting $F_A^{\mathrm{R}}(\rho)e^{jF_P^{\mathrm{R}}(\rho)} = A(\rho)$, the derivatives appearing in (3-4) are formulated as

$$\frac{\partial f_R(x,y)}{\partial x} = A'(\rho)\left(\cos^2\phi + j\sin\phi\cos\phi\right) + \frac{A(\rho)}{\rho}\left(\sin^2\phi - j\sin\phi\cos\phi\right) \tag{3-5}$$

$$-j\frac{\partial f(x,y)}{\partial y} = A'(\rho)\left(\sin^2\phi - j\sin\phi\cos\phi\right) + \frac{A(\rho)}{\rho}\left(\cos^2\phi + j\sin\phi\cos\phi\right)$$

and as a result,

$$K = \frac{1}{2}E\left\{A'(\rho) + \frac{A(\rho)}{\rho}\right\} \tag{3-6}$$

Also, the variance of the noise term in (3-3) can be obtained as

$$\sigma_d^2 = E\left\{|d|^2\right\} = E\left\{|o|^2\right\} - |K|^2 E\left\{|r|^2\right\} = E\left\{|A(\rho)|^2\right\} - |K|^2 E\left\{\rho^2\right\} \tag{3-7}$$

Applying Bussgang's theorem to (3-1),

$$y_k^{(n)} \approx K_k^{(n)}\Big|_{\mathbf{h}_k^{(n)},\mathbf{s}_k}\sqrt{\frac{1}{M}}\mathbf{h}_k^{(n)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k) + K_k^{(n)}\Big|_{\mathbf{h}_k^{(n)},\mathbf{s}_k} n_k^{(n)} + d_k^{(n)} \quad ;n = 1,\ldots,N \tag{3-8}$$

where K has been approximated with its instantaneous value

$$K_k^{(n)}\Big|_{\mathbf{h}_k^{(n)},\mathbf{s}_k} \approx \frac{1}{2}\left[A'\left(\left|\sqrt{\frac{1}{M}}\mathbf{h}_k^{(n)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k)\right|\right) + \frac{A\left(\left|\sqrt{\frac{1}{M}}\mathbf{h}_k^{(n)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k)\right|\right)}{\left|\sqrt{\frac{1}{M}}\mathbf{h}_k^{(n)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k)\right|}\right] \tag{3-9}$$

This approximation for K is used because the algorithm is to operate online, and the expectation from (3-6) is likely not available. Even if the process $A'(\rho) + \frac{A(\rho)}{\rho}$ is stationary, calculating an accurate estimate for the ensemble-average in (3-6)

by time-averaging is not appropriate for online system, because the channel is changing.

In order to recover the transmitted data vector from $\mathbf{y}_k = \begin{bmatrix} y_k^{(1)} & y_k^{(2)} & \cdots & y_k^{(N)} \end{bmatrix}$, the channel matrix $\mathbf{H}_k = \begin{bmatrix} \mathbf{h}_k^{(1)} & \mathbf{h}_k^{(2)} & \cdots & \mathbf{h}_k^{(N)} \end{bmatrix}^T$ has to be estimated. The idea is to estimate the channel matrix in a decision-directed manner at each symbol time without making use of pilot symbols or training sequences.

3.2 Decision-directed Channel Estimation

Many authors have tried to approximate the time variation of the fading channel by different dynamic models depending on the application. However, the results in [3.39] have shown that the first-order autoregressive model provides a sufficiently accurate model for time-selective fading channels. Therefore, the relation between $\mathbf{H}_k$ and $\mathbf{H}_{k-1}$ can be approximated as

$$\mathbf{H}_k \approx \rho_{\text{cor}}(T_s)\mathbf{H}_{k-1} + \sqrt{1-\left|\rho_{\text{cor}}(T_s)\right|^2}\,\mathbf{W}_k \tag{3-10}$$

where $\rho_{\text{cor}}(T_s)$ is the value of the channel coefficients' autocorrelation function, $\rho_{\text{cor}}(\Delta t)$, evaluated at $\Delta t = T_s$ and $\mathbf{W}_k$ is a matrix with zero-mean independent white Gaussian noise entries each with variance 1. When the fading statistics are unknown, $\rho_{\text{cor}}(T_s)$ can be usually estimated from the data in a training-assisted mode or decision-directed mode [3.40]. However, there is no attempt to estimate this statistic here, and for simulation, it is taken as

$J_0(2\pi f_D T_s)$ where J_0 is the zero-order Bessel function of the first kind and f_D is the maximum Doppler frequency. This form encompasses a commonly used, but major, assumption - a 2D-omnidirectional incident uncorrelated power distribution and 2D-omnidirectional antennas. The exact form of the correlation function is not important here in the sense that the algorithm below uses only one sample value of it. In practice, if this correlation sample value changes with time, it will do so slowly, and this allows time to track its estimate for its application in this algorithm. A sensitivity analysis (shown below) indicates that the proposed decision-directed technique is robust for a moderate mismatch of the channel's second-order statistics, i.e., the correlation function values, relative to the *a priori* information assumed by the receiver.

On reception of the *n*th data symbol, the receiver first employs a new decision algorithm to obtain a primary data vector estimate of the transmitted signal. The detail is descried below.

3.2.1 Decision Algorithm

Having the unbiased MMSE estimate of the channel matrix coefficients, $\hat{h}_{n,m}(k-1)$, and the corresponding error variances associated with the estimation process, $\sigma^2_{\hat{h}_{n,m}(k-1)}$, obtained from the previous symbol interval (this is a standard type of assumption in deriving iterative algorithms), the channel matrix estimation process at the $(k-1)$ th symbol time can be expressed by

$$\mathbf{H}_{k-1} \approx \hat{\mathbf{H}}_{k-1} + \mathbf{V}_{k-1} \tag{3-11}$$

where $\left[\hat{\mathbf{H}}_{k-1}\right]_{n,m} = \hat{h}_{n,m}(k-1)$ and $\mathbf{V}_{k-1}$ is a matrix with zero-mean independent white Gaussian noise entries having the corresponding variances $E\left\{ \left| \left[\mathbf{V}_{k-1}\right]_{n,m} \right|^2 \right\} = \sigma^2_{\hat{h}_{n,m}(k-1)}$, and also

$$E\left\{\mathbf{V}_{k-1}\mathbf{V}_{k-1}^H\right\} = \begin{bmatrix} \sum_{m=1}^{M} \sigma^2_{\hat{h}_{1,m}(k-1)} & 0 & \cdots & 0 \\ 0 & \sum_{m=1}^{M} \sigma^2_{\hat{h}_{2,m}(k-1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{m=1}^{M} \sigma^2_{\hat{h}_{N,m}(k-1)} \end{bmatrix} \quad \text{(3-12)}$$

Since the $\mathbf{V}_{k-1}$ is a matrix with zero-mean noise, taking expectation from both side of (3-11), we have $E\left\{\hat{\mathbf{H}}_{k-1}\right\} = E\left\{\mathbf{H}_{k-1}\right\}$ which models the unbaised channel matrix estimation process at the previous symbol time.

The approximation models of (3-10) and (3-11) are only used for deriving the proposed algorithm. In the simulations below, the correct structure of MMSE estimator (without approximation) is used for the channel estimation at each symbol time in (3-36).

Moreover, the optimal channel coefficients' linear predictions, given that the channel follows the AR(1) model of (3-10), but ignoring the addition of received data measurements $y_k^{(n)}; n = 1,..., N$, are

$$\overline{h}_{n,m}(k) = \rho_{\text{cor}}(T_s)\hat{h}_{n,m}(k-1) \quad \text{(3-13)}$$

with the corresponding prediction error variances

$$\sigma^2_{\bar{h}_{n.m}(k)} = \left|\rho_{\text{cor}}(T_s)\right|^2 \sigma^2_{\hat{h}_{n.m}(k-1)} + 1 - \left|\rho_{\text{cor}}(T_s)\right|^2 \tag{3-14}$$

From (3-10)-(3-13) and the set of equations (3-8) in matrix form, the received vector $\mathbf{y}_k$ is approximated as

$$\mathbf{y}_k \approx \frac{1}{\sqrt{M}} \rho_{\text{cor}}(T_s) \mathbf{K}_k \Big|_{\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, \ldots, \bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k} \hat{\mathbf{H}}_{k-1} \mathbf{f}_{\text{T}}(\mathbf{s}_k) + \tilde{\mathbf{v}}_k \tag{3-15}$$

where

$$\begin{aligned} \tilde{\mathbf{v}}_k &= \mathbf{K}_k \Big|_{\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, \ldots, \bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k} \mathbf{n}_k + \mathbf{d}_k \\ &+ \frac{1}{\sqrt{M}} \mathbf{K}_k \Big|_{\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, \ldots, \bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k} \left(\sqrt{1 - \left|\rho_{\text{cor}}(T_s)\right|^2} \mathbf{W}_k + \rho_{\text{cor}}(T_s) \mathbf{V}_{k-1} \right) \mathbf{f}_{\text{T}}(\mathbf{s}_k) \end{aligned} \tag{3-16}$$

in which

$$\mathbf{K}_k \Big|_{\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, \ldots, \bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k} = diag\left[K_k^{(1)} \big|_{\bar{\mathbf{h}}_k^{(1)}, \mathbf{s}_k} \quad K_k^{(2)} \big|_{\bar{\mathbf{h}}_k^{(2)}, \mathbf{s}_k} \quad \cdots \quad K_k^{(N)} \big|_{\bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k} \right] \tag{3-17}$$

$$\mathbf{d}_k \Big|_{\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, \ldots, \bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k} = \left[d_k^{(1)} \big|_{\bar{\mathbf{h}}_k^{(1)}, \mathbf{s}_k} \quad d_k^{(2)} \big|_{\bar{\mathbf{h}}_k^{(2)}, \mathbf{s}_k} \quad \cdots \quad d_k^{(N)} \big|_{\bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k} \right]^T \tag{3-18}$$

In the above expressions, the notation $(.)\Big|_{\bar{\mathbf{h}}_{k-1}^{(1)}, \bar{\mathbf{h}}_{k-1}^{(2)}, \ldots, \bar{\mathbf{h}}_{k-1}^{(N)}, \mathbf{s}_k}$ is used to emphasize that the noise vector $\mathbf{d}_k$ and matrix $\mathbf{K}_k$ depend on $\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, \ldots, \bar{\mathbf{h}}_k^{(N)}$, and $\mathbf{s}_k$. The equation of (3-15) can be re-written as

$$\mathbf{y}_k \approx \left(\frac{1}{\sqrt{M}} \rho_{\text{cor}}(T_s) \mathbf{K}_k \Big|_{\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, \ldots, \bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k^{pri}} \hat{\mathbf{H}}_{k-1} \, \mathrm{f}_{\text{T}}(\mathbf{s}_k^{pri}) \right) b + \tilde{\mathbf{v}}_k \tag{3-19}$$

where $b = 1 + j0$. In the arrangement of (3-19), b is introduced, as in (2.16), as an unknown deterministic variable to be estimated. The receiver searches over the *M*-dimensional transmit data vector constellation set $C^M = \left\{ \mathbf{s}_k^1 \quad \mathbf{s}_k^2 \quad \cdots \quad \mathbf{s}_k^{L^M} \right\}$

to see which candidate of $\mathbf{s}_k^{pri}$ yields the MMSE estimate of *b* according to the observations y_k constrained by $E\{\hat{b}\}=1+j0$.

The primary data vector estimate is

$$\hat{\mathbf{s}}_k^{pri} = \arg \min_{\substack{\mathbf{s}_k^i \in C^M \\ i=1,...,L^M}} \left(\left| b - \hat{b}_i \right|^2 \right) \tag{3-20}$$

where $\hat{b}_i$ is the unbiased LMMSE of b obtained by

$$\hat{b}_i = \left(\mathbf{f}_i{}^H \mathbf{R}_{\tilde{\mathbf{v}}_k^i}^{-1} \mathbf{f}_i \right)^{-1} \mathbf{f}_i{}^H \mathbf{R}_{\tilde{\mathbf{v}}_k^i}^{-1} \mathbf{y}_k \tag{3-21}$$

with

$$\mathbf{f}_i = \frac{1}{\sqrt{M}} \rho_{\mathrm{cor}}(T_s) \mathbf{K}_k \Big|_{\bar{\mathbf{h}}_k^{(1)}, \bar{\mathbf{h}}_k^{(2)}, ..., \bar{\mathbf{h}}_k^{(N)}, \mathbf{s}_k^i} \hat{\mathbf{H}}_{k-1} \mathbf{f}_{\mathrm{T}}(\mathbf{s}_k^i) \tag{3-22}$$

and $\mathbf{R}_{\tilde{\mathbf{v}}_k^i} = E\left\{ \tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^H \middle| \mathbf{s}_k = \mathbf{s}_k^i \right\}$ is the noise covariance matrix calculated at $\mathbf{s}_k = \mathbf{s}_k^i$. According to (3-12) and since the elements in $\mathbf{d}_k$ are independent, matrix $\mathbf{R}_{\tilde{\mathbf{v}}_k^i}$ is a $N \times N$ diagonal matrix with *n*th diagonal element as

$$\begin{aligned} \left[\mathbf{R}_{\tilde{\mathbf{v}}_k^i} \right]_{n,n} \approx & \left| A \left(\left| \sqrt{\frac{1}{M}} \rho_{\mathrm{cor}}(T_s) \hat{\mathbf{h}}_{k-1}^{(n)} \mathbf{f}_{\mathrm{T}}(\mathbf{s}_k^i) \right| \right) \right|^2 - \left| K_k^{(n)} |_{\bar{\mathbf{h}}_k^{(n)}, \mathbf{s}_k^i} \right|^2 \left| \sqrt{\frac{1}{M}} \rho_{\mathrm{cor}}(T_s) \hat{\mathbf{h}}_{k-1}^{(n)} \mathbf{f}_{\mathrm{T}}(\mathbf{s}_k^i) \right|^2 \\ & + \left| K_k^{(n)} |_{\bar{\mathbf{h}}_k^{(n)}, \mathbf{s}_k^i} \right|^2 \left(\sigma_n^2 + \frac{\left\| \mathbf{f}_{\mathrm{T}}(\mathbf{s}_k^i) \right\|^2}{M} \left(1 - \left| \rho_{\mathrm{cor}}(T_s) \right|^2 + \left| \rho_{\mathrm{cor}}(T_s) \right|^2 \sum_{m=1}^{M} \sigma_{\hat{h}_{n,m}(k-1)}^2 \right) \right)^2 \end{aligned} \tag{3-23}$$

In calculating $\left[\mathbf{R}_{\tilde{\mathbf{v}}_k^i} \right]_{n,n}$, the noise variance $\sigma_{d_k^{(n)}}^2$ is substituted by its instantaneous value since calculating the expectation in (3-7) is difficult. $\hat{b}_i$ is a function of $\mathbf{R}_{\tilde{\mathbf{v}}_k^i}$ which, in turn, depends on $\sigma_{\hat{h}_{n,m}(k-1)}^2; m = 1,...,M; n = 1,...,N$.

Consideration of these in the above equations makes the new decision algorithm yield a more accurate primary symbol data vector estimate $\hat{\mathbf{s}}_k^{pri}$.

3.2.2 Channel Estimation

Having the primary data vector estimate $\hat{\mathbf{s}}_k^{pri}$ at hand, the channel matrix estimate can now be refined in the form of two, alternative methods: using statistical linearization via unscented transformation; or using analytical linearization which results in a locally (nested) iterative scheme for updating the channel matrix estimate. Both of these methods are presented below.

3.2.2.1 Statistical Linearization

Here, the knowledge gained from observing the measurements $y_k^{(n)}; n = 1,..., N$ are used to refine the predicted channel vector $\overline{\mathbf{h}}_k^{(n)} = \left[\overline{h}_{n,1}(k) \quad \overline{h}_{n,2}(k) \quad \cdots \quad \overline{h}_{n,M}(k)\right]$ $; n = 1,..., N$. At first, the new set of observations is generated as

$$\begin{aligned} z_k^{(n)} &= \left| y_k^{(n)} \right| = \left| F_A^{\mathrm{R}}\left(\rho_k^{(n)}\right) e^{\,j F_P^{\mathrm{R}}\left(\rho_k^{(n)}\right)} . e^{\,j\phi(t)} \right| = F_A^{\mathrm{R}}\left(\rho_k^{(n)}\right) \\ &= F_A^{\mathrm{R}}\left(\left| \sqrt{\frac{1}{M}} \mathbf{h}_k^{(n)} \mathbf{f}_{\mathrm{T}}(\mathbf{s}_k) + n_k^{(n)} \right| \right); n = 1,..., N \end{aligned} \tag{3-24}$$

The Unscented Transformation (UT) [3.41]-[3.42] is to handle the nonlinearities in (3-24). However, the standard UT is characterized with real-valued random variables. As a result, to fit the UT principles to our problem involved with complex random variables, a summary of UT is described below.

Consider propagating a *p*-dimensional real-valued random vector a with mean $\bar{\mathbf{a}}$ and covariance $\mathbf{C}_{\mathbf{a}}$, through an arbitrary nonlinear function $\mathbf{g}:\Re^{n_{\mathbf{a}}=p}\to\Re^{n_{\mathbf{b}}}$, to produce a random variable $\mathbf{b}=\mathbf{g}(\mathbf{a})$. A set of $2p+1$ points, called sigma points, are generated by the following algorithm whose sample mean and sample covariance are $\bar{\mathbf{a}}$ and $\mathbf{C}_{\mathbf{a}}$, respectively

$$\begin{aligned}
&\boldsymbol{\alpha}_0=\bar{\mathbf{a}} \qquad\qquad ; \quad \xi_0=\frac{\kappa}{(p+\kappa)}\\
&\boldsymbol{\alpha}_i=\bar{\mathbf{a}}+\left(\sqrt{(p+\kappa)\mathbf{C}_{\mathbf{a}}}\right)_i \quad ; \quad \xi_i=\frac{1}{2(p+\kappa)} \qquad i=1,\ldots,p\\
&\boldsymbol{\alpha}_i=\bar{\mathbf{a}}-\left(\sqrt{(p+\kappa)\mathbf{C}_{\mathbf{a}}}\right)_i \quad ; \quad \xi_i=\frac{1}{2(p+\kappa)} \qquad i=p+1,\ldots,2p
\end{aligned} \tag{3-25}$$

where κ is a scaling parameter such that $\kappa+p\neq 0;\kappa\in\Re$, and $\left(\sqrt{(p+\kappa)\mathbf{C}_{\mathbf{a}}}\right)_i$ is the *i*th column of the matrix square root of $(p+\kappa)\mathbf{C}_{\mathbf{a}}$, and ξ_i is the weight that is associated with the *i*th point. The weights are normalized; that is, they satisfy $\sum_{i=0}^{2p}\xi_i=1$. The set of samples chosen by (3-25) are guaranteed to have the same sample mean, covariance, and all higher, odd-ordered central moments, as the distribution of the random vector $\mathbf{a}$. The matrix square root and κ affect the fourth and higher order sample moments of the sigma points [3.42]. Now each sigma point is propagated through the nonlinear function $\mathbf{g}$,

$$\boldsymbol{\beta}_i=\mathbf{g}(\boldsymbol{\alpha}_i) \qquad ; \qquad i=0,\ldots,2p \tag{3-26}$$

and the first two moments of $\mathbf{b}$ are computed as follows

$$\bar{\mathbf{b}} = \sum_{i=0}^{2p} \xi_i \boldsymbol{\beta}_i \tag{3-27}$$

$$\mathbf{C}_{\mathbf{b}} = \sum_{i=0}^{2p} \xi_i \left(\boldsymbol{\beta}_i - \bar{\mathbf{b}}\right)\left(\boldsymbol{\beta}_i - \bar{\mathbf{b}}\right)^H \tag{3-28}$$

To fit the UT principle to our problem, all the real and imaginary parts of $h_{n,m}(k); m = 1,..., M$ are stacked into $2M \times 1$ vector $\mathbf{a}_k^{(n)}$ as follows

$$\mathbf{a}_k^{(n)} = \left[\mathrm{Re}\{h_{n,1}(k)\} \quad \mathrm{Im}\{h_{n,1}(k)\} \quad \cdots \quad \mathrm{Re}\{h_{n,M}(k)\} \quad \mathrm{Im}\{h_{n,M}(k)\}\right]^T \tag{3-29}$$

Considering the above arrangement along with (3-13)-(3-14), $\mathbf{a}_k^{(n)}$ would be a random vector with mean $\bar{\mathbf{a}}_k^{(n)}$ and $2M \times 2M$ diagonal covariance matrix $\mathbf{C}_{\mathbf{a}_k^{(n)}}$ as

$$\bar{\mathbf{a}}_k^{(n)} = \left[\mathrm{Re}\{\bar{h}_{n,1}(k)\} \quad \mathrm{Im}\{\bar{h}_{n,1}(k)\} \quad \cdots \quad \mathrm{Re}\{\bar{h}_{n,M}(k)\} \quad \mathrm{Im}\{\bar{h}_{n,M}(k)\}\right]^T \tag{3-30}$$

$$\mathbf{C}_{\bar{\mathbf{a}}_k^{(n)}} = diag\left(\left[\sigma^2_{\mathrm{Re}\{\bar{h}_{n,1}(k)\}} \quad \sigma^2_{\mathrm{Im}\{\bar{h}_{n,1}(k)\}} \quad \cdots \quad \sigma^2_{\mathrm{Re}\{\bar{h}_{n,M}(k)\}} \quad \sigma^2_{\mathrm{Im}\{\bar{h}_{n,M}(k)\}}\right]\right) \tag{3-31}$$

With the assumption of a wide sense stationary uncorrelated scattering (WSSUS), $\mathrm{Re}\{h_{n,m}\}$ and $\mathrm{Im}\{h_{n,m}\}$ are uncorrelated random variables with equal variances. Finite-sample estimates of these variances are not equal, but are considered "asymptotically equal" by averaging over sufficient realizations [2.33]). As a result, for a complex channel coefficient, $h_{n,m} = \mathrm{Re}\{h_{n,m}\} + j\,\mathrm{Im}\{h_{n,m}\}$,

$$\begin{aligned} \sigma^2_{h_{n,m}} &\overset{\Delta}{=} \frac{1}{2} E\left\{\left(h_{n,m} - \bar{h}_{n,m}\right)\left(h_{n,m} - \bar{h}_{n,m}\right)^*\right\} \\ &= \frac{1}{2}\left(\sigma^2_{\mathrm{Re}\{h_{n,m}\}} + \sigma^2_{\mathrm{Im}\{h_{n,m}\}}\right) \approx \sigma^2_{\mathrm{Re}\{h_{n,m}\}} \approx \sigma^2_{\mathrm{Im}\{h_{n,m}\}} \end{aligned} \tag{3-32}$$

Thus, $\mathbf{C}_{\bar{\mathbf{a}}_k^{(n)}}$ in (3-31) is approximated by

$$\mathbf{C}_{\bar{\mathbf{a}}_k^{(n)}} = diag\left(\left[\sigma^2_{\bar{h}_{n,1}(k)} \quad \sigma^2_{\bar{h}_{n,1}(k)} \quad \cdots \quad \sigma^2_{\bar{h}_{n,M}(k)} \quad \sigma^2_{\bar{h}_{n,M}(k)}\right]\right) \tag{3-33}$$

Now, $F_A^{\mathrm{R}}(.)$ in (3-24) is taken as the nonlinear function with the corresponding output, $z_k^{(n)}$ and $4M+1$ sigma points, $\{\boldsymbol{\alpha}_i\}_{i=0}^{4M}$ are generated according to (3-25). These sigma points are passed through the nonlinear function $F_A^{\mathrm{R}}\left(\left|\sqrt{1/M}\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T\boldsymbol{\alpha}_i\right|\right)$ to yield the prediction measurement

$$\bar{z}_k^{(n)} = \sum_{i=0}^{4M}\xi_i F_A^{\mathrm{R}}\left(\left|\sqrt{\frac{1}{M}}\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T\boldsymbol{\alpha}_i\right|\right) \tag{3-34}$$

with variance

$$\sigma^2_{\bar{z}_k^{(n)}} = \sum_{i=0}^{4M}\xi_i\left(F_A^{\mathrm{R}}\left(\left|\sqrt{\frac{1}{M}}\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T\boldsymbol{\alpha}_i\right|\right)-\bar{z}_k^{(n)}\right)\left(F_A^{\mathrm{R}}\left(\left|\sqrt{\frac{1}{M}}\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T\boldsymbol{\alpha}_i\right|\right)-\bar{z}_k^{(n)}\right)^* \tag{3-35}$$

The problem of refining the channel vector estimate $\bar{\mathbf{h}}_k^{(n)}$, is finalized with a new estimation problem as follows. We wish to linearly estimate the random vector $\mathbf{a}_k^{(n)}$ as in (3-29) with mean $\bar{\mathbf{a}}_k^{(n)}$ and covariance matrix $\mathbf{C}_{\bar{\mathbf{a}}_k^{(n)}} = diag\left(\left[\sigma^2_{\bar{h}_{n,1}(k)} \quad \sigma^2_{\bar{h}_{n,1}(k)} \quad \cdots \quad \sigma^2_{\bar{h}_{n,M}(k)} \quad \sigma^2_{\bar{h}_{n,M}(k)}\right]\right)$, from measurement $\upsilon_k^{(n)} = \left(z_k^{(n)} - \bar{z}_k^{(n)}\right)$ in the minimum mean square sense ($\upsilon_k^{(n)}$ is the innovation term). The LMMSE estimator is [3.43]

$$\hat{\mathbf{a}}_k^{(n)} = \bar{\mathbf{a}}_k^{(n)} + \mathbf{C}_{\mathbf{a},\upsilon}\mathbf{C}_\upsilon^{-1}\upsilon_k^{(n)} = \bar{\mathbf{a}}_k^{(n)} + \mathbf{C}_{\mathbf{a},\upsilon}\mathbf{C}_\upsilon^{-1}\left(z_k^{(n)} - \bar{z}_k^{(n)}\right) \tag{3-36}$$

where $\mathbf{C}_\upsilon = \mathbf{C}_z = \sigma^2_{\bar{z}_k^{(n)}}$

$$\mathbf{C}_{\mathbf{a},\upsilon} = \sum_{i=0}^{4M} \xi_i \,(\boldsymbol{\alpha}_i - \overline{\mathbf{a}}_k^{\,(n)}) \left(F_A^{\mathrm{R}} \left(\left| \sqrt{1/M} \left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri}) \right)^T \boldsymbol{\alpha}_i \right| \right) - \overline{z}_k^{(n)} \right)^*$$

The associated estimation error covariance matrix is

$$\mathbf{C}_{\hat{\mathbf{a}}_k^{(n)}} = \mathbf{C}_{\overline{\mathbf{a}}_k^{(n)}} - \mathbf{C}_{\mathbf{a},\upsilon} \mathbf{C}_{\upsilon}^{-1} \mathbf{C}_{\upsilon,\mathbf{a}} \tag{3-37}$$

Finally, $\hat{h}_{n,m}(k)$ which is the *m*th element of vector $\hat{\mathbf{h}}_k^{(n)}$ is obtained as

$$\hat{h}_{n,m}(k) = \left[\hat{\mathbf{h}}_k^{(n)} \right]_m = \left[\hat{\mathbf{a}}_k^{(n)} \right]_{2m-1} + j \left[\hat{\mathbf{a}}_k^{(n)} \right]_{2m} \quad ; \quad m = 1,...,M \tag{3-38}$$

with approximated estimation variance $\sigma^2_{\hat{h}_{n,m}(k)} \approx \frac{1}{2} \left(\left[\mathbf{C}_{\hat{\mathbf{a}}_k^{(n)}} \right]_{2m-1,2m-1} + \left[\mathbf{C}_{\hat{\mathbf{a}}_k^{(n)}} \right]_{2m,2m} \right)$.

Each of the channel vectors $\hat{\mathbf{h}}_k^{(n)}; n = 1,...,N$ is obtained individually from the algorithm described above.

3.2.2.2 Analytical Linearization

Here, the set of equations (3-8) are rewritten as

$$y_k^{(n)} \approx \underbrace{\sqrt{\frac{1}{M}} K_k^{(n)} \Big|_{\overline{\mathbf{h}}_k^{(n)}, \hat{\mathbf{s}}_k^{pri}} \left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri}) \right)^T \left(\mathbf{h}_k^{(n)} \right)^T}_{\Omega\left(\left(\mathbf{h}_k^{(n)} \right)^T \right)} + \underbrace{K_k^{(n)} n_k^{(n)} + d_k^{(n)}}_{\tilde{w}_k} \tag{3-39}$$

where $K_k^{(n)} \Big|_{\overline{\mathbf{h}}_k^{(n)}, \hat{\mathbf{s}}_k^{pri}}$ is the value of $K_k^{(n)}$ evaluated at $(\overline{\mathbf{h}}_k^{(n)}, \hat{\mathbf{s}}_k^{pri})$.

For brevity, let $\mathbf{x} = \left(\mathbf{h}_k^{(n)} \right)^T$ and $\mathbf{Y}_k^{(n)} = \left\{ y_k^{(n)}, y_{k-1}^{(n)}, ..., y_1^{(n)} \right\}$ denote the set of all sampled received signal from *n*th antenna up to *k*th symbol time. Using Bayes' theorem on the conditional density for memoryless sensor systems yields

$$p(\mathbf{x}\left|\mathbf{Y}_k^{(n)}\right.) = p(\mathbf{x}\left|\mathbf{Y}_{k-1}^{(n)}, y_k^{(n)}\right.) = \frac{1}{c}\, p(y_k^{(n)}\left|\mathbf{x}, \mathbf{Y}_{k-1}^{(n)}\right.) p(\mathbf{x}\left|\mathbf{Y}_{k-1}^{(n)}\right.) = \frac{1}{c}\, p(y_k^{(n)}\left|\mathbf{x}\right.) p(\mathbf{x}\left|\mathbf{Y}_{k-1}^{(n)}\right.) \tag{3-40}$$

with *c* a normalization constant

$$c = \int_{\mathbf{x}} p(y_k^{(n)}\left|\mathbf{x}\right.) p(\mathbf{x}\left|\mathbf{Y}_{k-1}^{(n)}\right.) d\mathbf{x} \tag{3-41}$$

The approximation is made such that both the predicted channel matrix elements in (3-13) and the noise term in (3-39) are considered normally distributed. Thus, the posterior probability density $p(\mathbf{x}\left|\mathbf{Y}_k^{(n)}\right.)$, would be the product of two Gaussians. Therefore, the task is now to find the maximum of $p(\mathbf{x}\left|\mathbf{Y}_k^{(n)}\right.)$ for MAP estimation. Equivalently, its logarithm can be maximized. After the elimination of the irrelevant constants and factors, it boils down to minimization the following function

$$\Phi(\mathbf{x}) = \frac{1}{2}(\mathbf{x}\text{-}\bar{\mathbf{x}})^H \mathbf{C}_{\bar{\mathbf{x}}}^{-1}(\mathbf{x}\text{-}\bar{\mathbf{x}}) + \frac{1}{2}(y_k^{(n)}\text{-}\Omega(\mathbf{x}))^H \mathbf{C}_{\tilde{w}}^{-1}(y_k^{(n)}\text{-}\Omega(\mathbf{x})) \tag{3-42}$$

with $\bar{\mathbf{x}} = (\bar{\mathbf{h}}_k^{(n)})^T$ and

$$\mathbf{C}_{\bar{\mathbf{x}}} = diag\left[\sigma_{\bar{h}_{n,1}(k)}^2 \quad \sigma_{\bar{h}_{n,2}(k)}^2 \quad \cdots \quad \sigma_{\bar{h}_{n,M}(k)}^2\right] \tag{3-43}$$

and $\mathbf{C}_{\tilde{w}}$ the associated measurement noise variance.

The strategy of finding a minimum is to use Newton-Raphson iteration starting from $\hat{\mathbf{x}}_0 = \bar{\mathbf{x}}$. At the beginning of the *u*th iteration step, there is already an

estimate $\hat{\mathbf{x}}_{u-1}$ obtained from the previous step. Now the set of equations (3-8) is better approximated by

$$y_k^{(n)} \approx \underbrace{\sqrt{\frac{1}{M}} K_k^{(n)}\Big|_{\hat{\mathbf{x}}_{u-1}^T,\hat{\mathbf{s}}_k^{pri}} \left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T \mathbf{x}}_{\Omega(\mathbf{x})} + \underbrace{K_k^{(n)} n_k^{(n)} + d_k^{(n)}}_{\tilde{w}_k} \tag{3-44}$$

with noise variance

$$\begin{aligned} \mathbf{C}_{\tilde{w}} &= \left|K_k^{(n)}|_{\hat{\mathbf{x}}_{u-1}^T,\hat{\mathbf{s}}_k^{pri}}\right|^2 \sigma_n^2 + \sigma_d^2\Big|_{\hat{\mathbf{x}}_{u-1}^T,\hat{\mathbf{s}}_k^{pri}} \\ &\approx \left|A\left(\left|\sqrt{\frac{1}{M}}\hat{\mathbf{x}}_{u-1}^T \mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right|\right)\right|^2 + \left|K_k^{(n)}|_{\hat{\mathbf{x}}_{u-1}^T,\hat{\mathbf{s}}_k^{pri}}\right|^2 \left(\sigma_n^2 - \left|\sqrt{\frac{1}{M}}\hat{\mathbf{x}}_{u-1}^T \mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right|^2\right) \end{aligned} \tag{3-45}$$

At this point, $\Phi(\mathbf{x})$ is expanded around $\hat{\mathbf{x}}_{u-1}$ to a second order Taylor series approximation

$$\begin{aligned} \Phi(\mathbf{x}) \cong \Phi(\hat{\mathbf{x}}_{u-1}) &+ \nabla\Phi\Big|_{\hat{\mathbf{x}}_{u-1}} (\mathbf{x} - \hat{\mathbf{x}}_{u-1}) \\ &+ (\mathbf{x} - \hat{\mathbf{x}}_{u-1})^H \nabla^2\Phi\Big|_{\hat{\mathbf{x}}_{u-1}} (\mathbf{x} - \hat{\mathbf{x}}_{u-1}) \end{aligned} \tag{3-46}$$

where $\nabla\Phi = \dfrac{\partial\Phi}{\partial\mathbf{x}}$ and $\nabla^2\Phi = \dfrac{\partial(\nabla\Phi)}{\partial\mathbf{x}^*}$ denote the Gradient and Hessian matrix of $\Phi(\mathbf{x})$ with respect to complex vector $\mathbf{x}$, respectively.

The estimate $\hat{\mathbf{x}}_u$ is the minimum of the approximation of (3-46). It is found by equating the gradient of the approximation to zero. Differentiation of (3-46) with respect to complex vector $\mathbf{x}$ gives

$$\nabla\Phi\Big|_{\hat{\mathbf{x}}_{u-1}} + (\mathbf{x} - \hat{\mathbf{x}}_{u-1})^H \nabla^2\Phi\Big|_{\hat{\mathbf{x}}_{u-1}} = 0 \tag{3-47}$$

which results in

$$\hat{\mathbf{x}}_u = \hat{\mathbf{x}}_{u-1} - \left(\nabla^2\Phi\big|_{\hat{\mathbf{x}}_{u-1}}\right)^{-H}\left(\nabla\Phi\big|_{\hat{\mathbf{x}}_{u-1}}\right)^H \tag{3-48}$$

The Gradient and Hessian of $\Phi(\mathbf{x})$, in explicit form, are obtained from (3-42) as

$$\begin{aligned}\nabla\Phi = (\mathbf{x}-\bar{\mathbf{x}})^H\mathbf{C}_{\bar{\mathbf{x}}}^{-1} - \left(y_k^{(n)} - \Omega(\mathbf{x})\right)^H\mathbf{C}_{\tilde{w}}^{-1}\nabla\Omega = (\mathbf{x}-\bar{\mathbf{x}})^H\mathbf{C}_{\bar{\mathbf{x}}}^{-1} \\ -\left(y_k^{(n)} - \sqrt{\frac{1}{M}}K_k^{(n)}\Big|_{\hat{\mathbf{x}}_{u-1}^T,\hat{\mathbf{s}}_k^{pri}}\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T\mathbf{x}\right)^H\mathbf{C}_{\tilde{w}}^{-1}\sqrt{\frac{1}{M}}K_k^{(n)}\Big|_{\hat{\mathbf{x}}_{u-1}^T,\hat{\mathbf{s}}_k^{pri}}\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T\end{aligned} \tag{3-49}$$

$$\nabla^2\Phi = \mathbf{C}_{\bar{\mathbf{x}}}^{-1} + \frac{1}{M}\left|K_k^{(n)}\,|_{\hat{\mathbf{x}}_{u-1}^T,\hat{\mathbf{s}}_k^{pri}}\right|^2\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^*\mathbf{C}_{\tilde{w}}^{-1}\left(\mathbf{f}_{\mathrm{T}}(\hat{\mathbf{s}}_k^{pri})\right)^T \tag{3-50}$$

Substitution of (3-49)-(3-50) into (3-48) yields the following iteration scheme

$$\begin{aligned}\hat{\mathbf{x}}_u = \hat{\mathbf{x}}_{u-1} - \left(\mathbf{C}_{\bar{\mathbf{x}}}^{-1} + \left(\nabla\Omega\,|_{\hat{\mathbf{x}}_{u-1}}\right)^H\mathbf{C}_{\tilde{w}}^{-1}\nabla\Omega\,|_{\hat{\mathbf{x}}_{u-1}}\right)^{-1} \\ \times\left[\mathbf{C}_{\bar{\mathbf{x}}}^{-1}(\hat{\mathbf{x}}_{u-1}-\bar{\mathbf{x}}) - \left(\nabla\Omega\,|_{\hat{\mathbf{x}}_{u-1}}\right)^H\mathbf{C}_{\tilde{w}}^{-1}\left(y_k^{(n)} - \Omega(\hat{\mathbf{x}}_{u-1})\right)\right]\end{aligned} \tag{3-51}$$

We note that, $\hat{\mathbf{x}}_u$ converges very fast and it is possible to fix the number of iterations to some small number U, e.g. $U=3$. (The effect of U on system error performance is analyzed through simulation below). The final result is set to the last iteration, i,e, $\hat{\mathbf{h}}_k^{(n)} = \hat{\mathbf{x}}^T = \hat{\mathbf{x}}_U^T$.

The factor $\left(\mathbf{C}_{\bar{\mathbf{x}}}^{-1} + \left(\nabla\Omega\,|_{\hat{\mathbf{x}}_{u-1}}\right)^H\mathbf{C}_{\tilde{w}}^{-1}\nabla\Omega\,|_{\hat{\mathbf{x}}_{u-1}}\right)^{-1}$ in (3-51) can be regarded as the error covariance matrix associated with $\hat{\mathbf{x}}_u$, i.e.

$$\mathbf{C}_{\hat{\mathbf{x}}_u}^{(n)} = \left(\mathbf{C}_{\bar{\mathbf{x}}}^{-1} + \left(\nabla\Omega\,|_{\hat{\mathbf{x}}_{u-1}}\right)^H\mathbf{C}_{\tilde{w}}^{-1}\nabla\Omega\,|_{\hat{\mathbf{x}}_{u-1}}\right)^{-1} \tag{3-52}$$

and $\sigma^2_{\hat{h}_{n,m}(k)} = \left[\mathbf{C}_{\hat{\mathbf{x}}_U}^{(n)}\right]_{m,m}$.

This insight gives another connection to the last term in (3-51) because, in comparison with the standard Kalman filter recursion formulas, the term $\mathbf{C}_{\hat{\mathbf{x}}_u}^{(n)}\left(\nabla\Omega|_{\hat{\mathbf{x}}_{u-1}}\right)^H\mathbf{C}_{\tilde{w}}^{-1}$ can be regarded as the Kalman gain matrix during the *u*th iteration.

Again, each of the channel vectors $\hat{\mathbf{h}}_k^{(n)};n=1,...,N$ has to be obtained individually from the locally iterative algorithm described above.

3.2.3 Data Vector Recovery

Here, the receiver recovers the transmitted data vector using the channel matrix estimate $\hat{\mathbf{H}}_k=\left[\hat{\mathbf{h}}_k^{(1)}\quad\hat{\mathbf{h}}_k^{(2)}\quad\cdots\quad\hat{\mathbf{h}}_k^{(N)}\right]^T$ and the primary data vector estimate $\hat{\mathbf{s}}_k^{pri}$. The minimum distance receiver [3.45] chooses the vector $\hat{\mathbf{s}}_k$ that solves

$$\hat{\mathbf{s}}_k=\arg\min_{\mathbf{s}_k\in C^M}\|\mathbf{y}_k-\mathbf{o}_k\|_F^2 \tag{3-53}$$

where $\|.\|_F$ denotes the Frobenius norm and

$$\mathbf{o}_k=\left[f\left(\sqrt{\frac{1}{M}}\hat{\mathbf{h}}_k^{(1)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k)\right)\quad f\left(\sqrt{\frac{1}{M}}\hat{\mathbf{h}}_k^{(2)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k)\right)\quad\cdots\quad f\left(\sqrt{\frac{1}{M}}\hat{\mathbf{h}}_k^{(N)}\mathbf{f}_{\mathrm{T}}(\mathbf{s}_k)\right)\right]^T.$$

In (3-53) the search is performed over all the candidate vector symbols, $\mathbf{s}_k$, and the decoding complexity of the receiver is exponential in *M*. Note that in fact, (3-53) is the optimum ML decoder when the MIMO system is linear.

Finally, the estimated symbols are fed to the hard decision block to yield the detected transmitted symbols by setting the optimal thresholds in the constellation regions of the transmitted signal. The channel matrix estimate,

$\hat{\mathbf{H}}_k = \left[\hat{\mathbf{h}}_k^{(1)} \quad \hat{\mathbf{h}}_k^{(2)} \quad \cdots \quad \hat{\mathbf{h}}_k^{(N)}\right]^T$ and the corresponding variances associated with the estimation processes, $\sigma^2_{\hat{h}_{n,m}(k)}; n = 1,...,N; m = 1,...,M$, also help the decision algorithm in the subsequent (k+1)th symbol interval to obtain the subsequent primary data vector estimate $\hat{\mathbf{s}}_{k+1}^{pri}$. We note that for initialization, it is sufficient for the proposed algorithm to use any of the training approaches developed in [3.44] only once at the beginning of whole data sequence transmission. To obtain a meaningful estimate of $\mathbf{H}$ at the beginning, there is needed at least as many training measurements as unknowns, which implies that at the first Mth iteration, any kind of training-based scheme, e.g., LS or MMSE pilot-aided channel estimation, is used for an initial estimate of the channel, $\hat{\mathbf{H}}_{ini} = \hat{\mathbf{H}}_M$. From the ($M$+1)th iteration, the algorithm switches to the presented iterative algorithm and uses the channel estimate $\hat{\mathbf{H}}_M$ that is at hand. This means that the algorithm switches to a "blind" approach (in the sense that no further pilots are required) after at least the Mth iteration. The initial channel estimates obtained this way, are accurate enough for the algorithm to have fast convergence for the typical application of systems with fade rates of $f_D T_s \leq\sim 0.005$.

3.3 Simulation Results

The optimal receiver performance sets a lower bound on the error rate probability of sub-optimal receivers. However, no exact optimal analytical solution is available for error probability, even for linear MIMO systems. (For linear systems several upper bounds on error probability for the ML as well as other

sub-optimal receivers have been derived.) Thus, simulation for the SER performance results of coherent detection with perfect CSI is used for a benchmark, and evaluated and compared against the results of the presented decision-directed system with the associated perfect CSI reference ones.

We take i.i.d. QPSK data symbols with zero mean and variance 1, and for convenience of interpretation, the data symbol duration is $T_{\mathrm{s}} = 0.1\,\mathrm{ms}$ (i.e., 10k symbols/sec) and the channel fate rade is $f_{\mathrm{D}}T_{\mathrm{s}} = 0.001$. Also, throughout the simulations, the signal to noise ratio is

$$\mathrm{SNR(dB)} \overset{\Delta}{=} 10\log_{10}\frac{E\{\left|s_k\right|^2\}}{E\{\left|n_k\right|^2\}} = 10\log_{10}\frac{\sigma_s^2}{\sigma_n^2} \tag{3-54}$$

The proposed algorithm is formulated in a general form and it can be applied to any linear and nonlinear MIMO systems with nonlinearities introduced by $f_{\mathrm{T}}(.)$ and $f_{\mathrm{R}}(.)$ in (1). The last sub-section below, 3.3.2, is dedicated to linear MIMO systems since with this choice, more direct assessment of the proposed channel estimation technique on system error performance is possible. Also, this allows fair performance comparison with known results of the conventional decision-directed Kalman filtering [3.30] and two pilot-aided systems (least-squares (LS), and an MMSE pilot-aided system [3.11]).

Here, the value of $\rho_{\mathrm{cor}}(T_s) = J_0(2\pi f_{\mathrm{D}}T_s)$ is assumed to be known, i.e., estimation of these statistics is not included as part of the algorithm. Including this estimate is a relatively straightforward extension, but the goal here is to quantify the behaviour of the decision-directed estimation with known channel

statistics. The impact of any mismatch in the time-correlation coefficient mismatch (perhaps caused by an incorrect estimate of f_{D}, or simply because the model is incorrect) on the proposed system's error performance is analyzed by simulation below. We show that the proposed technique is robust for a mild mismatch of the channel's second-order statistics, i.e., the correlation function values which relate to f_{D}, relative to the a priori information assumed by the receiver. However, a large degradation in performance is demonstrated as the mismatch increases, as expected, and this is demonstrated below.

3.3.1 Nonlinear MIMO System

Despite nonlinearities being usually small, they are known to be very difficult to deal with [3.45]. Here, for both transmit and receive amplifiers some examples are provided from the family of Solid State Power Amplifier (SSPA) nonlinearity which is defined as [3.37]-[3.38]

$$F_A(\rho) = \frac{\rho}{\left[1+\left(\frac{\rho}{A_s}\right)^{2q}\right]^{\frac{1}{2q}}} \quad , \quad F_P(\rho) = 0 \tag{3-55}$$

where q defines the smoothness of the transition from linear operation to saturation and A_s is the saturation output amplitude where for simulation it is set as $A_s = 1$. q is a standard parameter which the designer must assign to the SSPA amplifier. This may require a calibration measurement. Designers would be able to decrease the cost of the amplifiers if a larger non-linearity could be tolerated. For large values of *q*, the SSPA model approaches the soft limiter

model that is commonly used to represent the clipping operation [3.46] to reduce the dynamic range of the OFDM signal. The soft limiter model is specified as

$$F_A(\rho) = \begin{cases} \rho & ; \ \rho \le A_s \\ A_s & ; \ \rho > A_s \end{cases} \quad , \qquad F_P(\rho) = 0 \, . \tag{3-56}$$

We note that the model used in (3-55) is just an example of amplifier nonlinearity provided here for simulation purposes. Any kind of memoryless AM/AM and AM/PM characteristics can be used instead. Furthermore, throughout the simulations, q_{T} and q_{R} are used as the corresponding transmit and receive amplifier smoothness parameter, respectively.

Figure 3-2 illustrates the SER performance of a 2×2 nonlinear MIMO system with different nonlinearity configurations for transmit and receive amplifiers when the statistical linearization technique is used. In general, as either of the parameters q_{T} or q_{R} increases, less degradation in performance is observed since the nonlinearity becomes smoother.

A useful result is that, for a given receiver nonlinearity represented by q_{R} parameter, the ensuring system performance curves are essentially shifted versions of the curves obtained with the ideal linear transmit amplifier ($q_{\mathrm{T}} \to \infty$). This is seen in Figure 3-3 where the effect of nonlinear parameter q_{T} (representing the nonlinearity at the transmitter) on the system SER performance with fixed $q_{\mathrm{R}} = 2$ is illustrated. As an example, there is a shift of 1.1 dB in SNR, except near the error floor region. As a result, in order to assess the impact of the receiver nonlinearity (represented by q_{R}) on system error performance more

directly, in the rest of the simulations, the SER performance is evaluated of nonlinear MIMO systems with ideal linear amplifiers at the transmitter by setting

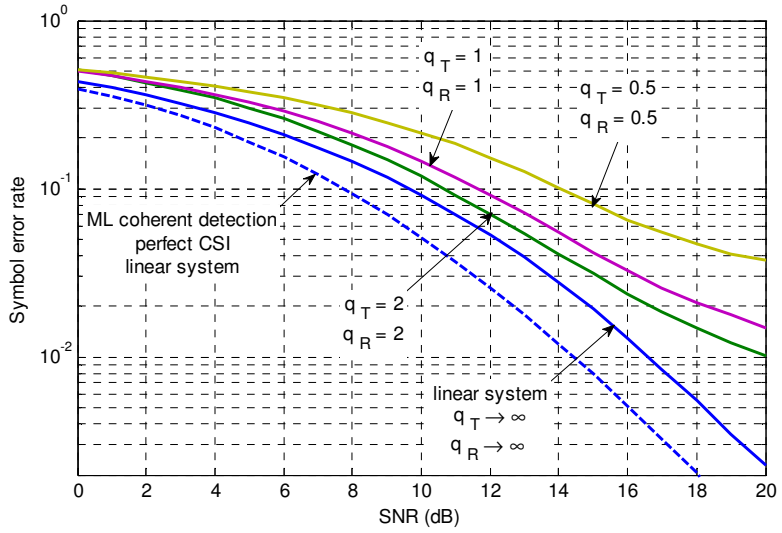

Figure 3-2 SER performance of a 2×2 nonlinear MIMO system with different nonlinearity configurations for transmit and receive amplifiers

The statistical linearization technique is used, and fade rate is 0.001

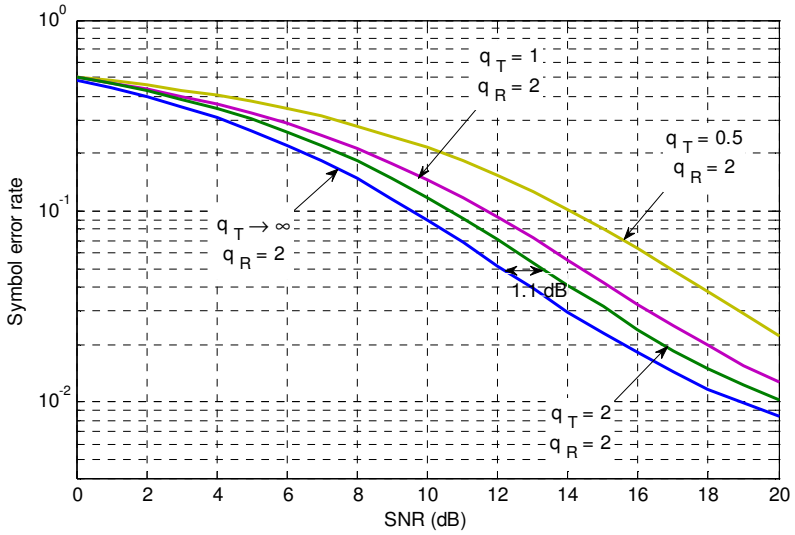

Figure 3-3 Effect of nonlinear parameter q_T (representing the nonlinearity at the transmitter) on the system SER performance with fixed $q_R = 2$

$q_T \to \infty$. Here we can also benchmark the two statistical/analytical linearization techniques with a look-up table (LUT) approach to the nonlinear problem.

3.3.1.1 Statistical/Analytical linearization approaches: $q_T \to \infty$

The impact of nonlinearity parameter q_R on the SER performance of a 2×2 nonlinear MIMO system using the statistical linearization technique is shown in Figure 3-4. Comparison is also made with the linear system error curve which corresponds to large values of q_R (soft limiter model). The nonlinearity is nearly compensated for values of SNR < 13 dB for $q_R \geq 0.5$. The performance curves with analytical linearization also follow the same trend as the curves with statistical linearization. The SNR value of 13 dB would change with different fading rates.

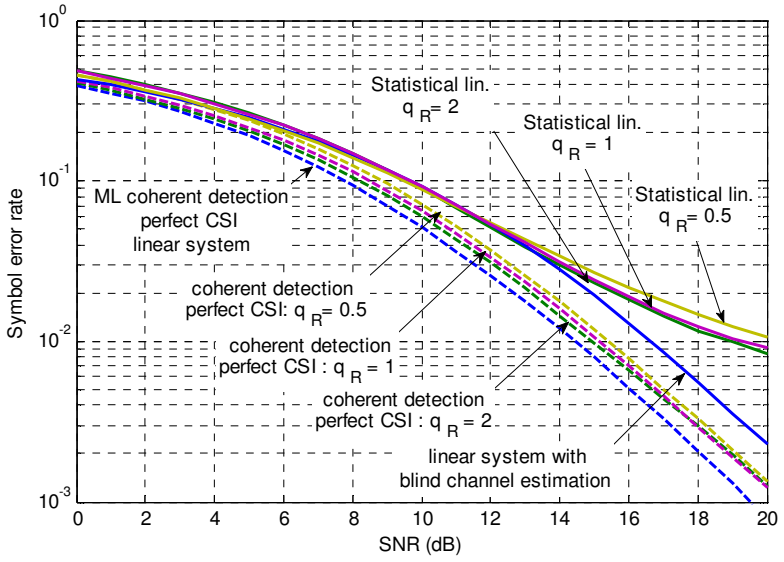

Figure 3-4 Impact of nonlinearity parameter q_R on SER performance of a 2×2 nonlinear MIMO system employing the statistical linearization technique while $q_T \to \infty$

At large values of SNR, the difference between the curves of the statistical/analytical linearization approaches and that for a linear system increases with SNR. The performance degradation associated with the statistical/analytical approaches can be explained via equation (3-8): it is found via simulations (not shown here) that the noise variance σ_d^2 (see (3-7)) increases with SNR. As a result, the Bussgangs' approximation in (3-8) becomes less accurate for larger values of SNR. In turn, this causes inaccuracies in the decision algorithm which is built up from (3-8).

Figure 3-5 shows the SER performance of a nonlinear SISO and a 2×2 nonlinear MIMO system with the nonlinearity parameters $q_{\mathrm{R}} = 0.5$ and perfect CSI at the receiver. The corresponding curves of the presented decision-directed

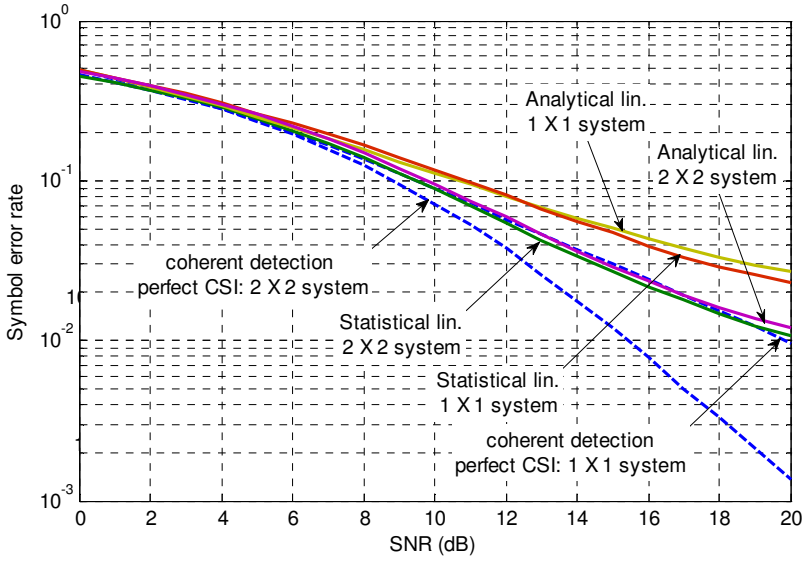

Figure 3-5 SER performance of a nonlinear SISO and a 2×2 nonlinear MIMO system with the nonlinearity parameter $q_{\mathrm{R}} = 0.5$ and $q_{\mathrm{T}} \to \infty$.

Dashed and solid lines correspond to coherent detection with perfect CSI and the proposed systems, respectively. The analytical linearization parameter is $U = 3$

system employing statistical and analytical linearization techniques are also illustrated. The analytical linearization technique finishes the locally iterative (nested) scheme in $U = 3$ steps.

The effect of U on system performance is studied below. The degradation in performance of a 2×2 nonlinear system is more than that in the nonlinear SISO system. This trend is expected since more nonlinear amplifiers contribute more nonlinearity as the number of receive antennas increases. Figure 3-5 reveals the details.

Finally, the impact of the analytical linearization parameter U on SER performance is shown in Figure 3-6. Typically, the performance gets better as the total number of iterations in the proposed local iterative scheme increases from 1 to 3.

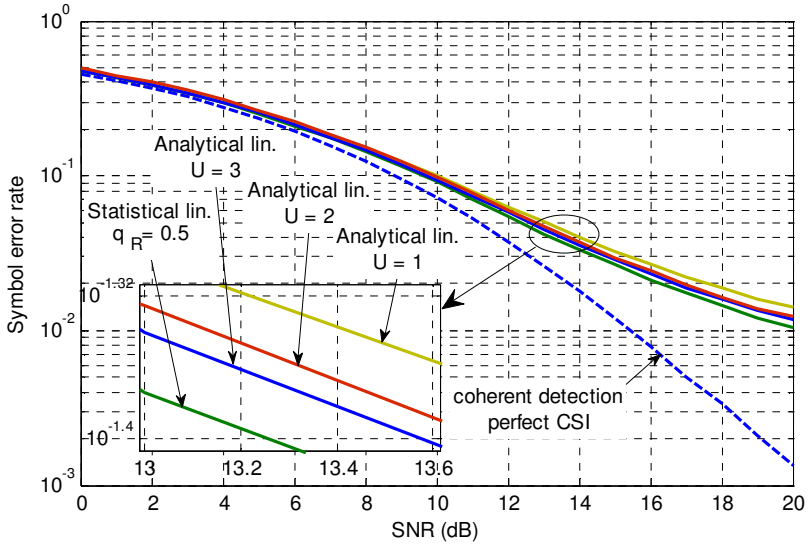

Figure 3-6 Impact of the analytical linearization parameter *U* on SER performance of a 2×2 nonlinear MIMO system for $q_{\mathrm{T}} \to \infty$

For $U \geq 3$, the local iterative scheme nearly converges to a steady state which is the locally optimum value and as a result no further improvement in performance is observed.

As in Section 2.4.2, and repeated here for convenience, it is assumed that the normalized time-correlation function of the channel coefficients is known, and only one value of the channel correlation coefficient function, $\rho_{\text{cor}}(T_s) = J_0(2\pi f_{\text{D}} T_s)$, is required. As already noted, a real-world, local autocorrelation is seldom exactly J_0 (or some other model), even in the main lobe region, and almost never J_0 (or another model) away from the main lobe. Furthermore, estimating a correlation coefficient from finite samples introduces uncertainty depending on the number of samples used. In order to get a reliable estimate where the function becomes small, a very large number of samples is required [3.47].

In order to test the performance sensitivity to the second-order statistics, the simple percentage correlation function sample mismatch, as in (2.35), is

$$\eta_{miss}(T_s) = \left| \frac{\rho_{\text{cor}}(T_s) - \hat{\rho}_{\text{cor}}(T_s)}{\rho_{\text{cor}}(T_s)} \times 100 \right| = \left| \frac{J_0(2\pi f_{\text{D}} T_s) - \hat{\rho}_{\text{cor}}(T_s)}{J_0(2\pi f_{\text{D}} T_s)} \times 100 \right| \tag{3-57}$$

where the normalized time-correlation function used at the receiver takes on different values from $\rho_{\text{cor}}(T_s) = J_0(2\pi f_{\text{D}} T_s)$. Figure 3-7 illustrates the effect of the mismatch parameter $\eta_{miss} = \eta_0$ on SER performance, given as the relative SER change, $\left(SER_{\eta_{miss}=0} - SER_{\eta_{miss}=\eta_0}\right) / SER_{\eta_{miss}=0}$, at SNR = 20 dB using the statistical linearization. The system remains moderately insensitive to the mismatch of (3-57) - up to 20% for fate rate 0.005, but for higher mismatch, there will be a

larger degradation in the SER performance. Also, there is more sensitivity to the mismatch for higher fade rates. The reason is that, for a given mismatch, the higher fade rate results in the addition of an error term to equation (3-19) which in turn causes more signal estimation inaccuracies in the decision algorithm. Thus, for the typical applications with fade rate $f_{\mathrm{D}}T_s \leq\sim 0.005$, the proposed technique is robust to mild mismatches of the channel's second-order statistics. The figure helps to quantify the sensitivity of the proposed receiver to the modelling of the temporal second order statistics of the channel.

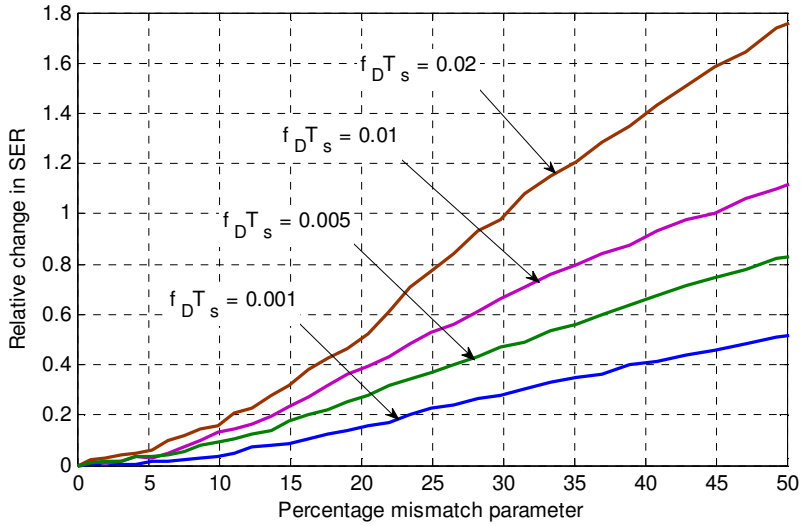

Figure 3-7 The effect of mismatch parameter on the relative SER performance of system employing the presented decision-directed channel estimation.
The statistical linearization is used at SNR = 20 dB.

3.3.1.2 Look-up table approach followed by linear channel estimation: $q_{\mathrm{T}} \to \infty$

We may also attempt to directly compensate the nonlinearity at the receiver before proceeding to channel estimation and data recovery by means of an LUT approach. First, by processing $y_n \equiv y_k^{(n)}; n = 1,...,N$ as shown in Figure 3-8,

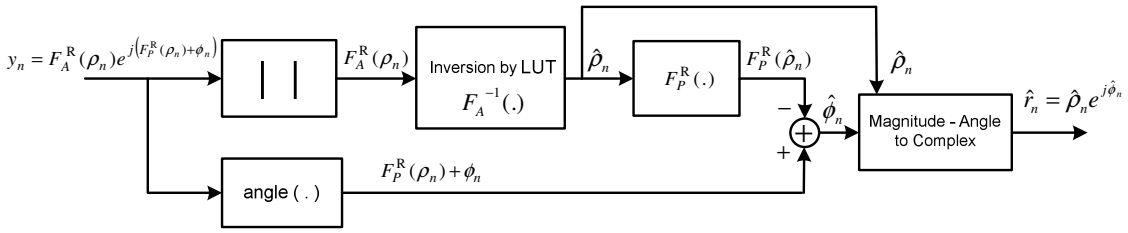

Figure 3-8 Receiver nonlinearity compensation by means of LUT, here for the *n*th antenna branch and $q_{\mathrm{T}} \to \infty$.

The outputs of all branches, $r_1, r_2, ..., r_N$, are fed to the multiple antenna decoder block for channel estimation and data recovery, designed for linear MIMO systems.

the estimation values of $\hat{r}_n \equiv \hat{r}_k^{(n)}; n = 1,...,N$ are obtained. The estimation process at this step, can be modelled as

$$\hat{\mathbf{r}}_k = \mathbf{r}_k + \sigma_e \tilde{\mathbf{w}}_k = \mathbf{r}_k + \mathbf{e}_k \tag{3-58}$$

where $\tilde{\mathbf{w}}_k$ is a vector with i.i.d. complex Gaussian entries each with variance 1 and σ_e^2 is the associated estimation error variance reflecting the inaccuracies in estimating $\hat{\mathbf{r}}_k$.

Next, these estimates are fed to the multiple antenna decoder block, designed for linear MIMO systems where decision-directed channel estimation and data recovery is carried out. From the SNR point of view,

$$SNR_{\hat{\mathbf{r}}} = SNR_{\mathbf{r}} - 10\log_{10}\sigma_e^2 \tag{3-59}$$

Equation (3-59) states that the proposed channel estimation/data recovery designed for the linear system is going to operate on the signal (i.e., $\hat{\mathbf{r}}_k$) that

has the reduced SNR value of $10\log_{10}\sigma_e^2$ in comparison to $\mathbf{r}_k$. Simulations show that σ_e^2 depends mainly on the gridding size (quantization for LUT) and does not change with SNR. As a result, for a fixed gridding size, the curves of LUT-based system are essentially shifted versions of the curve for a linear system. We note that the inversion operation $F_A^{-1}(.)$ shown in Figure 3-8 is performed by means of LUT.

The SER performance curves of system based on LUT compensation are shown in Figure 3-9 for the nonlinearity parameters $q_{\mathrm{R}} = 0.5$. The effect of different values of gridding sizes (quantization for the LUT) is also illustrated. Here, the input signal full scale amplitude is taken as 0 to 4, so a grid size of 0.005 means there are 800 LUT entries.

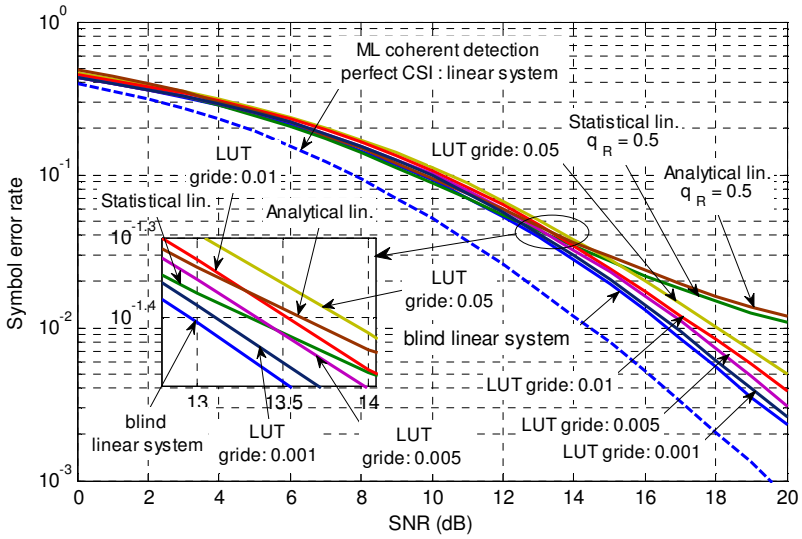

Figure 3-9 SER performance curves of a 2×2 system with nonlinearity compensation using LUT for $q_{\mathrm{R}} = 0.5$ and $q_{\mathrm{T}} \to \infty$

The effect of gridding size on the error performance is also illustrated and comparison is made with the presented statistical/analytical linearization approaches of section III.

The curves are essentially shifted versions of the curve for a linear system which confirm equation (3-59). However, as the gridding size increases (less LUT entries), a larger shift is observed.

The comparison can also be made with the SER results of the techniques proposed in section 6.2 using statistical/analytical linearization and Bussgang's theorem. The proposed statistical/analytical linearization approaches outperform the system based on LUT for SNR values smaller than about 13.5 dB. Since the error performance of statistical/analytical linearization approaches coincide with the curves of linear system for low values of SNR (where the nonlinearity is completely compensated), it is concluded that the main reason that the performance of LUT-based approach falls behind the statistical/analytical approaches, is the gridding size. In particular, for very small values of gridding sizes, $\sigma_e^2 \to 0$, the resultant LUT-based curve coincides with the corresponding curve for the linear system (see (3-59) and the LUT-based approach and the statistical/analytical approaches yield the same performance for small values of SNR (SNR < 13.5 dB for the example provided). On the other hand, at large values of SNR, the difference between the curves of the statistical/analytical approaches and that for a linear system, increases with SNR while the LUT-based curves experience a constant degradation in all SNRs. As a result, for large values of SNR the LUT-based approach outperforms the presented statistical/analytical linearization approaches.

Comparison is also made between the performance of the presented algorithm and that of LUT-based approach in a 2×2 system with nonlinear

amplifiers at both transmitter and receiver (both q_{T} and q_{R} are small). The results are illustrated in Figure 3-10 for $q_{\mathrm{R}}=1$ and various values of the transmit nonlinearity parameter, q_{T}. The LUT has a grid size of 0.005. For small values of q_{T}, the presented algorithm outperforms the LUT-based approach in SER performance for all values of SNR. In particular, the difference between SER performances becomes larger as q_{T} decreases. This is simply because the presented LUT-based approach is not designed to compensate the transmit nonlinearity. The result in Figure 3-10 lays out that the presented algorithm is more effective than the LUT-based approach in MIMO systems with nonlinearities at both transmit and receiver amplifiers.

We note that all presented techniques require knowledge of the nonlinearity characteristics, $F_A^{\mathrm{T/R}}(.)$ and $F_P^{\mathrm{T/R}}(.)$. However, the systems based

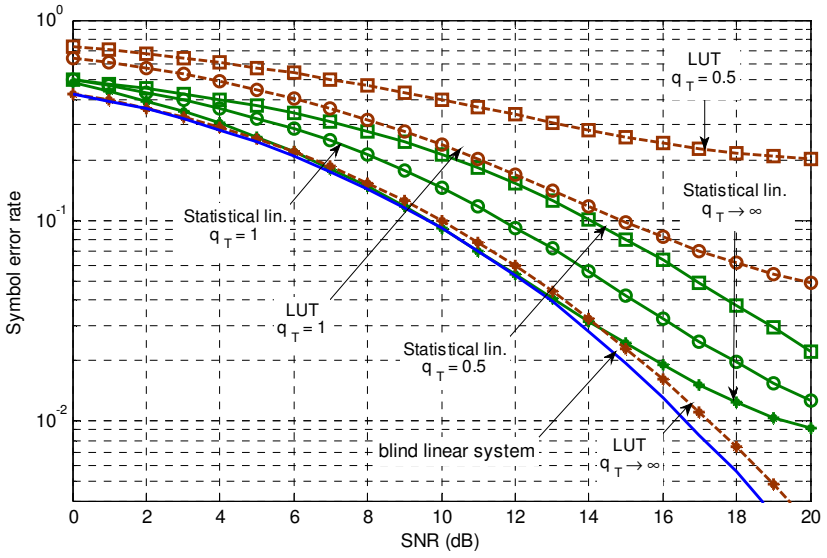

Figure 3-10 Comparison of SER performance of the presented algorithm with that of LUT-based approach in a 2×2 system with nonlinear transmit and receive amplifiers ($q_{\mathrm{R}}=1$ and various values of q_{T})

on LUT require the extra memory of the table. The size of the extra memory increases, as the gridding size decreases.

In order to assess the impact of the proposed channel estimation technique on system error performance more directly, the next section consideres linear MIMO systems. This also allows fair performance comparison with a known decision-directed system employing Kalman filtering to track the channel [3.30] and two pilot-aided systems (least-squares (LS), and an MMSE pilot-aided system [3.11]).

3.3.2 Linear MIMO System: $q_{\mathrm{T}} \to \infty$, $q_{\mathrm{R}} \to \infty$

For the linear MIMO systems ($q_{\mathrm{T}} \to \infty$ and $q_{\mathrm{R}} \to \infty$), the signal model (3-1) is simplified to

$$\mathbf{y}_k = \sqrt{\frac{1}{M}}\mathbf{H}_k \mathbf{s}_k + \mathbf{n}_k \tag{3-60}$$

The proposed channel estimation can be joined with any decoding structure (e.g., ML, MMSE, ZF and MMSE-OSUC) at the receiver. Each decoding structure provides a specific order of diversity. The zero forcing (ZF) receiver provides $N - M + 1$ order diversity [3.48] (the same as MMSE and successive cancellation (SUC) receivers but with different SNR loss). As an alternative, one may take advantage of an ordered successive cancellation (OSUC) receiver [3.49] which may have more than $N - M + 1$ order diversity because of the ordering (selection) process [3.48], or use the optimal ML receiver which extracts N order diversity with the expense of high decoding complexity

(exponential in M). For different decoding structures, the SER performance of the system employing the proposed channel estimation is compared with the associated coherent detection curve with perfect CSI. The results are illustrated in Figure 3-11 for a 2×2 system. Each pair of curves bearing the same color (also labeled) corresponds to a specific decoding structure (the "dotted line" is associated with the "coherent detection with perfect CSI" and the "solid line with marker" is the performance when the effect of decision-directed channel estimation is added to that system).

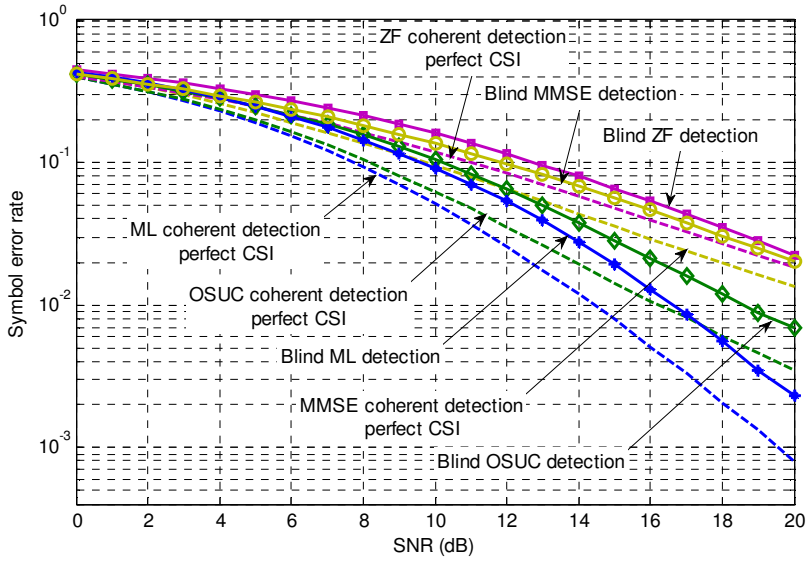

Figure 3-11 SER performance of a linear 2×2 MIMO system employing ML, MMSE, ZF, and MMSE-OSUC receivers.

Dashed and solid lines correspond to coherent detection with perfect CSI and proposed system, respectively; $q_{\mathrm{T}}\to\infty$, $q_{\mathrm{R}}\to\infty$.

The statistical and analytical linearization techniques yield the same performance in this case since no nonlinearity exists. With each receiver type (decoding structure), the presented system exploits the same diversity order as the coherent detection with perfect CSI but experiences an SNR loss. However,

the SNR loss depends on the system parameters and the type of decoding structure. In particular, it is within 3 dB (specifically, 2.5 dB) of the perfect CSI curves in an ML (in MMSE-OSUC) receiver at fade rate $f_{\mathrm{D}}T_{\mathrm{s}} = 0.001$ and SER of 10^{-2}.

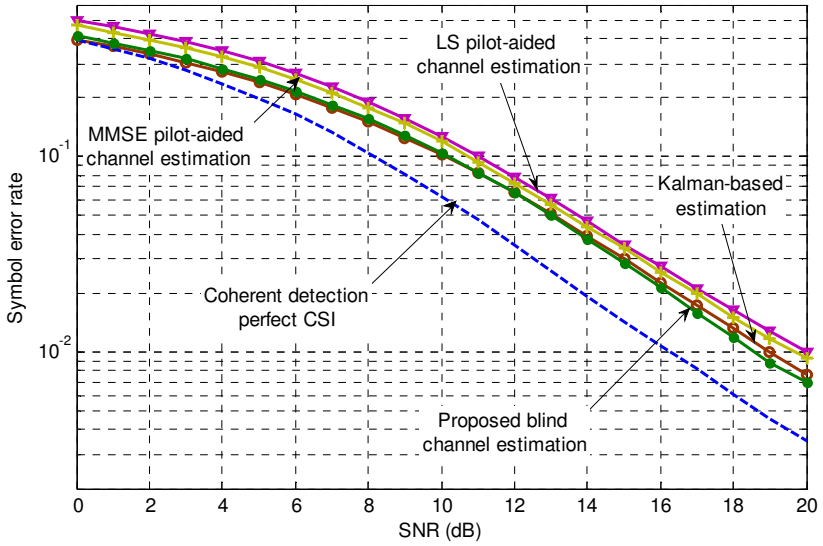

Figure 3-12 Comparison of SER performances of linear 2×2 MMSE-OSUC receivers employing different channel estimation techniques.

The system uses SM-HE, QPSK signaling at fade rate $f_{\mathrm{D}}T_{s} = 0.001$

We may also compare the results of a system employing the presented decision-directed channel estimation with that of a system using other channel estimation techniques for a given decoding structure. Here, a fair comparison is possible with a known channel estimation based on Kalman filtering (conventional decision-directed Kalman filtering [3.30]), and two known pilot-aided channel estimations: Least Squares (LS) and an MMSE pilot-aided system [3.11]. The results are shown in Figure 3-12 using the MMSE-OSUC receiver. The presented system outperforms both pilot-aided channel estimation

techniques but with just a slight improvement in performance over the decision-directed Kalman-based estimation at moderate-high values of SNR.

Recall that the technique of [3.30] is a decision-directed algorithm which uses channel predictions to obtain the coarse symbol estimate for formulating the measurement equations in the Kalman filtering. However, obtaining the coarse symbol estimate is undertaken without considering the associated channel estimation error variances, $\sigma^2_{\hat{h}_{n,m}(k-1)}$, whereas these have been taken into consideration in the proposed decision algorithm ((3-15)-(3-22)). Compared to pilot-aided systems, the proposed decision-directed system has superior performance for all SNRs, at least for the fade rate used here. This is expected, because in pilot-aided systems, the channel is estimated only after at least M consecutive symbols, whereas in the presented technique, the channel is estimated at each symbol time which results in more accurate channel estimates in the time-varying environment.

The robustness of the proposed system to the time-variations of the channel can also be quantified by simulation and compared with that of other conventional channel estimation techniques. The MMSE-OSUC decoding structure is used for all the systems. The effect of fade rate (maximum Doppler shift) on error performance is shown in Figure 3-13 for $10\,\mathrm{Hz} \le f_\mathrm{D} \le 300\,\mathrm{Hz}$ at SNR = 20 dB. For the proposed system, there is a maximum performance degradation of 15% in the SER (compared to 60% in pilot-aided systems) for a fading rate $f_\mathrm{D}T_s \le\sim 0.01$ or $f_\mathrm{D} <\sim 100\,\mathrm{Hz}$, relative to the case of block fading in a quasi-static

channel (the Kalman-based estimation has the same trend as with the proposed channel estimation). The pilot-aided systems are more sensitive to time-variations of the channel. However, in general, the system performances depend on many factors such as channel fade rate, SNR and pilot scheme used.

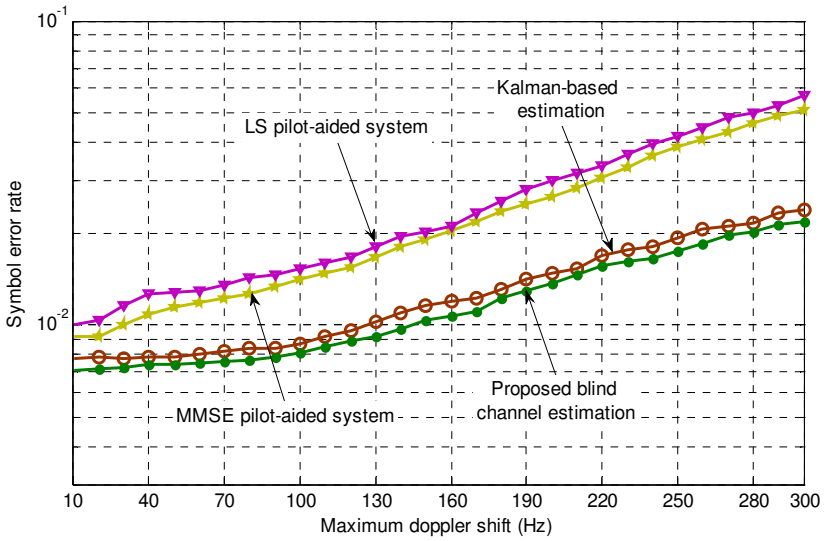

Figure 3-13 The effect of maximum Doppler shift on decision-directed/pilot-aided system error performance employing MMSE-OSUC receiver at SNR = 20 dB.

For example, the LS pilot-aided system is always worse than the MMSE pilot-aided system on this measure, and at very low fade rates (static fading) the all the algorithms (two-pilot aided systems, the Kalman-based estimation, and the proposed channel estimation technique) yield nearly the same performance at all values of SNR. However, as the fade rate increases, the deference between the SER curves increases.

Techniques that are based on the three conventional channel estimators under comparison (the two LS/MMSE pilot-aided channel estimator and decision-directed Kalman filtering) and/or detectors such as ZF, MMSE and OSUC,

become ineffective in nonlinear MIMO systems. This is because these channel estimators and detectors are designed based on the linearity in a system. As another example, it is well known that the Kalman filter is the optimum linear state estimator when the set of state and observation equations are linear [3.43], [3.30]. The performance of a Kalman filter degrades significantly even when a mild nonlinearity exists in the system. More severely, in most nonlinear cases the Kalman filter does not converge at all if an inappropriate initial state value is set within the algorithm [3.43]. As a result, a decision-directed Kalman-based channel estimator is not capable of tracking the time-variation in the channel when the system in nonlinear, and the system remains running without appropriate channel estimates, and the receiver fails to detect the transmitted symbols correctly.

This is not the case with the presented algorithm. As shown earlier, the presented algorithm is also effective for nonlinear MIMO systems. As a result, this can be considered as one of the advantages of the presented algorithm over the other conventional techniques (including detection and channel estimation) designed exclusively for linear systems.

3.4 Summary and Conclusions

The amplifiers in a practical MIMO system can be performance-limiting. Amplifier non-linearity is modelled here with a memoryless, AM/PM amplifier characterization. The theory and signal-processing is presented for compensating the link degradation caused by the non-linearity. This includes formulating the first proposed decision-directed approach in the form of two

alternative channel estimation/tracking methods (statistical and analytical linearization) for nonlinear MIMO systems. The MIMO performance is estimated by simulation. For a fixed nonlinearity in each receiver amplifier, the resultant system error curves are essentially shifted versions of the curves obtained with ideal linear transmit amplifiers. Furthermore, for the examples provided, the proposed approach is capable of nearly compensating the nonlinearity induced at the receiver side for values of SNR less than about 13 dB. That is, in this range of SNR, the error curves coincide with those of the linear MIMO system when the amplifiers are linear. The two channel estimation methods, statistical linearization and analytical linearization, have similar performance, being within a fraction of a dB in SNR. They both follow the same trend as coherent systems with perfect CSI for small- to mid-sized SNR, and develop an error floor for large SNR. The look-up table (LUT) approach also offers good performance, being more effective than the presented algorithm in compensating the receiver nonlinearity at high SNR when the transmit amplifier is linear. This is at the expense of extra memory. For smaller values of the SNR which are typical of wireless at extended ranges, the performance is similar, with the linearization outperforming the LUT approach by about a half dB. However, with nonlinearities at both transmit and receiver amplifiers, the presented algorithm is more effective than the LUT-based approach in MIMO systems at all values of SNR.

As the MIMO system becomes linear, typically possible by using expensive amplifiers, the performances of the statistical and analytical linearization approaches converge. The scheme presented here compares

favourably against known decision-directed schemes such as conventional Kalman filtering, and two established, pilot-aided systems. The proposed channel estimation is also more robust to the time-variations of the channel compared to LS/MMSE pilot-aided systems.

The comparison of simulation results with benchmark and other known results puts a focus on the importance of checking the channel modeling. We show that the proposed technique is robust for a moderate mismatch of the channel's second-order statistics, i.e., the correlation function values, relative to the *a priori* information assumed by the receiver. However, a large degradation in performance is demonstrated as the mismatch increases, as expected.

REFERENCE LIST

[3.1] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *J. Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[3.2] R. O’Neil and L. B. Lopes, “Performance of amplitude limited multitone signals,” *IEEE Veh. Technol. Conf.*, pp. 1675–1679, Jun. 1994.

[3.3] D. Dardari, V. Tralli, and A. Vaccari, “A Theoretical Characterization of Nonlinear Distortion Effects in OFDM Systems,” *IEEE Trans. Commun.,* vol. 48, no, 10, pp. 1755-1764, Oct. 2000.

[3.4] W. A. Morgan, “Minimize IM distortion in GaAs FET amplifiers,” *Microwaves and RF.,* vol. 25, no. 10, pp. 107-110, 1986.

[3.5] V.M. Vladimirov, S.N. Kulinich, Yu.Y.Shikhov, “LNA – Active bandpass filter for receiver-indicator of Glonass+GPS,” in *Proc. Int. Conf. Inf. Commun. Energy Sys. Technol.,* Oct. 2002.

[3.6] T. H. Lee, *The Design of CMOS Radio Frequency Integrated Circuits*, Cambridge, U.K.: Cambridge University Press, 1998.

[3.7] T. H. Lee, “5-GHz CMOS wireless LANs,” *Trans. Microwave Theory and Technique*, vol. 50, no. 1, pp. 268-280, Jan. 2002.

[3.8] M. Ibnkahla, et al*., "*High-Speed satellite mobile communications: Technologies and challenges,” *IEEE Proceedings*, vol. 92, no. 2, Feb. 2004.

[3.9] Q. Sun, D. C. Cox, H. C. Huang, and A. Lozano, “Estimation of continuous flat fading MIMO channels,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 549–553, Oct. 2002.

[3.10] X. Ma, G. B. Giannakis, and S. Ohno, “Optimal training for block transmission over doubly selective wireless fading channels,” *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[3.11] M. Biguesh, and A. B. Gershman, “Training-Based MIMO Channel Estimation: A Study of Estimator Tradeoffs and Optimal Training Signals,” *IEEE Trans. Signal Process.,* vol. 54, no. 3, pp. 884-893, Mar. 2006.

[3.12] B. Hassibi and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links?,” *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 2515–2528, Apr. 2003.

[3.13] C.-Y. Chi and C.-H. Chen, “Cumulant based inverse filter criteria for MIMO blind deconvolution: Properties, algorithms, and application to DS/CDMA systems in multipath,” *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1282-1299, July 2001.

[3.14] Z. Ding and T. Nguyen, “Stationary points of a kurtosis maximization algorithm for blind signal separation and antenna beamforming,” *IEEE Trans. Signal Process.,* vol. 48, no. 6, pp. 1587-1596, Jun. 2000.

[3.15] C. Y. Chi, C. Y. Chen, CH Chen, C. C. Feng “Batch processing algorithms for blind equalization using higher-order statistics,” *IEEE Signal Process. Magazine,* vol. 2, no. 1, pp. 25-49, Jan. 2003.

[3.16] J. Liang, and Z. Ding, “Blind MIMO system identification based on cumulant subspace decomposition,” *IEEE Trans. Signal Process.,* vol. 51, no. 6, pp. 1457-1468, Jun. 2003.

[3.17] C. Shin, R. W. Heath, and E. J. Powers, “Blind Channel Estimation for MIMO-OFDM Systems”, *IEEE Trans. Veh. Technol.,* vol. 56, no. 2, pp. 670-685, Mar 2007.

[3.18] B. Chen, A. P. Petropulu, “Frequency domain blind MIMO system identification based on second and higher order statistics,” *IEEE Trans. Signal process.,* vol. 49, no. 8, pp. 1677-1688, Aug. 2001.

[3.19] T. Acar, Y. Yu, and A. P. Petropulu, “Blind MIMO system estimation based on PARAFAC decomposition of higher order output tensors,” *IEEE Trans. Signal Process.,* vol. 54, no. 11, pp. 4156-4168, Nov. 2006.

[3.20] M. A. Khalighi, and S. Bourennane, “Semi-blind channel estimation based on superimposed pilots for single-carrier MIMO systems”, *IEEE Veh. Technol. Conf.* pp. 1480-1484, Apr 2007.

[3.21] J. Gao, and H. Liu, “Low-complexity MAP channel estimation for mobile MIMO-OFDM systems”, *IEEE Trans. Wireless Commun.,* vol. 7, no. 3, pp. 774-780, Mar 2008.

[3.22] Z. Ding, T. Ratnarajah, and C. F. N. Cowan, “HOS-based semi-blind spatial equalization for MIMO Rayleigh fading channels,” *IEEE Trans. Signal Process.,* vol. 56, no. 1, pp. 248-255, Jan. 2008.

[3.23] T. Dahl, N. Christophersen, and D. Gesbert, “Blind MIMO eigenmode transmission based on the algebraic power method,” *IEEE Trans. Signal Process.,* vol. 52, no. 9, pp. 2424-2431, Sep. 2004.

[3.24] A. K. Jagannatham, and B. D. Rao, “Whitening-rotation-based semi-blind MIMO channel estimation,” *IEEE Trans. Signal Process.,* vol. 54, no. 3, pp. 861-869, Mar. 2006.

[3.25] A. Medles and D. T. M. Slock, “Semiblind channel estimation for MIMO spatial multiplexing systems,” in *Proc. Veh. Technol. Conf.*, vol. 2, pp. 1240–1244, Oct. 2001.

[3.26] R. Everson and S. Roberts, *Independent Component Analysis*, *Principles and Practice*, Cambridge, U.K.: Cambridge University Press 2001.

[3.27] V. Zarzoso, and A. K. Nandi, “Blind MIMO equalization with optimum delay using independent component analysis,” *Int. J. Adapt. Control Signal Process.,* Vol. 18, No. 3, pp. 245-263, Mar. 2004.

[3.28] A. Mansour, “A mutually referenced blind multiuser separation of convolutive mixture algorithm,” *Signal Processing,* Vol. 81, No. 11, pp. 2253-2266, 2001.

[3.29] V. Zarzoso, and A. K. Nandi, “Exploiting non-Gaussianity in blind identification and equalization of MIMO FIR channels,” *IEE Proc. in Vision, Image and Signal Process.,* Vol. 151, No. 1, pp. 69-75, Feb. 2004.

[3.30] A. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, “Multi-input multi-output fading channel tracking and equalization using Kalman estimation”, *IEEE Trans. Signal Process.,* Vol. 50, No. 5, pp. 1065-1076, May, 2002.

[3.31] W. H. Chin, D. B. Ward, A.G. Constantinides “Semi-blind MIMO channel tracking using auxiliary particle filtering,” *Proc. GLOBECOM,* Vol. 1, pp. 322-325, Nov. 2002.

[3.32] S. Haykin, A. H. Sayed, J. R. Zeidler, P. Yee, and P. C. Wei. “Adaptive tracking of linear time-variant systems by extended RLS algorithms.” *IEEE Trans. Signal Process.,* Vol. 45, No. 5, pp. 11 18-1 128, May 1997.

[3.33] S. A. Banani, and Vaughan, R. G., "Compensating for non-linear amplifiers in MIMO communications systems", accepted in *IEEE Trans. Antennas and Propagation*, Jul. 2011.

[3.34] J. J. Bussgang, “Cross correlation function of amplitude-distorted Gaussian input signals,” *Res. Lab Electron., M.I.T., Cambridge, MA, Tech. Rep. 216,* Vol. 3, Mar. 1952.

[3.35] J. Minkoff, “The role of AM-to-PM conversion in memoryless nonlinear systems,” *IEEE Trans. Commun.*, Vol. 33, No. 2, pp. 139 – 144, Feb. 1985.

[3.36] S. A. Banani, and R. G. Vaughan, “Blind channel estimation for MIMO systems with nonlinearities at the receiver,” *IEEE Veh. Technol. Conf.* pp. 1-5, May 2010.

[3.37] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[3.38] A. M. Saleh, “Frequency independent and frequency dependent nonlinear model of TWT amplifiers,” *IEEE Trans. Commun.*, Vol. COM-29, pp. 1715–1720, Nov. 1981.

[3.39] H.Wang and P. Chang, “On verifying the first-order Markovian assumption for a Rayleigh fading channel model,” *IEEE Trans. Veh. Technol.*, Vol. 45, No. 2, pp. 353–357, May 1996.

[3.40] L. M. Davis, I. B. Collings, and R. J. Evans, “Coupled estimators for equalization of fast-fading mobile channels”, *IEEE Trans. Commun.,* Vol. 46, No. 10, pp. 1262 – 1265, Oct. 1998.

[3.41] S. Julier, J. Uhlmann, and H.F. Durrant-White, “A new method for nonlinear transformation of means and covariances in filters and estimators,” *IEEE Trans. Automatic Control*, Vol. 45, No. 3, pp.477-482, Mar. 2000.

[3.42] E. A. Wan and R. van der Merwe, “The unscented Kalman filter for nonlinear estimation,” *IEEE Symp. on Adaptive Systems for Signal Proc., Comm. And Control (AS-SPCC),* pp.153-158, 2000.

[3.43] A. H. Sayed, *Fundamentals of Adaptive Filtering,* Hoboken, NJ: Wiley, 2003.

[3.44] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, “Estimation and equalization of fading channels with random coefficients,” *Signal Process.*, Vol. 53, No. 2–3, pp. 211–229, Sep. 1996.

[3.45] J. G. Proakis, *Digital Communications,* 5th ed. New York: McGraw-Hill, 2008.

[3.46] R. Prasad, *OFDM for Wireless Communications Systems*, Norwood, MA: Artech House, 2004.

[3.47] R. Vaughan, and J. Bach Anderson, *Channels, Propagation and Antennas for Mobile and Personal Communications*, London, U.K.: IEE, 2003.

[3.48] A. Paulraj, R. Nabar and D. Gore, *Introduction to space-time wireless communications*, Cambridge, U.K.: Cambridge University Press, 2003.

[3.49] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, “Simplified processing for high spectral efficiency wireless communication employing multi-element arrays,” *IEEE JSAC*, Vol. 17, No. 11, pp. 1841-1852, Nov. 1999.

4: JOINT DECISION-DIRECTED CHANNEL ESTIMATION AND DISCRETE SPEED TRACKING IN WIRELESS SYSTEMS

The communications performance of many time-varying wireless systems depends on the accuracy of the estimation of the unknown parameters, such as the fading channel coefficients, by the mobile station (MS). The channel coefficients must be estimated by either pilot symbols (training sequences) or in a decision-directed manner. In all channel estimation techniques, one significant source of channel estimate inaccuracy is the Doppler phenomenon. The maximum Doppler frequency (Doppler shift) is proportional to the relative speed of transmit and receive terminals. Most channel estimation techniques are designed and implemented for a fixed maximum Doppler frequency which corresponds to a mobile terminal moving with a constant speed (within fixed scatterers). However, if the speed or direction of the moving terminal changes, the parameters of the channel estimation algorithm have to be reset accordingly in offline mode.

A mobile trajectory can be divided into segments, each corresponding to a different behavioral mode of movement. That is, the mobile terminal may stand still or change its speed, etc. In such cases, a single motion model cannot characterize the mobile dynamics of the terminal. The motion can be modeled by one of a pre-specified, finite number of modes, each of which represents a different mobile dynamic. Here, the terminal movement is approximated by D

different constant velocities (or speeds, since the scatterers are considered as fixed direction). Simulations here have the actual speeds of the mobile terminal quantized to these values.

In this chapter [4.1] a new ICA-based decision-directed channel estimation technique (second proposed decision-directed approach in the thesis) is presented for spatial multiplexing MIMO schemes where the relative speed of the transmit/receive terminals may change. This changing speed is called maneuvering. This extends the work in [4.2] where the system with non-maneuvering terminals only is considered. We also present a new hard decision switching algorithm to handle the system under maneuvering scenario. It is shown by simulation that if the switching block is disabled in the algorithm within the maneuvering scenario, a large degradation in performance is observed. This further demonstrates the advantage of the switching block. In addition, as the switching algorithm selects between different speed modes of the maneuvering terminals at the symbol rate, the speed can also be tracked via the data communication link. This is accomplished based on only the received data information signal, i.e., no other information (for example from a speed sensor or radar, etc.) is required. This work is new in the sense that it is the first to address tracking of the mobile terminal in combination with communications reception.

The formulation is for general MIMO, and the specific examples considered are single-input single-output (SISO), maximum ratio combining (MRC) and a 2 by 2 MIMO. The channels are assumed to be time-selective Rayleigh flat fading. Non-stationary ICA with a generalized exponential density

function is used to separate each source signal and particle filtering is used to track the channel. Performance is evaluated by simulation and is compared with optimal coherent detection as benchmark. Moreover, under a non-maneuvering scenario (the mobile is moving with only one constant speed), a fair comparison with other channel estimation techniques (a known decision-directed technique based on Kalman-filter tracking [4.3] and two known pilot-aided systems [4.4]) is also possible. Here, improved performance is observed. Clearly, the presented system requires more computational complexity. Complete evaluation of the computational complexity is difficult; however; some insight is possible into the computational complexity of one iteration (symbol time), through an approximate operations count of the implementation equations.

The rest of the chapter is organized as follows. Section 4.1 describes the system model. The ICA based decision-directed estimation is formulated in section 4.2 followed by complexity evaluation in section 4.3. Simulations results are presented in section 4.4, and section 4.5 is the summary and conclusions.

4.1 System Model and Formulation

Throughout, an uncoded spatial multiplexing (SM) scheme is assumed, i.e. the transmitter demultiplexes the uncoded data stream to the M antennas at each symbol time with symbol time duration T_s. This is a common and reasonable assumption. Assuming a Rayleigh flat fading MIMO channel, assuming perfect synchronization, the signal model is given by

$$\mathbf{y}_k = \sqrt{\frac{1}{M}}\mathbf{H}_k\mathbf{s}_k + \mathbf{n}_k = \sqrt{\frac{1}{M}}\begin{bmatrix} h_{1,1}(k) & \cdots & h_{1,M}(k) \\ \vdots & \ddots & \vdots \\ h_{N,1}(k) & & h_{N,M}(k) \end{bmatrix}\mathbf{s}_k + \mathbf{n}_k \qquad (4\text{-}1)$$

where $\mathbf{y}_k$ is the received $N \times 1$ vector, $\mathbf{s}_k$ is the $M \times 1$ transmit signal vector with i.i.d. data symbol entries, and $\mathbf{n}_k$ is the zero mean white Gaussian noise with covariance matrix $\mathbf{R} = \sigma_n^2 \mathbf{I}_N$ at *k*th symbol time. $\mathbf{H}_k$ is the channel $N \times M$ matrix with i.i.d. complex Gaussian entries, $h_{i,j}; i = 1,...,N; j = 1,...,M$, with associated variances $\sigma_{h_{i,j}}^2 = 1$. The factor $1/\sqrt{M}$ accounts for the even division of the available energy between the transmit antennas. This choice is suboptimal in an information theoretic sense, but is used since there is no channel state information (CSI) at the transmitter. The optimum eigen-channel MIMO requires extensive channel sounding and its efficient deployment for fast-fading channels is not yet developed. The index *k* shows that, at each iteration (symbol time), the channel matrix is changing.

The terminal movement is approximated by D different dynamic modes described by known D different constant speeds, $\{\mathrm{v}_1, \mathrm{v}_2, ..., \mathrm{v}_D\}$, where each corresponds to a specific maximum Doppler frequency, $f_{\mathrm{m}}^{\,d}$, via the relation

$$f_{\mathrm{m}}^{\,d} = \frac{\mathrm{v_d}}{\lambda_{\mathrm{c}}} \ \mathrm{m/s} \quad ; \ d = 1,...,\mathrm{D} \qquad (4\text{-}2)$$

with λ_{c} the nominal wavelength. The same formula is used in Doppler radar [4.5]-[4.6] to measure the speed of a detected object using the Doppler shift of a target, and similarly for sonar.

We are not trying to estimate different values of fade rates here. The goal of the work is to present a new switching scheme to adaptively select between different speed modes. This makes it possible to track the discretized relative speed of transmit/receive terminals.

It is also assumed that the time-varying channel matrix coefficients, during the *d*th speed mode, each have the same form of normalized autocorrelation function,

$$\rho_{h_{i,j}}(\Delta t) = J_0(2\pi f_{\mathrm{m}}^{\,d} \Delta t) \tag{4-3}$$

with J_0 is the zero order Bessel function of the first kind. This model implies a specific physical scenario (the product of the distribution of incoming waves and the antenna pattern is uniform and omnidirectional), but other forms could be used in more advanced algorithms. The results in [4.7] show that the first-order autoregressive (AR) model provides a sufficiently accurate model for time-selective fading channels. Therefore, from (4-3) at $\Delta t = T_s$, the relation between $h_{i,j}(k)$ and $h_{i,j}(k-1)$ for the *d*th speed mode, can be approximated by

$$h_{i,j}(k) \approx J_0(2\pi f_{\mathrm{m}}^{\,d} T_s) h_{i,j}(k-1) + \sqrt{1 - \left| J_0(2\pi f_{\mathrm{m}}^{\,d} T_s) \right|^2}\, v_{i,j}(k); \quad d = 1,...,\mathrm{D} \tag{4-4}$$

where $v_{i,j}(k)$ is the AR white Gaussian noise process with unity variance.

The model in (4-1) suits the standard noisy ICA problem where there are *M* independent source components and *N* observations [4.8]. Maximum likelihood estimation, perhaps the most commonly used statistical estimation

principle, can be used to estimate the ICA model. It is closely related to the infomax principle described widely in the neural network literature. If the densities of the independent components are known in advance, a very simple gradient algorithm can be derived. To speed up convergence, the natural gradient version, and especially the Fast-ICA fixed-point algorithm, can be used. If the densities of the independent components are not known, the situation is somewhat more complicated. Fortunately, a very rough density approximation is seems to work, and in an extreme approximation case, just two types of density models (sub- and supergaussian) are sufficient to approximate the densities of the independent components. The choice of density can then be based on whether the independent components are sub- or supergaussian. This estimate (choice) can be simply added to the gradient methods, and this is automatically done in Fast-ICA algorithms.

Since multiple sources are involved in this problem and it is difficult to use a switching model for separating the sources with light-tailed probability density functions [4.9], a model is needed which is more flexible than the traditional $1/\cosh$ model [4.10]. We use the generalized exponentials model [4.8] as the source model. This model, like the mixture Gaussian model [4.11], can be used to model non-Gaussian sources for multi-source tracking, but it is not as computationally complex as a mixture Gaussian model in target tracking applications [4.8]. Here, the ICA approximates each source density by

$$p_{s_i}(s_i|\theta_i) = b_i \exp\left(-\left|\frac{s_i - m_i}{\omega_i}\right|^{r_i}\right) \quad ; \; i = 1,2,...,M \tag{4-5}$$

where

$$b_i = \frac{r_i}{2\omega_i \Gamma(1/r_i)} \tag{4-6}$$

is the normalizing constant with $\Gamma(x)$ denoting the Gamma function. The density depends upon parameters $\theta_i = \{m_i, \omega_i, r_i\}$. The location of the distribution is set by m_i, its width by ω_i, and its tails by r_i. Densities with $r < 2$ are called super-Gaussian, while those with $r > 2$ are called sub-Gaussian. With $r = 1$, $p_{s_i}(s_i|\theta_i)$ is Laplacian; and as $r \to \infty$, $p_{s_i}(s_i|\theta_i)$ becomes uniform. Generalized exponential source models in static ICA can separate mixtures of Laplacian, Gaussian and uniformly distributed sources, while methods using a static $\tanh$ nonlinearity cannot separate such mixtures [4.10]. If there is *a priori* knowledge of the sources (data symbols) available, then the distribution parameters can be selected accordingly. For example, if the receiver knows that the modulation is D-QAM, then the density in (4-5) can be set to equi-likely delta functions, i.e., the probability mass function (pmf) of the modulation, and the value of the pmf is (discrete-) uniform and 1/D. Section 4.4, below, also notes the usage of the pmf in the simulation. Using the ICA allows channel estimation even without knowledge of the modulation type at the receiver. If there is no such knowledge, then the ICA must check its stability condition (to ensure that the distribution is suitable) at each iteration [4.8].

Here, as in other ICA implementations, the scaling ambiguity between variances of the sources and the scaling of columns of $\mathbf{H}_k$ is resolved by

normalizing the variance of each source [4.8], i.e. setting $\omega_i = 1$ for all *i* and allowing all the scaled information to reside in the columns of $\mathbf{H}_k$. Furthermore, the mean of each source density can be set to zero, using *a priori* knowledge, and this reduces complexity.

4.2 Decision-directed Channel Estimation / Tracking

The block diagram of the presented decision-directed technique applied to a SM-MIMO system is illustrated in Figure 4-1. At each iteration (symbol time), first, the dominant speed mode, say mode *L*, is detected by the hard decision switching block described below in subsection 4.2.3. Having selected speed mode *L*, the corresponding set of equations (4-4) associated with the speed mode *L* is taken as the state (dynamic) equations and the ICA model (4-1) as the measurement equations. These set of equations make a state-estimation problem with the channel matrix coefficients considered as the state variables. Here, a particle filter combined with ICA is used to track the channel.

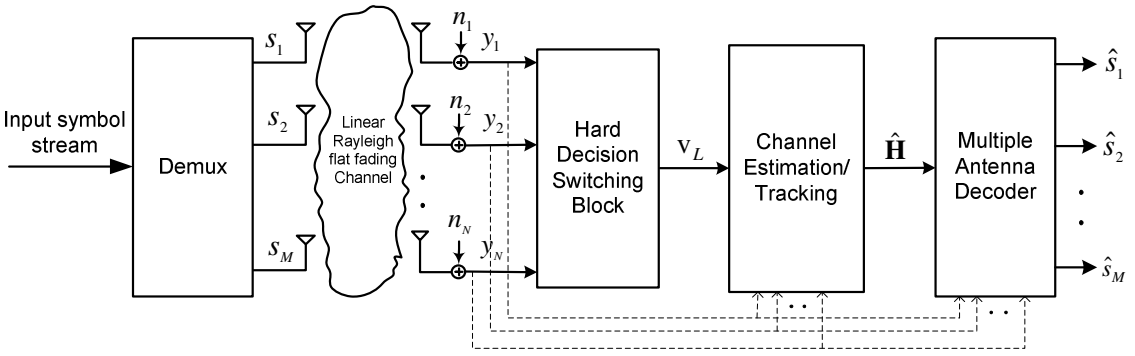

Figure 4-1 Block diagram of a SM-MIMO system employing the presented technique

4.2.1 Particle Filtering

The problem is now to track $\mathbf{H}_k$, as the new observation vector $\mathbf{y}_k$ is received. If $\mathbf{y}_{1:k} = \{\mathbf{y}_k \quad \mathbf{y}_{k-1} \quad \cdots \quad \mathbf{y}_1\}$ denotes the set of all previous observations up to time *k,* then the goal of the filtering is to estimate the probability density function (PDF) of the states, $p(\mathbf{H}_k | \mathbf{y}_{1:k})$, where $\mathbf{H}_k$ consists the set of all the state variables at iteration *k*, i.e., $\mathbf{H}_k = \{h_{n,m}(k); n = 1,..., N; m = 1,..., M\}$.

Particle filters from the sampling importance resampling (SIR) filter [4.12] firstly represent the state density $p(\mathbf{H}_k | \mathbf{y}_{1:k-1})$ by a set of "particles", each with a probability mass. Then, the probability mass of each particle is modified using the state and observation equations after a new independent sample is obtained from the posterior $p(\mathbf{H}_k | \mathbf{y}_{1:k})$ before proceeding to the next prediction/observation step. The particle filtering facilitated by ICA is summarized below.

At the end of $(k-1)$ th iteration, the SIR algorithm finds a set of N_{p} equally weighted particles that approximate the *a posteriori* $p(\mathbf{H}_{k-1} | \mathbf{y}_{1:k-1})$. We denote $\{\mathbf{H}_{k-1}^n\}_{n=1}^{N_{\mathrm{p}}}$ as the set of N_{p} particles where each is distributed as an independent sample from $p(\mathbf{H}_{k-1} | \mathbf{y}_{1:k-1})$. At time instant *k*, the filtering proceeds as follows.

4.2.1.1 Prediction stage

Having N_{p} samples, $\{\mathbf{H}_{k-1}^n\}_{n=1}^{N_{\mathrm{p}}}$, at time instant $k-1$, N_{p} samples are

drawn, $\left\{\mathbf{H}_{k|k-1}^{1},\mathbf{H}_{k|k-1}^{2},...,\mathbf{H}_{k|k-1}^{N_p}\right\}$, from the importance density $q(\mathbf{H}_k\left|\mathbf{H}_{1:k-1},\mathbf{y}_k\right.)$ [4.12]. The most popular choice for the importance density is the transitional prior [4.13] which can be obtained from formulation in (4-4) associated with mode *L*

$$q(\mathbf{H}_k\left|\mathbf{H}_{1:k-1},\mathbf{y}_k,L\right.)=q(\mathbf{H}_k\left|\mathbf{H}_{k-1},\mathbf{H}_{1:k-2},\mathbf{y}_k,L\right.)=p(\mathbf{H}_k\left|\mathbf{H}_{k-1},L\right.) \tag{4-7}$$

If the particles $\left\{\mathbf{H}_{k-1}^{n}\right\}_{n=1}^{N_p}$ are the independent samples from $p(\mathbf{H}_{k-1}\left|\mathbf{y}_{1:k-1}\right.)$, then the particles $\left\{\mathbf{H}_{k|k-1}^{n}\right\}_{n=1}^{N_p}$ will be the independent samples from $p(\mathbf{H}_k\left|\mathbf{y}_{1:k-1}\right.)$.

4.2.1.2 Update stage

By receiving the new observation, $\mathbf{y}_k$, the prediction states represented by the set $\left\{\mathbf{H}_{k|k-1}^{n}\right\}_{n=1}^{N_p}$ are updated (corrected). Each particle is weighted by the likelihood of the observation $\mathbf{y}_k$. The observation is generated from (4-1) with the elements of $\mathbf{H}_k$ assembled from the state variables, $\mathbf{H}_{k|k-1}^{n}$. We note since the MIMO system in our case, is operating over a flat fading channel (described by set of measurement equation (4-1) sampled at symbol time and the data vector symbols are independent), the current observation $\mathbf{y}_k$ is independent of the previous observations $\mathbf{y}_{k-1}$, $\mathbf{y}_{k-2}$, $\cdots$. Thus

$$p(\mathbf{y}_k\left|\mathbf{H}_{1:k},\mathbf{y}_{1:k-1}\right.)=p(\mathbf{y}_k\left|\mathbf{H}_k,\mathbf{H}_{1:k-1},,\mathbf{y}_{1:k-1}\right.)=p(\mathbf{y}_k\left|\mathbf{H}_k\right.)$$

where $\mathbf{H}_{1:k}=\left\{\mathbf{H}_k\quad\mathbf{H}_{k-1}\quad\cdots\quad\mathbf{H}_1\right\}$. This makes it possible to update the importance weights according to [4.12]

$$\begin{aligned}\tilde{w}_k^n &= w_{k-1}^n \frac{p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.) p(\mathbf{H}_{k|k-1}^n \left| \mathbf{H}_{k-1}^n, L \right.)}{q(\mathbf{H}_{k|k-1}^n \left| \mathbf{H}_{1:k-1}^n, \mathbf{y}_k, L \right.)} \quad ; \quad n = 1,..., N_{\mathrm{p}} \\ &= w_{k-1}^n p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.)\end{aligned} \tag{4-8}$$

and the corresponding normalized weights are

$$w_k^n = \frac{w_{k-1}^n p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.)}{\sum_{n=1}^{N_p} w_{k-1}^n p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.)} \quad ; \; n = 1,..., N_{\mathrm{p}} \tag{4-9}$$

Note that for each sample the likelihood function $p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.)$ is

$$p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.) = \int_{\mathbf{s}} p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n, \mathbf{s}_k \right.) p(\mathbf{s}_k) d\mathbf{s}_k \tag{4-10}$$

where $\mathbf{s}_k$ is the vector of sources (data symbols) at time instant *k,* modelled by (4-5). Owing to the assumed independence of original sources,

$$p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.) = \int_{\mathbf{s}} p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n, \mathbf{s}_k \right.) \prod_{i=1}^{M} p_{s_i}(s_i) d\mathbf{s}_k \tag{4-11}$$

Equation (4-11) may be evaluated using the Laplace approximation [4.14] or Monte Carlo integration. Monte Carlo integration is more general, but more computationally expensive [4.15]. The Laplace approximation is appropriate when the observation noise is small [4.14].

For the *N*-dimensional integral

$$I = \int f(\mathbf{s}) \exp\bigl(-h(\mathbf{s})\bigr) d\mathbf{s} \tag{4-12}$$

when $h(\mathbf{s})$ is sharply peaked, the Laplace approximation is

$$I \approx \int f(\tilde{\mathbf{s}}) \exp\left(-h(\tilde{\mathbf{s}}) - (\mathbf{s}-\tilde{\mathbf{s}})^T \mathbf{A}(\mathbf{s}-\tilde{\mathbf{s}})/2\right) d\mathbf{s} \tag{4-13}$$

where $\mathbf{A}$ is the Hessian matrix with the matrix elements $[\mathbf{A}]_{i,j} = \frac{\partial^2 h(\mathbf{s})}{\partial s_i \partial s_j^*}$ and $\tilde{\mathbf{s}}$ is the $\mathbf{s}$ that minimizes $h(\mathbf{s})$. Thus

$$I \approx \sqrt{\det(2\pi \mathbf{A}^{-1})} f(\tilde{\mathbf{s}}) \exp\left(-h(\tilde{\mathbf{s}})\right) \tag{4-14}$$

To relate the integral in (4-11) to the Laplace approximation, the additive observation noise is assumed to be Gaussian with zero mean and covariance matrix $\mathbf{R}$, as noted above. Thus, (4-11) becomes

$$p(\mathbf{y}_k \left| \mathbf{H}_{k|k-1}^n \right.) = \int_{\mathbf{s}} \frac{1}{(2\pi)^{N/2} (\det(\mathbf{R}))^{1/2}} \exp\left(-\frac{1}{2} \left(\sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n \mathbf{s}_k - \mathbf{y}_k \right)^H \mathbf{R}^{-1} \left(\sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n \mathbf{s}_k - \mathbf{y}_k \right) \right) \prod_{m=1}^{M} p_m(s_m)\, d\mathbf{s}_k \tag{4-15}$$

Using the family of generalized exponentials as the source model, the Gaussian term in (4-15) is considered as $\exp(-h(\mathbf{s}))$ and the product of generalized exponentials and normalizing factors as $f(\mathbf{s})$. This results in

$$h(\mathbf{s}_k) = (\sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n \mathbf{s}_k - \mathbf{y}_k)^H \mathbf{R}^{-1} (\sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n \mathbf{s}_k - \mathbf{y}_k)/2 \tag{4-16}$$

which is minimized at

$$\tilde{\mathbf{s}}_k^n = (\sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n)^H \mathbf{R}^{-1} \sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n)^{-1} \sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n)^H \mathbf{R}^{-1} \mathbf{y}_k \tag{4-17}$$

and the Hessian matrix is

$$\mathbf{A}_k^n = \frac{1}{M}\mathbf{H}_{k|k-1}^n{}^H \mathbf{R}^{-1}\mathbf{H}_{k|k-1}^n \tag{4-18}$$

Consequently the likelihood is approximated as

$$p(\mathbf{y}_k \mid \mathbf{H}_{k|k-1}^n) \approx$$
$$c\exp\left(-\left(\sqrt{\frac{1}{M}}\mathbf{H}_{k|k-1}^n\tilde{\mathbf{s}}_k^n - \mathbf{y}_k\right)^H \mathbf{R}^{-1}\left(\sqrt{\frac{1}{M}}\mathbf{H}_{k|k-1}^n\tilde{\mathbf{s}}_k^n - \mathbf{y}_k\right)/2\right)\exp\left(-\sum_{m=1}^{M}\left|\tilde{s}_m\right|^{r_i}\right) \tag{4-19}$$

where the normalizing factor *c* is

$$c = \frac{B}{\sqrt{\det(\frac{1}{M}\mathbf{H}_{k|k-1}^n{}^H \mathbf{R}^{-1}\mathbf{H}_{k|k-1}^n)\det(\mathbf{R})}} \tag{4-20}$$

and B is the product of the source normalizing constants, i.e. $B = b_1 b_2 ... b_M$.

Comparison with Monte-Carlo integration shows that this approximation is suitable over a wide range of noise covariance [4.14].

4.2.1.3 Resampling stage

For the resampling stage, the systematic resampling algorithm (SR) introduced in [4.16] is used. The particles, $\mathbf{H}_{k|k-1}^n$, and the weights, w_k^n, approximate $p(\mathbf{H}_k \mid \mathbf{y}_{1:k})$. With resampling, the particles with large weights are replicated and the ones with negligible weights are removed. Thus, the resampling process is completed with N_{p} replacements to form an approximate sample from $p(\mathbf{H}_k \mid \mathbf{y}_{1:k})$, where all particles are carrying equal weights. This new sample can now be used as the basis for the next prediction step.

Knowledge of the *a posteriori* density $p(\mathbf{H}_k \mid \mathbf{y}_{1:k})$, enables one to compute an optimal state estimate with respect to any criterion. For example, the MMSE estimate is the conditional mean of $\mathbf{H}_k$, i.e. $\hat{\mathbf{H}}_k = E\{\mathbf{H}_k \mid \mathbf{y}_{1:k}\}$, while the maximum *a posteriori* (MAP) estimate is the maximum of $p(\mathbf{H}_k \mid \mathbf{y}_{1:k})$, i.e. $\hat{\mathbf{H}} = \arg\max_{\mathbf{H}_k} p(\mathbf{H}_k \mid \mathbf{y}_{1:k})$. However, since a linear-Gaussian case (linear set of state and observation equations with additive Gaussian noises) is considered here, the mean state estimate is equivalent to the MAP state estimate. Note that $\hat{h}_{i,j}(k)\ ; i = 1,...,N; j = 1,...,M$ are the elements of $\hat{\mathbf{H}}_k$ which yield the channel matrix estimate $\hat{\mathbf{H}}_k$.

4.2.2 Source / Data Recovery

Having the estimate, $\hat{\mathbf{H}}_k$, the source vector $\mathbf{s}_k$ can be estimated in different ways. For example, using MAP estimation, $\hat{\mathbf{s}}_k$ is found by maximizing $\log\left(p(\mathbf{y}_k \mid \hat{\mathbf{H}}_k, \mathbf{s}_k)\, p(\mathbf{s}_k)\right)$, which is equivalent to minimizing

$$(\mathbf{y}_k - \sqrt{\frac{1}{M}}\hat{\mathbf{H}}_k \mathbf{s}_k)^H \mathbf{R}^{-1} (\mathbf{y}_k - \sqrt{\frac{1}{M}}\hat{\mathbf{H}}_k \mathbf{s}_k) + \sum_{i=1}^{M} \left| \frac{s_i - m_i}{\omega_i} \right|^{r_i} \qquad (4\text{-}21)$$

Using the pseudo-Newton method [4.17] when the noise variance is small, the estimated source would be

$$\hat{\mathbf{s}}_k \approx \sqrt{M}\,(\hat{\mathbf{H}}_k^H \hat{\mathbf{H}}_k)^{-1} \hat{\mathbf{H}}_k^H \mathbf{y}_k \qquad (4\text{-}22)$$

If there is no *a priori* knowledge about the source distribution, then the stability criterion needs checking at this point. If the criterion is not satisfied, then the ICA algorithm needs to try the other source distribution model (super- instead of sub-Gaussian, or vice versa), and start again from (4-15).

The estimator (4-22) has the same structure as a zero forcing (ZF) receiver which provides only $N-M+1$ order diversity [4.18] (the same as MMSE and successive cancellation (SUC) receivers but with different SNR loss). As an alternative, one may take advantage of ordered successive cancellation (OSUC) receiver [4.19] which may have more than $N-M+1$ order diversity because of the ordering (selection) process [4.18], or use the optimal ML receiver which extracts N order diversity with the expense of high decoding complexity (exponential in M). We simulate the presented decision-directed system with different decoding structures (ML, MMSE, ZF and OSUC) below. Finally, the estimated symbols are fed to the hard decision block to yield the detected transmitted symbols by setting the optimal thresholds in the constellation regions of the transmitted signal.

4.2.3 Hard Decision Switching Algorithm

The goal is to process observation $\mathbf{y}_k$ to determine the dominant speed mode of the mobile station in time sample *k*. At the start of each iteration (before proceeding to prediction/update stage of filtering), an artificial observation vector $\mathbf{z}_k$ is generated as follows

$$\mathbf{z}_k = \mathbf{y}_k + \mathbf{1} = \sqrt{\frac{1}{M}}\mathbf{H}_k\mathbf{s}_k + \mathbf{1} + \mathbf{n}_k = \left[\sqrt{\frac{1}{M}}\mathbf{H}_k \quad \mathbf{1}\right]\begin{bmatrix}\mathbf{s}_k \\ b\end{bmatrix} + \mathbf{n}_k \tag{4-23}$$

where $\mathbf{1}$ denotes a $N\times 1$ column vector with all elements of 1 and $b = 1 + j0$. In the arrangement of (4-23), *b* is again introduced and treated as an unknown deterministic variable for estimating. In this visualization, one may think of b as the signal generated from a hypothetical source. In essence, the artificial data vector $\mathbf{z}_k$ in (4-23), represents the effect of all real sources $\mathbf{s}_k = \begin{bmatrix} s_1 & s_2 & \cdots & s_M \end{bmatrix}^T$ including the hypothetical source *b* transmitted over an $N\times(M+1)$ MIMO channel with channel matrix $\tilde{\mathbf{H}}_k = \left[\sqrt{1/M}\,\mathbf{H}_k \quad \mathbf{1}\right]$. Since the source *b* does not exist in the real world but it affects the observation vector $\mathbf{z}_k$, it can be called a hypothetical source. The ZF estimate of *b* constrained to $E\{\hat{b}\} = 1 + j0$ [4.22] is

$$\hat{b} = \left[\left(\tilde{\mathbf{H}}_k^H\tilde{\mathbf{H}}_k\right)^{-1}\tilde{\mathbf{H}}_k^H\mathbf{z}_k\right]_{(M+1),1} \tag{4-24}$$

Note that $\tilde{\mathbf{H}}_k$ is an $N\times(M+1)$ matrix. As a result, the necessary condition for the matrix $\left(\tilde{\mathbf{H}}_k^H\tilde{\mathbf{H}}_k\right)^{-1}$ to exist is $M+1\le N$; This means that the above formulation, while applicable to MRC, is not applicable for a SISO or for the $N\times N$ MMO systems. For these cases, as an alternative, two artificial observation vectors $\mathbf{z}_k^{(1)}$ and $\mathbf{z}_k^{(2)}$ are generated as follows:

$$\mathbf{z}_k^{(1)} = \mathbf{y}_k + \mathbf{1} = \sqrt{\frac{1}{M}}\mathbf{H}_k\mathbf{s}_k + \mathbf{1} + \mathbf{n}_k \tag{4-25}$$

$$\mathbf{z}_k^{(2)} = \mathbf{y}_k - \mathbf{1} = \sqrt{\frac{1}{M}}\mathbf{H}_k \mathbf{s}_k - \mathbf{1} + \mathbf{n}_k \tag{4-26}$$

The two above equations can be rewritten in one formulation

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{z}_k^{(1)} \\ \mathbf{z}_k^{(2)} \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{1/M}\,\mathbf{H}_k & \mathbf{1} \\ \sqrt{1/M}\,\mathbf{H}_k & -\mathbf{1} \end{bmatrix}}_{\tilde{\mathbf{H}}_k} \begin{bmatrix} \mathbf{s}_k \\ b \end{bmatrix} + \begin{bmatrix} \mathbf{n}_k \\ \mathbf{n}_k \end{bmatrix} \tag{4-27}$$

Using ZF criterion yields an estimate of b as

$$\hat{b} = \left[\left(\tilde{\mathbf{H}}_k^H \tilde{\mathbf{H}}_k \right)^{-1} \tilde{\mathbf{H}}_k^H \mathbf{z}_k \right]_{(M+1),1} = \left[\left[\begin{array}{c|c} \frac{2}{M}\mathbf{H}_k^H \mathbf{H}_k & \mathbf{o} \\ \hline \mathbf{o}^T & 2N \end{array} \right]^{-1} \begin{bmatrix} \sqrt{1/M}\,\mathbf{H}_k & \mathbf{1} \\ \sqrt{1/M}\,\mathbf{H}_k & -\mathbf{1} \end{bmatrix} \mathbf{z}_k \right]_{(M+1),1} \tag{4-28}$$

with $\mathbf{o}$ denoting a $M \times 1$ column vector with all zero elements. Here $\tilde{\mathbf{H}}_k$ is a $2N \times (M+1)$ matrix and the necessary condition for the matrix $\left(\tilde{\mathbf{H}}_k^H \tilde{\mathbf{H}}_k \right)^{-1}$ to exist is $M + 1 \le 2N$ which is suitable for any $N \times N$ MIMO systems as well. This concludes the basic scheme.

To estimate the speed mode, first, a set $N_c << N_p$ of particles, $\left\{ \mathbf{H}_{k-1}^n \right\}_{n=1}^{N_c}$, is taken from time iteration $k-1$ and N_c samples $\left\{ \mathbf{H}_{k|k-1}^1, \mathbf{H}_{k|k-1}^2, ..., \mathbf{H}_{k|k-1}^{N_c} \right\}$ are generated using the prediction stage of the filtering for each of the possible dynamic modes $d = 1,...,\mathrm{D}$. That is, the samples $\left\{ \left(\mathbf{H}_{k|k-1}^n \right)_d \right\}_{n=1}^{N_c}; d = 1,...,\mathrm{D}$ are predicted. Then, for each dynamic mode, the corresponding mixing matrix $\left(\tilde{\mathbf{H}}_{k|k-1}^n \right)_d$ is formed using the sample point $\left(\mathbf{H}_{k|k-1}^n \right)_d$, and the estimation values $\left(\hat{b}_k^n \right)_d$ are obtained using (4-28) (or (4-24) for system with $M + 1 \le N$), accordingly.

Next, the estimated values of *b* for different speed modes, $\left(\hat{b}_k\right)_d$, are compared with the true value of $b = 1 + j0$, and the corresponding errors are recorded. The mode with the minimum root mean square error (RMSE) is then identified as the speed of the mobile for that iteration. The estimated RMSE for dynamic mode d is calculated as

$$\mathrm{RMSE}_k(d) = \sqrt{\frac{1}{N_c}\sum_{n=1}^{N_c}\left(\left(\hat{b}_k^n\right)_d - b\right)\left(\left(\hat{b}_k^n\right)_d - b\right)^*} \quad ; \; d = 1,\ldots,\mathrm{D} \tag{4-29}$$

If mode *L* has the least RMSE, then it is selected in the *k*th iteration, i.e.,

$$\text{mode } L = \arg\min_d \left\{\mathrm{RMSE}_k(d)\right\} \tag{4-30}$$

A small N_c allows the above decision with low complexity. After selecting mode *L,* the filtering begins by generating an N_p set of particles for the desired dynamic mode in the prediction step.

We note that for initializing the presented iterative algorithm, it is sufficient to use any of the approaches developed in [4.20] only once at the beginning of whole data sequence transmission. The receiver requires knowledge of the different values of mobile terminal's relative speed, the channel matrix variances, and one value of the channel normalized time correlation function, $\rho(T_s)$. We do not include estimation of these statistics as part of the algorithm. Including these estimates is a relatively straightforward extension, but the goal here is to quantify the behavior of the decision-directed estimation with known channel statistics.

4.3 Complexity Evaluation

Again, complete evaluation of the computational complexity is too complicated and only some insight can be offered into the complexity in one iteration (symbol time), through an approximate operations count for the implementation equations. An operation in this accounting is roughly equivalent to a multiply-and-accumulate. The joint decision-directed channel estimation and speed tracking technique comprises four stages: the switching algorithm, the prediction stage, the update stage, and the resampling stage. The complexity evaluation of the data/source recovery stage is not considered here since this depends on the type of decoder (ZF, MMSE, OSUC, ML, etc.) used.

Each iteration begins with the switching algorithm (described in section 4.2.3). Let N_s be the number of operations needed for drawing a sample from a statistical distribution function. Since the total number of state variables is NM, the complexity of predicting the samples $\left\{(\mathbf{H}^n_{k|k-1})_d\right\}_{n=1}^{N_c}; d=1,...,\mathrm{D}$ is $O\left(DN_cNMN_s\right)$. Now, for each speed mode candidate the estimation values $(\hat{b}_k^n)_d; n=1,...,N_c$ are obtained from (4-28). A breakdown of operations counts in implementing (4-28) is as follows (the formulas (in angular brackets) above the matrix formulas give the rough operation counts for implementing this formula)

$$\hat{b}=\left[\underbrace{\left(\overbrace{\tilde{\mathbf{H}}_k^H\times\tilde{\mathbf{H}}_k}^{\langle(M+1)^2N\rangle}\right)^{-1}}_{\langle(M+1)^3\rangle}\overbrace{\times\tilde{\mathbf{H}}_k^H}^{\langle(M+1)^2N\rangle}\underbrace{\times\mathbf{z}_k}_{\langle N\rangle}\right]_{(M+1),1} \tag{4-31}$$

This requires the total of $DN_c((M+1)^2N+(M+1)^3+(M+1)^2N+N)$ operations for all the speed mode candidates.

Equation (4-29) has one nested summation for each of the speed mode candidates. In addition, let N_{r} be the number of operations needed for calculating $\mathrm{sqrt}(.)$ function. Thus equation (4-29) has the complexity of $O\left(D(N_c+N_{\mathrm{r}})\right)$. Finally, the overall complexity of the switching algorithm is $O\left(D\,N_c((M+1)^2(M+N)+NM\,N_s)+DN_{\mathrm{r}}\right)$.

Once the dominant speed mode is determined, the prediction stage of particle filtering has the complexity of $O\left(N_{\mathrm{p}}NM\,N_s\right)$ in order to obtain the N_{p} predicted samples $\left\{\mathbf{H}_{k|k-1}^{1},\mathbf{H}_{k|k-1}^{2},...,\mathbf{H}_{k|k-1}^{N_{\mathrm{p}}}\right\}$.

In the update stage, importance weights associated with N_{p} predicted particles have to be updated according to (4-8). In the worst case (if there is no *a priori* knowledge about the source distribution), this involves calculating the N_{p} likelihood functions $p(\mathbf{y}_k\left|\mathbf{H}_{k|k-1}^{n}\right.)$ (according to (4-19)) for each combination of source distribution models that is assigned within the ICA. Here, for each particle, $\tilde{\mathbf{s}}_k^n$ has to be obtained via implementing (4-17), first. A breakdown of operations counts in implementing (4-17) and $h(\tilde{\mathbf{s}}_k^n)$ is shown below, for the case $\mathbf{R}=\sigma_n^2\mathbf{I}_N$:

$$\tilde{\mathbf{s}}_k^n=\sqrt{M}\underbrace{\left(\overbrace{\mathbf{H}_{k|k-1}^{n}{}^{H}\times\mathbf{H}_{k|k-1}^{n}}^{\langle M^2N\rangle}\right)^{-1}}_{\langle M^3\rangle}\overbrace{\times\mathbf{H}_{k|k-1}^{n}{}^{H}}^{\langle M^2N\rangle}\underbrace{\times\mathbf{y}_k}_{\langle MN\rangle} \tag{4-32}$$

$$h(\tilde{\mathbf{s}}_k^n) = \underbrace{(\underbrace{\sqrt{\frac{1}{M}}}_{\langle N \rangle} \times \overbrace{\mathbf{H}_{k|k-1}^n \times \tilde{\mathbf{s}}_k^n}^{\langle NM \rangle} - \mathbf{y}_k)^H \times (\sqrt{\frac{1}{M}} \mathbf{H}_{k|k-1}^n \tilde{\mathbf{s}}_k^n - \mathbf{y}_k)}_{\langle N \rangle} / 2\sigma_n^2 \tag{4-33}$$

Furthermore, if N_{f} denotes the total number of operations needed for calculating all functions within (4-19) (e.g. $\exp(.)$, $\det(.)$, $\mathrm{sqrt}(.)$, etc.) the overall complexity of (4-19) calculated for all N_{p} particles is $O\left(N_{\mathrm{p}}((M)^2(M+N)+N_{\mathrm{f}})\right)$. We note that this operations count is only for one combination of the source distributions assigned within the ICA.

As mentioned above, if there is no *a priori* knowledge about the source distribution, then the stability criterion needs checking once the sources are detected at the data recovery stage (Section 4.2.2), for each iteration. If the criterion is not satisfied, then the ICA algorithm needs to try the other combinations of source distributions (sub- or super-Gaussian), and repeat the update stage of particle filtering. Since there is a total of 2^M combinations of source distribution models, in the worst case, the update stage has the overall complexity of $O\left(2^M N_{\mathrm{p}}((M)^2(M+N)+N_{\mathrm{f}})\right)$. However, if some prior knowledge of the data is available at the receiver, a proper choice for the source distributions can be set within the ICA only once at the beginning (e.g. with QPSK signaling, the source densities in (4-5) are chosen to be uniform), and this reduces the complexity significantly. This is the usual case - with most receivers there is some prior knowledge of the signaling type.

After the update stage, there is the resampling stage. In general, any resampling algorithm (e.g. systematic resampling (SR), residual systematic

resampling (RSR), etc.) can be used here. The complexity of the SR and RSR algorithms is of $O(N_{\mathrm{p}})$ [4.21]. So the resampling stage has little contribution to the overall system complexity.

4.4 Simulation Results

As noted above, the ML receiver performance lower bounds the error rate probability of other sub-optimal (ZF, MMSE, OSUC, and etc.) receivers. No exact, optimal, analytical solution is available. However, several upper bounds on error probability for the ML as well as other sub-optimal receivers have been derived. Thus, the BER performance results are simulated using perfect CSI, as the benchmark, and these are evaluated and compared to the results of the presented decision-directed system with the associated perfect CSI reference ones.

In all the simulations, $N_{\mathrm{p}} = 200$ and $N_c = 20$. For convenience, i.i.d. QPSK data symbols with zero mean and variance 1 are used, but the presented technique is applicable to any signaling. Using this prior knowledge on the data, the source densities in (4-5) are chosen to be uniform. In other words, in view of the ICA, *M* sources, each with uniform density are mixed to yield the observation. For convenience of interpretation, the data symbol duration is $T_{\mathrm{s}} = 0.1\,\mathrm{ms}$ (i.e., 10k symbols/sec); and the channel matrix coefficients all have the unity variance, i.e., $\sigma^2_{h_{i,j}} = 1$. The presented algorithm is formulated in a general format and it can be applied to any MIMO system with maneuvering or non-maneuvering terminals. In order to assess the impact of the presented channel estimation

technique on system error performance more directly, the first subsection below is dedicated to MIMO systems with non-maneuvering terminals (the speed is not changing). This choice also allows fair performance comparison with a known system employing Kalman filtering to track the channel [4.3] and two pilot-aided systems (least-squares (LS), and an MMSE pilot-aided system [4.4]). Throughout the simulations, the signal-to-noise ratio (SNR) refers to each of the receivers.

4.4.1 MIMO System with Non-Maneuvering Terminals

Figure 4-2 illustrates the symbol error rate (SER) performance of the presented system along with the coherent detection curves for ML, MMSE, ZF, and MMSE-OSUC receivers employing specially multiplexing and horizontal

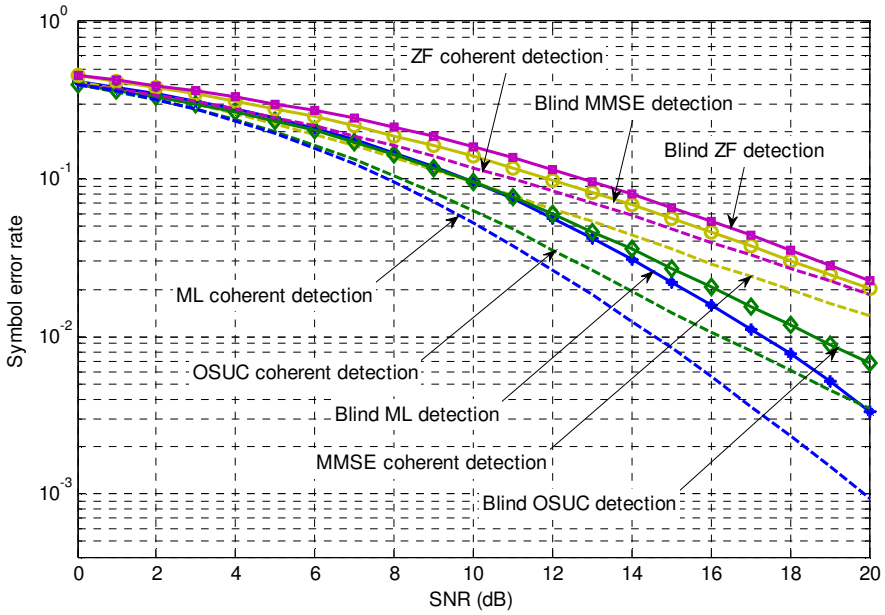

Figure 4-2 SER performance of a 2×2 MIMO system employing the presented decision-directed channel estimation technique, along with the coherent detection curves for ML, MMSE, ZF, and MMSE-OSUC receivers

The system employs specially multiplexing and horizontal encoding (SM-HE), at fade rate $f_{\mathrm{D}}T_s = 0.001$ (non-maneuvering terminals).

encoding (SM-HE), with $M = N = 2$ and fade rate $f_{\mathrm{D}}T_{\mathrm{s}} = 0.001$ (this fade rate corresponds to maximum Doppler frequency $f_{\mathrm{m}} = 10\,\mathrm{Hz}$ or the vehicle speed $\mathrm{v} \approx 4.5\,\mathrm{km/h}$ at $2.4\,\mathrm{GHz}$ band). With each receiver type, the presented system exploits the same diversity order as coherent detection, but experiences an SNR loss which depends on the system parameters and the type of decoder used. In particular, it is within 3 dB (specifically, 2.5) of the perfect CSI curves in an ML (OSUC) decoder at fade rate $f_{\mathrm{D}}T_{\mathrm{s}} = 0.001$ and symbol error probability of 10^{-2}.

We may compare the results of the presented decision-directed system with a known decision-directed system [4.3] which employs Kalman-based estimation with channel prediction, and with two pilot-based systems: least-squares (LS), and an MMSE pilot-aided system [4.4].

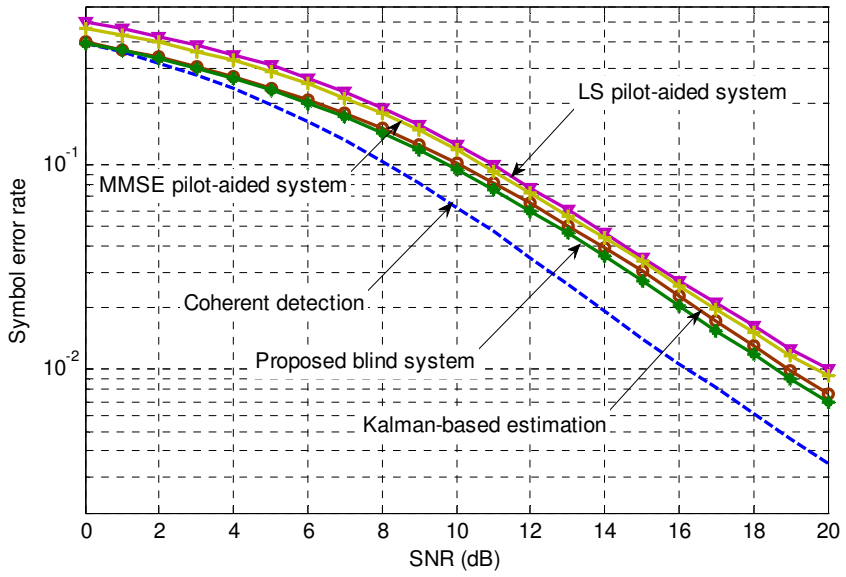

Figure 4-3 SER performance of a 2×2 MIMO system employing MMSE-OSUC decoder with different channel estimation techniques.

The system uses SM-HE, at fade rate $f_{\mathrm{D}}T_{s} = 0.001$ (non-maneuvering terminals).

The results are shown in Figure 4-3 for fade rate 0.001, and using MMSE-OSUC receiver. The presented system outperforms the other techniques, including a slight improvement in performance over the Kalman-based estimation at moderate-high values of SNR. The reason for this slight improvement is not clear (the ICA operation is complicated) but it is noted below that the Kalman filter has considerably less complexity.

Compared to pilot-aided systems, the presented ICA-based system has superior performance for all SNRs, at least for the fade rate used here. This is expected, because in pilot-aided systems, the channel is estimated only after at least M consecutive symbols, whereas in the presented technique, the channel is estimated at each symbol time which results in more accurate channel estimates in the time-varying environment.

4.4.2 MIMO System with Maneuvering Terminals

Consider a mobile station (vehicle) in a data communication link operating at $2.4\,\mathrm{GHz}$ band with three constant speed modes $\mathrm{v}_1 = 45\ \mathrm{km/h}$, $\mathrm{v}_2 = 67.5\ \mathrm{km/h}$ and $\mathrm{v}_3 = 90\ \mathrm{km/h}$. According to (4-2), these velocities correspond to the maximum Doppler frequencies $f_{\mathrm{m}}^{(1)} = 100\,\mathrm{Hz}$, $f_{\mathrm{m}}^{(2)} = 150\,\mathrm{Hz}$ and $f_{\mathrm{m}}^{(3)} = 200\,\mathrm{Hz}$ (fade rates $f_{\mathrm{m}}^{(1)} T_s = 0.01$, $f_{\mathrm{m}}^{(2)} T_s = 0.015$ and $f_{\mathrm{m}}^{(3)} T_s = 0.02$), respectively.

As an example, the relative speed of mobile terminals is tracked for $60\,\mathrm{s}$ in a SISO data communication link with the following discrete speed trajectory

$$v = \begin{cases} 67.5\,\text{km/h} & ; \quad 0\,\text{s} \le t \le 10\,\text{s} \\ 45\ \text{km/h} & ; \quad 10\,\text{s} \le t \le 20\,\text{s} \\ 67.5\ \text{km/h} & ; \quad 20\,\text{s} \le t \le 30\,\text{s} \\ 90\,\text{km/h} & ; \quad 30\,\text{s} \le t \le 40\,\text{s} \\ 67.5\ \text{km/h} & ; \quad 40\,\text{s} \le t \le 60\,\text{s} \end{cases} \tag{4-34}$$

At each symbol time $T_s = 0.1\,\text{ms}$, the dominant speed mode is detected by the presented hard decision switching algorithm; however, for better illustration an instantaneous result of the detected speed is shown every 0.1s in Figure 4-4 at $\text{SNR} = 20\,\text{dB}$. For this SNR, the probability of false detection would be approximately $20\,\%$.

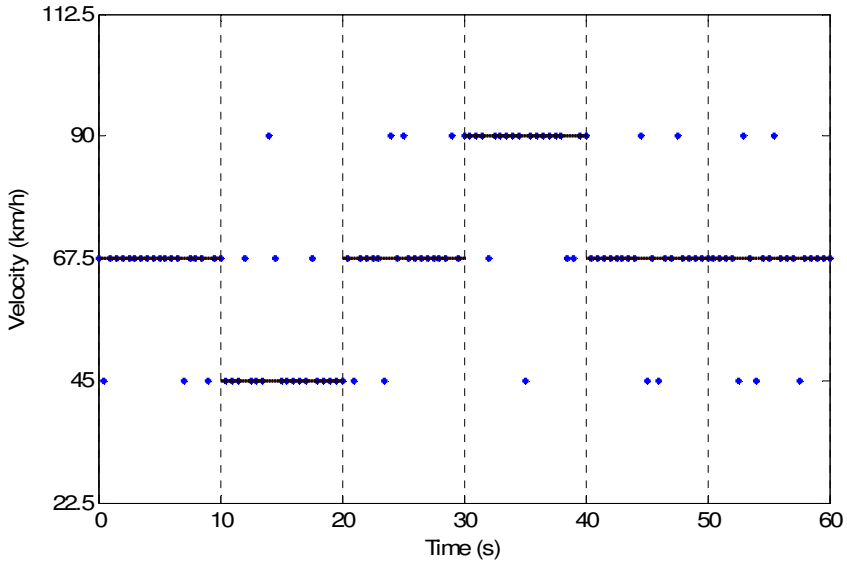

Figure 4-4 Instantaneously detected discrete speeds for a SISO system with maneuvering terminals operating at $\text{SNR} = 20\,\text{dB}$.

The results are shown every 0.1s. The true speed is shown as a solid line.

To further analyze the effect of the maneuvering terminals on SER performance, the presented decision-directed channel estimation technique equipped with the new switching algorithm is applied to different systems (a

SISO, a MRC system with two receive antennas and a 2×2 MIMO system) with the mobile speed motion as in (4-31). The results are illustrated in Figure 4-5, Figure 4-6, and Figure 4-7.

The comparison is also made with SER results of the systems under non-maneuvering scenario having constant relative speed $\mathrm{v}\approx 67.5\,\mathrm{km/h}$ at 2.4 GHz band ($\mathrm{v}\approx 67.5\,\mathrm{km/h}$ is the average speed of real speed motions in (4-31)). As the SNR increases, the difference between the SER curves obtained in the maneuvering and non-maneuvering scenarios decreases for low-to-mid values of SNR and stays within a fraction of dB for higher SNRs. This is because for smaller values of SNR, the switching block experiences a larger estimation error which in turn increases the probability of false detection.

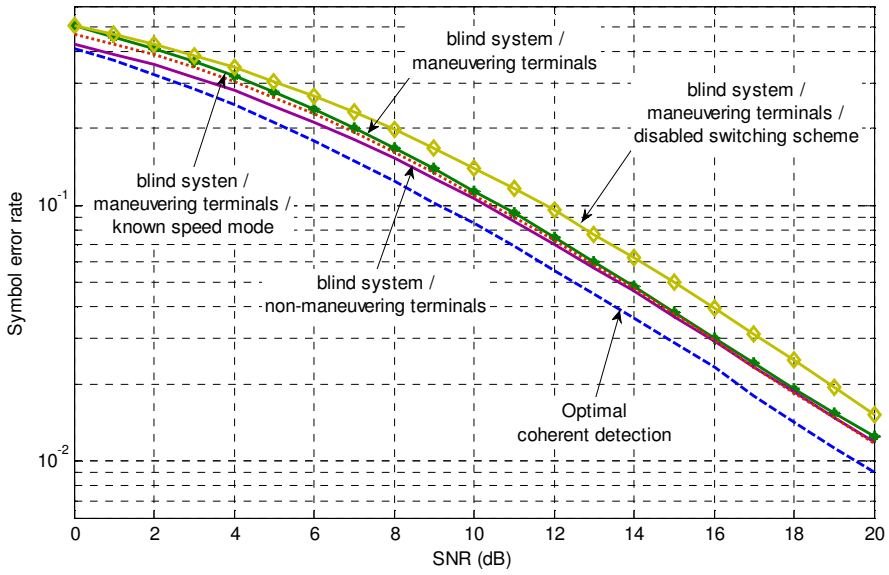

Figure 4-5 SER performance of a SISO system employing the presented decision-directed channel estimation in maneuvering and non-maneuvering scenarios. The comparison is also made with the case where the switching dynamic scheme is disabled in the algorithm. For the non-maneuvering case, the relative speed of terminals is considered to be $\mathrm{v}\approx 67.5\,\mathrm{km/h}$ in 2.4 GHz band.

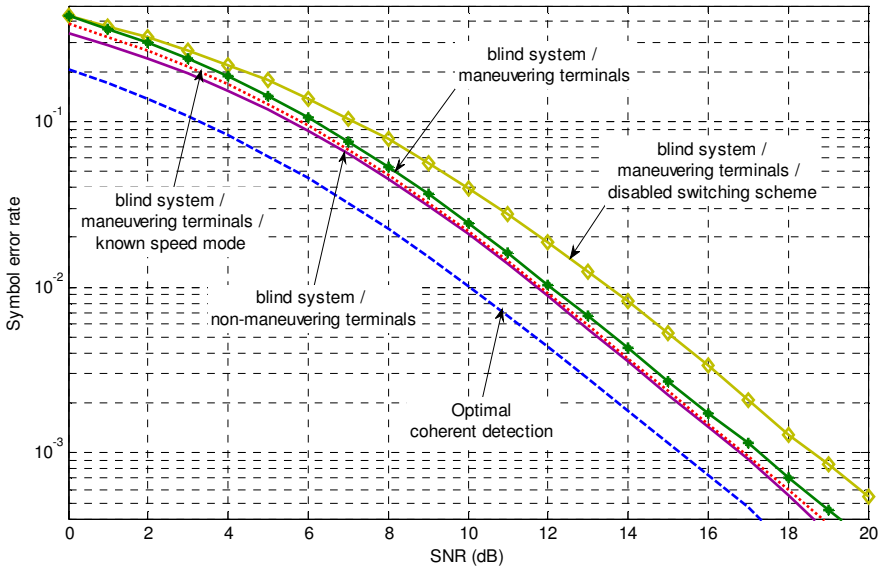

Figure 4-6 SER performance of an MRC system with two receive antennas with the presented decision-directed channel estimation in maneuvering and non-maneuvering scenarios.

The comparison is also made with the case where the switching dynamic scheme is disabled in the algorithm.

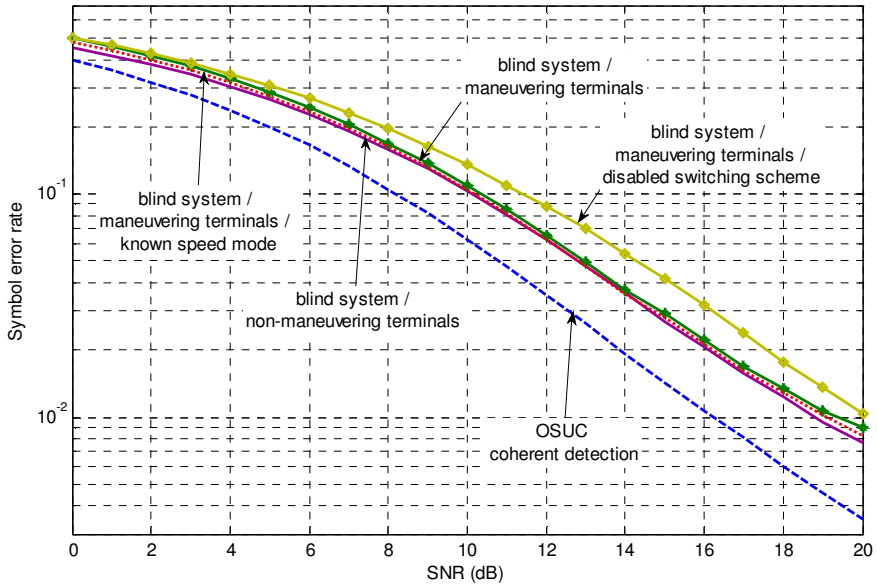

Figure 4-7 SER performance of a 2×2 MIMO system employing MMSE-OSUC detection and the presented decision-directed channel estimation in maneuvering and non-maneuvering scenarios.

The comparison is also made with the case where the switching dynamic scheme is disabled in the algorithm.

We can also benchmark our results with the one obtained when the switching algorithm knows perfectly the speed of the MS (dotted line) using the same channel estimation technique. With known speed modes, the resultant curves nearly coincide with the ones obtained under non-maneuvering for mid-high values of SNR.

Finally, if the switching block is disabled in the algorithm within the maneuvering scenario, a large degradation in performance is observed. This is also illustrated in Figure 4-5, Figure 4-6, and Figure 4-7 where the decision-directed channel estimation technique with parameters set for fix fade rate 0.015 (corresponding to the average speed $v \approx 67.5\,\mathrm{km/h}$) is applied without using switching dynamic scheme. This demonstrates the advantage of the switching block in systems with maneuvering transmit/receive terminals.

4.5 Summary and Conclusions

In this chapter, a new ICA-based decision-directed channel estimation is presented which employs particle filtering to track the time-varying channel matrix in spatially multiplexed, horizontally encoded MIMO systems having maneuvering terminals. The introduction of the new switching scheme in conjunction with the decision-directed channel estimation technique makes it possible to track the discrete relative speed of the maneuvering terminals. This is accomplished in parallel with channel estimation and data recovery based on only the received data information signal. The performance is very promising. The advantage of introducing the switching dynamic block is also demonstrated by disabling it within the algorithm which results in large degradation in

performance. The presented system also performs well in non-maneuvering scenario where it outperforms (by a small amount) known systems using Kalman-based estimation and established pilot-aided systems.

REFERENCE LIST

[4.1] S. A. Banani, and R. G. Vaughan, “Blind Channel Estimation and Speed Tracking in Wireless Systems using ICA with Particle Filtering,” accepted in *IET Communications*, Manuscript ID: COM-2010-0836R1, Oct. 2011.

[4.2] S. A. Banani, and R. G. Vaughan, “ICA with particle filtering for blind channel estimation in high data-rate MIMO systems,” in *Proc. IEEE Wireless Commun. and Network. Conf.,* pp. 1-6 , Apr. 2010.

[4.3] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, “Multi-input multi-output fading channel tracking and equalizationu using Kalman estimation”, *IEEE Trans. Signal Process.,* vol. 50, no. 5, pp. 1065-1076, May, 2002.

[4.4] M. Biguesh, and A. B. Gershman, “Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals,” *IEEE Trans. Signal Process.,* vol. 54, no. 3, pp. 884-893 , Mar. 2006.

[4.5] J. Sok Son, G. Thomas, and B. C. Flores, *Range-Doppler Radar Imaging and Motion Compensation,* Artech House, 2001.

[4.6] G. Morris, L. Harkness, *Airborne Pulsed Doppler Radar,* 2nd edition Artech House, 2006.

[4.7] H.Wang and P. Chang, “On verifying the first-order Markovian assumption for a Rayleigh fading channel model,” *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 353–357, May 1996.

[4.8] R. Everson and S. Roberts, *Independent component analysis*, *principles and practice*, Cambridge, UK: Cambridge University Press, 2001.

[4.9] T-W. Lee, M. Girolami, and T.J. Sejnowski, "Independent component analysis using an extended infomax Algorithm for mixed-sub-Gaussian and super-Gaussian sources," *Neural Computation,* vol. 11, pp. 417-441, 1999.

[4.10] R.M. Everson and S.J. Roberts, "ICA: A flexible non-linearity and decorrelating manifold approach," *Neural Computation,* vol. 11, pp. 1957-1983, 1999.

[4.11] H. Attias, “Independent factor analysis,” *Neural Computation*, vol. 11, pp. 803 -852, 1999.

[4.12] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” *IEEE Trans. Signal Process.,* vol. 50, no. 2, pp. 174–188, Feb. 2002.

[4.13] A. Doucet, S. Godsill, and C. Andrieu, “On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing,* vol. 10, no. 3, pp. 197-208, 2000.

[4.14] J. J. K. O' Ruanaidth and W.J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, Cambridge, UK: Springer, 1996.

[4.15] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, New York: Academic Press, 1984.

[4.16] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian non-linear state space models," *J. Computational and Graphical Statistics,* vol. 5, no.1, pp. 1-25, 1996.

[4.17] S. S. Oren, “On quasi-Newton and pseudo-Newton algorithms,” *J. Opt. Theory and Application,* vol. 20, no. 2, pp. 155-170, Oct. 1976.

[4.18] A. Paulraj, R. Nabar and D. Gore, *Introduction to space-time wireless communications,* Cambridge, UK: Cambridge University Press, 2003.

[4.19] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, “Simplified processing for high spectral efficiency wireless communication employing multi-element arrays,” *IEEE JSAC,* vol. 17, no. 11, pp. 1841-1852, Nov. 1999.

[4.20] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, “Estimation and equalization of fading channels with random coefficients,” *Signal Process.*, vol. 53, no. 2–3, pp. 211–229, Sep. 1996.

[4.21] S. Hong, P. M. Djuric, and M. Bolic, “Resampling algorithms for particle filters: A computational complexity perspective,” *EURASIP J. Applied Signal Process.,* vol. 2004, pp. 2267-2277, Jan. 2004.

[4.22] A. H. Sayed, *Fundamentals of Adaptive Filtering,* Hoboken, NJ: Wiley, 2003.

5: PRACTICABLE CAPACITY IN EIGEN-MIMO WITH CHANNEL ESTIMATION AND MEAN FEEDBACK

To achieve high capacity in wireless communication systems, multiple antennas at both ends of the fading channel are used to increase the spectral efficiency. The foundation for increasing capacity efficiency (bits per sec per Hz) is understanding the behavior of the information-theoretic capacity (strictly should be expressed as bits per channel use). It is possible, in principle, for the capacity to increase linearly with the number of antennas even when the channel state information (CSI) is unknown at the transmitter [5.1]. A further increase in capacity is possible when CSI is known at the transmitter [5.2]. These capacity results require many assumptions, including the need for the CSI to be known perfectly. In practice, perfect CSI is never known perfectly a priori and its estimation, without using an ideal blind technique (none are available), requires channel resource which reduces the capacity efficiency. So the practical comparison of open- and closed-loop MIMO capacities requires care because of their information theoretic nature. Fine differences may be overwhelmed by the engineering compromises inherent in a practical design for a given channel type. Such compromises include the deployment of practical communications techniques including choices of modulation such as QAM signaling (instead of Gaussian signals) and coding, pulse shaping and filter sizes, and finite block lengths (instead of infinitely long codes), and the associated guardbands which contribute to the required bandwidth, etc. Furthermore, eigen-MIMO requires

channel state information at the transmitter, and the capacity is compromised by factors such as channel estimation error and imperfect feedback. In particular, a closed loop capacity should strictly include the bandwidth resource required for the feedback link, but most of the current literature does not address this as part of the capacity. This chapter addresses that feedback requirement.

When the channel is reciprocal (as can be arranged in some time division duplex systems), the transmitter can use the uplink signal to estimate the downlink channel [5.3]. But in general the CSI estimated at the receiver must be linked back to the transmitter via CSI *feedback-link* mechanisms. The required channel resource for the feedback link can be relatively small in terms of bandwidth consumption, but its inclusion requires considerable processing and, in particular, protocol (i.e., spectral) resource. In this sense, the CSI feedback link becomes an expensive aspect of the MIMO system. The feedback link is imperfect in practice, and this is sometimes referred to as partial feedback. There have been previous investigations of MIMO performance degradation from channel estimation and imperfect feedback, for example, [5.4]-[5.13], and other references too numerous to list owing to the topicality of the research. In these examples, the optimal feedback and transmission strategy in the presence of a feedback link capacity constraint is considered in [5.4]. The optimal MIMO multiuser transmission design with a partial power feedback constraint is investigated [5.5] and [5.6]; and the influence of CSI feedback errors on capacity of linear multi-user MIMO systems is studied [5.7]-[5.9]. Furthermore, the effect of channel estimation error on the capacity of MIMO systems is investigated in

[5.10]-[5.11], while the optimal training and feedback design is given in [5.12]-[5.13]. However, none of the previous works investigate the joint impact of channel estimation and imperfect feedback on the practicable capacity in a digital link; although the capacity penalty resulting from using finite block length and uncoded QAM with the assumption of perfect CSI is investigated, e.g. in [5.17]-[5.18].

In this work [5.19]-[5.20], we investigate the simultaneous effect of training-based channel estimation and imperfect CSI feedback link (in particular, mean feedback) on both the practicable capacity and the information-theoretic capacity of eigen-MIMO systems with a feedback throughput constraint which imposes a maximum number of feedback bits per symbol duration. The capacities are given as a function of receiver branch signal-to-noise ratio (SNR), number of training symbols, channel estimation variance, feedback delay, feedback quantization noise, and the feedback throughput constraint. Then the capacities are maximized by solving different individual/joint optimization problems, using both analytical and simulation approaches.

An analytical formulation is given for the optimal allocation of transmission power between the training and the data payload. Furthermore, by considering the feedback codeword size (number of VQ indexing bits) as a variable, the capacity is jointly optimized over the number of training symbols and the feedback delay interval length. A comparison is also possible with the maximized theoretic capacity of an open-loop MIMO system such as that given in [5.14]. The comparison is only truly fair if the bandwidth resource required for the CSI

feedback link is accounted for. For channels that change extremely slowly, the feedback link can have a negligible capacity compared to the MIMO transmission, and in such a case the capacity requirement of the feedback link can be omitted. Even though the total time duration for training and data transmission and the SNR are reduced (effective values are used in the formulation) as the result of feedback delay time and feedback noise, respectively, the closed-loop theoretic capacity remains larger than the open-loop one, at least for low (SNR~6 dB) to mid (SNR ~14 dB) values of SNR.

The rest of the chapter is organized as follows. Section 5.1 describes the closed-loop MIMO system model. The information-theoretic/practicable capacity and the optimization problems are formulated in section 5.2. The simulations are presented in section 5.3, with the conclusions in section 5.4. The channel matrix $\mathbf{H}$ is treated as random with i.i.d. zero-mean complex Gaussian entries each with variance 1. Estimators of $\mathbf{H}$ are for a particular channel realization corresponding to the current block of the received data.

5.1 The Closed-Loop MIMO System Model

We consider an $N \times M$ MIMO quasi-static, flat block-fading channel model. The channel state remains static over a fading block of T symbols duration, but becomes independent across different blocks. For every block, there is a header of training symbols for CSI estimation at the receiver.

In order to estimate the channel matrix $\mathbf{H}$, let $T_\tau \geq M$ symbol durations be dedicated to channel sounding, i.e., training. Denote $\mathbf{S}_\tau$ as the $M \times T_\tau$ matrix of

training symbols transmitted with power constraint $\mathrm{tr}\mathbf{S}_\tau\mathbf{S}_\tau^H = MT_\tau$. The assumption of perfect symbol synchronization allows the received $N\times T_\tau$ matrix of training symbols to be written as the usual linear, additive noise, MIMO model,

$$\mathbf{Y}_\tau = \sqrt{\frac{\rho_\tau}{M}}\,\mathbf{H}\mathbf{S}_\tau + \mathbf{V}_\tau \tag{5-1}$$

where $\mathbf{V}_\tau$ is an $N\times T_\tau$ matrix representing the noise with $\mathbf{V}_\tau \sim CN(0,1)$, and ρ_τ is the SNR for the training symbols. Further below, a different SNR is introduced for the data symbols, and it is convenient to keep these SNRs different for now.

The channel matrix can be estimated under different optimization criteria, such as maximum likelihood, minimum-mean square error (MMSE), least squares, etc. We use the MMSE estimator because of its pragmatic tradeoff between simplicity and performance. Then the channel matrix estimate is written as

$$\hat{\mathbf{H}}_{\mathrm{r}} = \sqrt{\frac{M}{\rho_\tau}}\,\mathbf{Y}_\tau\mathbf{S}_\tau^H\left(\frac{M}{\rho_\tau}\mathbf{I}_M + \mathbf{S}_\tau\mathbf{S}_\tau^H\right)^{-1} \tag{5-2}$$

with the associated channel estimation error $\tilde{\mathbf{H}}_{\mathrm{r}} = \mathbf{H} - \hat{\mathbf{H}}_{\mathrm{r}}$. Since $\hat{\mathbf{H}}_{\mathrm{r}}$ is zero-mean, its variance can be written as $\sigma^2_{\hat{\mathbf{H}}_{\mathrm{r}}} = (NM)^{-1}E\,\mathrm{tr}\hat{\mathbf{H}}_{\mathrm{r}}\hat{\mathbf{H}}_{\mathrm{r}}^H$, and from the orthogonality principles of MMSE estimates,

$$\sigma^2_{\hat{\mathbf{H}}_{\mathrm{r}}} = 1 - \sigma^2_{\tilde{\mathbf{H}}_{\mathrm{r}}} \tag{5-3}$$

where $\sigma^2_{\tilde{\mathbf{H}}_{\mathrm{r}}} = (NM)^{-1}E\,\mathrm{tr}\tilde{\mathbf{H}}_{\mathrm{r}}\tilde{\mathbf{H}}_{\mathrm{r}}^H$. Thus, the channel estimation can be approximated by the following random process ($\hat{\mathbf{H}}_r$ is the estimate of $\mathbf{H}$)

$$\hat{\mathbf{H}}_{\mathrm{r}} \approx \mathbf{H} - \sigma_{\tilde{\mathbf{H}}_{\mathrm{r}}} \tilde{\mathbf{W}}_{\tilde{\mathbf{H}}_{\mathrm{r}}} \tag{5-4}$$

where $\tilde{\mathbf{W}}_{\tilde{\mathbf{H}}_{\mathrm{r}}}$ is a $N \times M$ matrix with $\tilde{\mathbf{W}}_{\tilde{\mathbf{H}}_{\mathrm{r}}} \sim CN(0,1)$. As with (2-11) and (3-11), the approximation model of (5-4) is only used for deriving the proposed algorithm. In the simulations, the correct structure of MMSE estimator (without approximation) is used for the channel estimation in (5-2).

Dedicating a larger interval for training improves the channel estimation and reduces $\sigma^2_{\tilde{\mathbf{H}}_{\mathrm{r}}}$; however, for a fixed feedback delay, less channel time is available for data transmission.

Once the channel is estimated at the receiver, a quantized form of the channel matrix (using a linear quantizer or rate distortion method) is fed back to the transmitter, via the CSI feedback link. The number of bits of the quantizer is *b* for the real and the imaginary part, i.e., there are 2*b* bits per complex channel coefficient, and there are MN channel coefficients. This means that the total number of bits for each estimate of CSI feedback is $B = 2MNb$ bits.

These B bits are to be fed back to the transmitter with a throughput constraint of R_{fb} bits per symbol duration. As a result, it takes a time of $T_{\mathrm{fb}} = B / R_{\mathrm{fb}}$ symbol times to feed back the codeword, excluding any protection coding bits for it, although the protection can be accounted by including a non-unity coding rate in T_{fb}. Such channel feedback coding is not included here. This is because it is assumed that the CSI feedback errors are dominated by the quantization process [5.7]. Here, a special form of imperfect feedback is

considered, called *mean feedback* [5.6] and [5.15], where the estimated channel at the receiver is viewed from the transmitter end as the following random process ($\hat{\mathbf{H}}_t$ is an estimate of $\hat{\mathbf{H}}_r$)

$$\hat{\mathbf{H}}_{\mathrm{t}} \approx \hat{\mathbf{H}}_{\mathrm{r}} + \sigma_{\mathrm{fb}} \mathbf{W}_{\mathrm{fb}} \tag{5-5}$$

with corresponding variance

$$\sigma^2_{\hat{\mathbf{H}}_{\mathrm{t}}} = \sigma^2_{\hat{\mathbf{H}}_{\mathrm{r}}} + \sigma^2_{\mathrm{fb}} = 1 - \sigma^2_{\tilde{\mathbf{H}}_{\mathrm{r}}} + \sigma^2_{\mathrm{fb}} \tag{5-6}$$

where σ^2_{fb} is the overall feedback noise variance, and $\mathbf{W}_{\mathrm{fb}}$ is an $N \times M$ matrix with $\mathbf{W}_{\mathrm{fb}} \sim CN(0,1)$. Thus, the channel matrix can be approximated as, *cf.*, (5-4),

$$\mathbf{H} \approx \hat{\mathbf{H}}_{\mathrm{t}} + \sigma_{\tilde{\mathbf{H}}_{\mathrm{r}}} \tilde{\mathbf{W}}_{\tilde{\mathbf{H}}_{\mathrm{r}}} - \sigma_{\mathrm{fb}} \mathbf{W}_{\mathrm{fb}} . \tag{5-7}$$

The signs on the rhs are different in this zero-mean formulation, but it doesn't matter because the added terms are random processes (i.e., not deterministic).

With this formulation, it is possible to relate the feedback noise variance to the various feedback parameters. For example, if the feedback noise is purely from linear quantization error, then

$$\frac{\sigma^2_{\mathrm{fb}}}{\sigma^2_{\hat{\mathbf{H}}_r}} \geq \frac{1}{2^{2b}} \tag{5-8}$$

where b is the number of quantizer bits for either of the real or the imaginary part of each of the channel matrix coeffients (there are 2*b* bits per complex channel coefficient). Since there are MN channel coefficients, the total number of bits for each estimate of CSI feedback is $B = 2MNb$ bits.

The equality in (5-8) corresponds to the case where the quantizer input (the estimated channel at the receiver) is uniformly distributed, covering all the quantization levels. For the Gaussian case, the proportionality coefficient is often set to 6. This means that $\sigma_{\text{fb}}^2 / \sigma_{\hat{\mathbf{H}}_r}^2 = 6/2^{2b}$ [5-22]. Now σ_{fb}^2 can be related to the feedback delay and the feedback throughput constraint, R_{fb} by

$$\frac{\sigma_{\text{fb}}^2}{\sigma_{\hat{\mathbf{H}}_r}^2} = \frac{6}{2^{2b}} = \frac{6}{2^{B/MN}} = \frac{6}{2^{T_{\text{fb}}R_{\text{fb}}/MN}} \tag{5-9}$$

As an alternative, based on the rate distortion method we can obtain an upper bound on the capacity (equivalently, lower bound on the required feedback bits) by using $\sigma_{\text{fb}}^2 / \sigma_{\hat{\mathbf{H}}_r}^2 \geq 1/2^{2b} = 1/2^{B/MN}$ [5-22]. In the simulations below, the performance is analyzed using each of the two quantization equations: linear quantizer and rate distortion.

As the VQ indexing bits (feedback codeword length, *B*) increases, the quantization error, σ_{fb}^2, decreases. However, with a feedback throughput constraint, it takes longer for the codeword to be fed back. Thus the feedback delay increases. This, in turn, given a fixed training interval, decreases the time available for data transmission. This relation is used below in simulations to obtain an optimal number of VQ index bits (via finding the optimal feedback delay time) that maximizes the theoretic/practicable capacity.

With channel knowledge at the transmitter, improvement of the capacity in the data transmission is now the focus. A transmit filter, or beamformer, comprising weights $\mathbf{W}$ (not to be confused with the white noise notation above,

denoted by a subscripted $\mathbf{W}$, such as $\mathbf{W}_{\text{fb}}$) derived using channel knowledge, is shown in Figure 5-1.

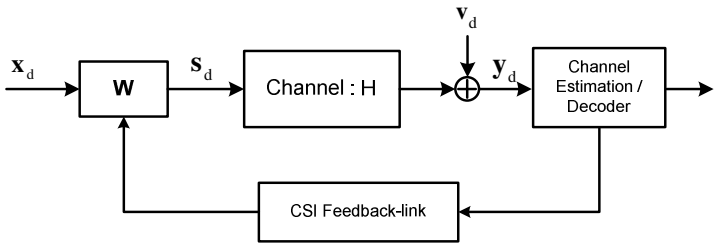

Figure 5-1 A closed-loop MIMO system with a transmit beamformer for using channel knowledge which is fed back from the receiver

At each symbol time, the $M \times 1$ data vector signal to be transmitted, $\mathbf{x}_{\text{d}}$, is multiplied by the $M \times M$ weight matrix $\mathbf{W}$, before transmission. From $\mathbf{x}_{\text{d}}$ satisfying $E\,\mathbf{x}_{\text{d}}\mathbf{x}_{\text{d}}^{H} = \mathbf{I}_{M}$ and $\mathbf{s}_{\text{d}} = \mathbf{W}\mathbf{x}_{\text{d}}$, the covariance matrix of the transmitted signal $\mathbf{R}_{\text{ss}} = E\,\mathbf{s}_{\text{d}}\mathbf{s}_{\text{d}}^{H}$ is given by

$$\mathbf{R}_{\text{ss}} = \mathbf{W}\mathbf{W}^{H}. \tag{5-10}$$

Thus, the input-output relation for the MIMO system in the data transmission can be written

$$\mathbf{Y}_{\text{d}} = \sqrt{\frac{\rho_{\text{d}}}{M}}\,\mathbf{H}\mathbf{W}\mathbf{X}_{\text{d}} + \mathbf{V}_{\text{d}} = \sqrt{\frac{\rho_{\text{d}}}{M}}\,\mathbf{H}\mathbf{S}_{\text{d}} + \mathbf{V}_{\text{d}} \tag{5-11}$$

where: $\mathbf{X}_{\text{d}}$ is the matrix of data symbols to be sent over T_{d} symbol durations with power constraint $E\,\text{tr}\,\mathbf{X}_{\text{d}}\mathbf{X}_{\text{d}}^{H} = T_{\text{d}}$; matrix $\mathbf{V}_{\text{d}}$ represents the additive noise with $\mathbf{V}_{\text{d}} \sim CN(0,1)$; and ρ_{d} is the SNR for the data transmission.

Similarly, $\mathbf{W}$ satisfies $\|\mathbf{W}\|_F^2 = \mathrm{tr}\,\mathbf{R}_{\mathbf{ss}} = M$. As a result, with $\mathbf{S}_\mathrm{d} = \mathbf{W}\mathbf{X}_\mathrm{d}$, we have $E\,\mathrm{tr}\,\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H = MT_\mathrm{d}$. Moreover, the cycle starting from the training transmission, including the channel estimation, quantization, codeword feedback, and calculation of the new beamformer, and detection of the received symbols is to be completed within T symbol durations. Thus,

$$T_\tau + T_\mathrm{fb} + T_\mathrm{d} = T \tag{5-12}$$

We note that T_d, T_τ, and T_fb are all integers which denote a time in symbol durations. From a time-partitioning of the energy, an average (over the training and data transmission) SNR at each receiver branch is

$$\rho = \frac{\rho_\tau T_\tau + \rho_\mathrm{d} T_\mathrm{d}}{T - T_\mathrm{fb}} \tag{5-13}$$

Given ρ and T, values are sought for the SNR during training and data transmission (ρ_τ and ρ_d, respectively), and the optimum number of training and data intervals (T_τ and T_d, respectively) that maximize the information theoretic/practicable capacity, under the feedback throughput constraint, R_fb. Furthermore, joint optimization over the number of training symbol and the feedback delay time can be undertaken by varying B.

5.2 Information-Theoretic/Practicable Capacity

First, the information-theoretic capacity and the practicable capacity are formulated in subsections 5.2.1 and 5.2.2, respectively. Then, the capacities are maximized over various parameters such as the training and data power

allocation, the number of training symbols, and the feedback delay length interval in subsection 5.2.3.

Although the channel is stochastic, the capacity of a sample realization of the channel is treated. Finally, the ergodic capacity is obtained as the ensemble average of the capacity achieved when the optimization is performed for each realization of $\mathbf{H}$, i.e. $C = E\, C_{\mathbf{H}}$.

5.2.1 Information-Theoretic Capacity

The information theoretic capacity of the MIMO system (expressed in bits per sec/Hz for comparison) with a fixed channel $\mathbf{H}$ is obtained by maximizing the mutual information, denoted below with the help of the usual *I*(.), between the known and received signals ($\mathbf{S}_\tau, \mathbf{Y}_\tau, \mathbf{Y}_\mathrm{d}$) and the unknown transmitted data signal, $\mathbf{S}_\mathrm{d}$, over the distribution of the transmit signal, denoted $p\mathbf{S}_\mathrm{d}(.)$, with $E\,\mathrm{tr}\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H = MT_\mathrm{d}$ and feedback-link throughput constraint R_fb. This can be written as

$$\begin{aligned} C_{\mathbf{H}} &= \max_{p\mathbf{S}_\mathrm{d}(.),\, E\ \mathrm{tr}\,\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H = MT_\mathrm{d}} \frac{1}{T} I\left(\mathbf{S}_\tau, \mathbf{Y}_\tau, \mathbf{Y}_\mathrm{d}; \mathbf{S}_\mathrm{d} \middle| R_\mathrm{fb}\right) \\ &= \max_{p\mathbf{S}_\mathrm{d}(.),\, E\ \mathrm{tr}\,\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H = MT_\mathrm{d}} \frac{1}{T} I\left(\mathbf{Y}_\mathrm{d}; \mathbf{S}_\mathrm{d} \middle| \mathbf{S}_\tau, \mathbf{Y}_\tau, R_\mathrm{fb}\right) \end{aligned} \tag{5-14}$$

Using (5-7) in ((5-11),

$$\mathbf{Y}_\mathrm{d} = \sqrt{\frac{\rho_\mathrm{d}}{M}}\,\hat{\mathbf{H}}_\mathrm{t}\mathbf{S}_\mathrm{d} + \tilde{\mathbf{V}}_\mathrm{d} \tag{5-15}$$

where $$\tilde{\mathbf{V}}_\mathrm{d} = \sqrt{\frac{\rho_\mathrm{d}}{M}}\,\sigma_{\tilde{\mathbf{H}}_\mathrm{r}}\,\tilde{\mathbf{W}}_{\tilde{\mathbf{H}}_\mathrm{r}}\mathbf{S}_\mathrm{d} - \sqrt{\frac{\rho_\mathrm{d}}{M}}\,\sigma_\mathrm{fb}\,\mathbf{W}_\mathrm{fb}\mathbf{S}_\mathrm{d} + \mathbf{V}_\mathrm{d}$$

is the resultant noise matrix with the associated noise variance

$$\begin{aligned}\sigma^2_{\tilde{\mathbf{V}}_\mathrm{d}} &= \frac{1}{NT_\mathrm{d}}\mathrm{tr}\,E\,\tilde{\mathbf{V}}_\mathrm{d}\tilde{\mathbf{V}}_\mathrm{d}^H = \frac{1}{NT_\mathrm{d}}E\,\mathrm{tr}\left(\frac{\rho_\mathrm{d}}{M}\sigma^2_{\tilde{\mathbf{H}}_\mathrm{r}}\tilde{\mathbf{W}}_{\tilde{\mathbf{H}}_\mathrm{r}}\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H\tilde{\mathbf{W}}^H_{\tilde{\mathbf{H}}_\mathrm{r}}\right)\\ &+\frac{1}{NT_\mathrm{d}}E\,\mathrm{tr}\left(\frac{\rho_\mathrm{d}}{M}\sigma^2_{\mathrm{fb}}\mathbf{W}_{\mathrm{fb}}\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H\mathbf{W}^H_{\mathrm{fb}}\right)+\frac{1}{NT_\mathrm{d}}E\,\mathrm{tr}\mathbf{V}_\mathrm{d}\mathbf{V}_\mathrm{d}^H\\ &=\frac{1}{NT_\mathrm{d}}\mathrm{tr}\left(\frac{\rho_\mathrm{d}}{M}\sigma^2_{\tilde{\mathbf{H}}_\mathrm{r}}E(\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H)E(\tilde{\mathbf{W}}^H_{\tilde{\mathbf{H}}_\mathrm{r}}\tilde{\mathbf{W}}_{\tilde{\mathbf{H}}_\mathrm{r}})\right)\\ &+\frac{1}{NT_\mathrm{d}}\mathrm{tr}\left(\frac{\rho_\mathrm{d}}{M}\sigma^2_{\mathrm{fb}}E(\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H)E(\mathbf{W}^H_{\mathrm{fb}}\mathbf{W}_{\mathrm{fb}})\right)+1\end{aligned} \tag{5-16}$$

Finally, since $E\,\mathrm{tr}\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H = MT_\mathrm{d}$, (5-16) reduces to

$$\sigma^2_{\tilde{\mathbf{V}}_\mathrm{d}} = 1+\rho_\mathrm{d}\left(\sigma^2_{\tilde{\mathbf{H}}_\mathrm{r}}+\sigma^2_{\mathrm{fb}}\right) \tag{5-17}$$

and the capacity in ((5-14) becomes [5.14], [5.16]

$$\begin{aligned}C_{\mathbf{H}} &= \min_{p\tilde{\mathbf{V}}_\mathrm{d}(.)}\ \max_{p\mathbf{S}_\mathrm{d}(.),\ E\ \mathrm{tr}\mathbf{S}_\mathrm{d}\mathbf{S}_\mathrm{d}^H=MT_\mathrm{d}}\ \frac{1}{T}I\left(\mathbf{Y}_\mathrm{d};\mathbf{S}_\mathrm{d}\middle|\hat{\mathbf{H}}_\mathrm{t}\right)\\ &=\max_{E\ \mathrm{tr}\mathbf{s}_\mathrm{d}\mathbf{s}_\mathrm{d}^H=M}\ \frac{T_\mathrm{d}}{T}\log_2\det\left(\mathbf{I}_N+\frac{\rho_\mathrm{d}}{M\sigma^2_{\tilde{\mathbf{V}}_\mathrm{d}}}\hat{\mathbf{H}}_\mathrm{t}\mathbf{R}_{\mathbf{ss}}\hat{\mathbf{H}}_\mathrm{t}^H\right)\\ &=\max_{E\ \mathrm{tr}\mathbf{s}_\mathrm{d}\mathbf{s}_\mathrm{d}^H=M}\ \frac{T_\mathrm{d}}{T}\log_2\det\left(\mathbf{I}_N+\frac{\rho_\mathrm{d}\sigma^2_{\hat{\mathbf{H}}_\mathrm{t}}}{M\sigma^2_{\tilde{\mathbf{V}}_\mathrm{d}}}\overline{\mathbf{H}}_\mathrm{t}\mathbf{R}_{\mathbf{ss}}\overline{\mathbf{H}}_\mathrm{t}^H\right)\end{aligned} \tag{5-18}$$

where $\overline{\mathbf{H}}_\mathrm{t} = \hat{\mathbf{H}}_\mathrm{t}/\sigma_{\hat{\mathbf{H}}_\mathrm{t}}$ is the normalized channel estimate at the transmitter with $\overline{\mathbf{H}}_\mathrm{t} \sim CN(0,1)$. Applying singular value decomposition, the matrix $\overline{\mathbf{H}}_\mathrm{t}$ with rank $r=\min(M,N)$ can be rewritten as $\overline{\mathbf{H}}_\mathrm{t} = \overline{\mathbf{U}}_\mathrm{t}\overline{\mathbf{\Lambda}}_\mathrm{t}^{1/2}\overline{\mathbf{V}}_\mathrm{t}^H$ where $\overline{\mathbf{V}}_\mathrm{t}$ and $\overline{\mathbf{U}}_\mathrm{t}$ are the unitary matrices containing the corresponding input and output singular vectors, respectively and $\overline{\mathbf{\Lambda}}_\mathrm{t}^{1/2}$ is a non-negative $N\times M$ diagonal matrix with *i*th diagonal element as $\overline{\lambda}_i^{1/2}$ (the square root of *i*th eigenvalue). In addition the diagonal elements satisfy $\overline{\lambda}_i \geq \overline{\lambda}_{i+1}$. As a result, $\overline{\mathbf{H}}_\mathrm{t}^H\overline{\mathbf{H}}_\mathrm{t} = \overline{\mathbf{V}}_\mathrm{t}\overline{\Lambda}_\mathrm{t}\overline{\mathbf{V}}_\mathrm{t}^H$, and

$$C_{\mathbf{H}} = \max_{E\ \mathrm{tr}\,\mathbf{s}_{\mathrm{d}}\mathbf{s}_{\mathrm{d}}^{H}=M} \frac{T_{\mathrm{d}}}{T}\log_2 \det\left(\mathbf{I}_N + \frac{\rho_{eff}}{M}\overline{\mathbf{\Lambda}}_{\mathrm{t}}^{1/2}\overline{\mathbf{V}}_{\mathrm{t}}^{H}\mathbf{R}_{\mathbf{ss}}\overline{\mathbf{V}}_{\mathrm{t}}\overline{\mathbf{\Lambda}}_{\mathrm{t}}^{1/2}\right) \tag{5-19}$$

where $\rho_{eff} = \rho_{\mathrm{d}}\sigma^2_{\hat{\mathbf{H}}_{\mathrm{t}}} / \sigma^2_{\tilde{\mathbf{V}}_{\mathrm{d}}}$ is an effective SNR, and note that this is averaged over the channel realizations, i.e., it does not depend on the specific channel matrix realization (whereas, maximizations of (5-19) relate to a particular channel realization). Also, note that $\tilde{\mathbf{R}}_{\mathbf{ss}} = \overline{\mathbf{V}}_{\mathrm{t}}^{H}\mathbf{R}_{\mathbf{ss}}\overline{\mathbf{V}}_{\mathrm{t}}$ is non-negative definite because $\mathbf{R}_{\mathbf{ss}}$ is, and that $\mathrm{tr}\,\tilde{\mathbf{R}}_{\mathbf{ss}} = \mathrm{tr}\,\mathbf{R}_{\mathbf{ss}}$. As a result, the maximization over $\mathbf{R}_{\mathbf{ss}}$ with $E\,\mathrm{tr}\,\mathbf{s}_{\mathrm{d}}\mathbf{s}_{\mathrm{d}}^{H} = M$, can be over $\tilde{\mathbf{R}}_{\mathbf{ss}}$ with $E\,\mathrm{tr}\,\tilde{\mathbf{s}}_{\mathrm{d}}\tilde{\mathbf{s}}_{\mathrm{d}}^{H} = M$. Since for any non-negative definite matrix $\mathbf{A}$, $\det\mathbf{A} \le \prod_i [\mathbf{A}]_{i,i}$,

$$\det\left(\mathbf{I}_N + \rho_{eff}\overline{\mathbf{\Lambda}}_{\mathrm{t}}^{1/2}\tilde{\mathbf{R}}_{\mathbf{ss}}\overline{\mathbf{\Lambda}}_{\mathrm{t}}^{1/2}\right) \le \prod_i\left(1 + \frac{\rho_{eff}}{M}[\tilde{\mathbf{R}}_{\mathbf{ss}}]_{i,i}\overline{\lambda}_i\right) \tag{5-20}$$

with the equality for when $\tilde{\mathbf{R}}_{\mathbf{ss}}$ is diagonal. Thus,

$$C_{\mathbf{H}} = \max_{E\ \mathrm{tr}\,\tilde{\mathbf{s}}_{\mathrm{d}}\tilde{\mathbf{s}}_{\mathrm{d}}^{H}=M} \frac{T_{\mathrm{d}}}{T}\sum_{i=1}^{r}\log_2\left(1 + \frac{\rho_{eff}}{M}[\tilde{\mathbf{R}}_{\mathbf{ss}}]_{i,i}\overline{\lambda}_i\right) \tag{5-21}$$

The optimal diagonal entries of $\tilde{\mathbf{R}}_{\mathbf{ss}}$ can be found via water-filling [5.2], i.e.,

$$p_i = [\tilde{\mathbf{R}}_{\mathbf{ss}}]_{i,i}^{\mathrm{opt}} = \left(\mu - \frac{M}{\rho_{eff}\overline{\lambda}_i}\right)^{+};\ i = 1,...,r \qquad , \qquad \sum_{i=1}^{r} p_i = M \tag{5-22}$$

where $(x)^{+}$ implies $(x)^{+} = \begin{cases} x & ;\ x \ge 0 \\ 0 & ;\ x < 0 \end{cases}$ and μ is a constant calculated iteratively as

$$\mu = \frac{M}{r-q+1}\left(1 + \frac{1}{\rho_{eff}}\sum_{i=1}^{r-q+1}\frac{1}{\overline{\lambda}_i}\right) \tag{5-23}$$

with the iteration count q, starting from 1. If the power allocated to the eigenchannel with the lowest gain is negative, i.e. $\left[\tilde{\mathbf{R}}_{\mathbf{ss}}\right]_{r-q+1} < 0$, this eigenchannel is discarded by setting $\left[\tilde{\mathbf{R}}_{\mathbf{ss}}\right]_{r-q+1}^{opt} = 0$ and μ is again calculated with the iteration count q incremented by 1. Once $\tilde{\mathbf{R}}_{\mathbf{ss}}^{opt}$ is found, an optimal beamformer $\mathbf{W}^{\text{opt}}$ is obtained from ((5-10) and $\mathbf{R}_{\mathbf{ss}}^{\text{opt}} = \overline{\mathbf{V}}_{\text{t}} \tilde{\mathbf{R}}_{\mathbf{ss}}^{\text{opt}} \overline{\mathbf{V}}_{\text{t}}^{H}$ as

$$\mathbf{W}^{\text{opt}} = \overline{\mathbf{V}}_{\text{t}} \left(\tilde{\mathbf{R}}_{\mathbf{ss}}^{\text{opt}}\right)^{1/2} \tag{5-24}$$

Now the capacity (5-21) can be rewritten as

$$C_{\mathbf{H}} = \frac{T_{\text{d}}}{T} \sum_{i=1}^{r-q+1} \log_2\left(1 + \frac{\rho_{eff}}{M} p_i \overline{\lambda}_i\right) \tag{5-25}$$

Therefore, the information-theoretic capacity of MIMO system is obtained by summing the capacity of each SISO eigenchannel which is governed by the SNR imposed on each eigenchannel (the SNR at the receiver of the ith eigenchannel is $(\rho_{eff}/M) p_i \overline{\lambda}_i$). The remaining task is to maximize (5-25) with respect to various parameters such as ρ_τ, ρ_{d}, T_τ, T_{d}, and T_{fb}. This is presented in subsection 5.2.3.

5.2.2 Practicable Capacity

The formulation of (5-25) requires many assumptions including the need for independent Gaussian signals at the input and infinitely long data codes. However, the practicable capacity offers a MIMO performance limit which accounts for the use of signal processing. This specifically includes digital

modulation and finite block coding. The impact of finite block length and digital modulation (e.g., QAM signaling) can be evaluated as follows. For uncoded D-QAM where D denotes the number of points in each signal constellation, the attainable normalized throughput (practicable capacity) in bit per sec/Hz for a SISO channel can be given in terms of the block error rate (BLER) for block length *BL* as $R = \log_2 \mathrm{D}.\left[1 - BLER(SNR)\right] = \log_2 \mathrm{D}.\left[1 - SER(SNR)\right]^{BL}$, *cf.*, [5.17], which uses BER instead of SER, although both forms turn out to yield the same result. Here, SER is the probability of symbol error (symbol-error rate) for an AWGN channel with D-QAM modulation and coherent detection. For example, for a uniform D-QAM, the SER is approximated as [5.21]

$$SER(SNR) \approx 4Q\left(\sqrt{\frac{3\,SNR}{\mathrm{D}-1}}\right) \tag{5-26}$$

with SNR is the signal-to-noise ratio per symbol.

The above practicable capacity, R, formula assumes perfect error detection, wherein blocks are correctly detected if and only if all bit decisions are error-free. Also, the formula assumes BL independent symbols per block, independent of the signal constellation. An alternative is to keep the number of bits per block fixed [5.17], but that would require using different block lengths for different substreams.

Following the SVD analysis, the MIMO channel is transformed into parallel SISO eigenchannels with unequal instantaneous SNRs (the SNR of the *i*th eigenchannel is $(\rho_{eff} / M) p_i \overline{\lambda}_i$). Therefore, the throughput (practicable capacity) of MIMO system (with a fixed channel $\mathbf{H}$) is the sum of the throughputs

corresponding to eigenchannels. For example, for a system with no adaptive modulation and no coding, the throughput with uniform D-QAM signaling is

$$R_{\mathbf{H}} = \sum_{i=1}^{r-q+1} R_i = \frac{T_{\mathrm{d}}}{T} \sum_{i=1}^{r-q+1} \log_2 \mathrm{D}.\left[1 - SER((\rho_{eff} / M) p_i \overline{\lambda_i})\right]^{BL}$$
$$= \frac{T_{\mathrm{d}}}{T} \sum_{i=1}^{r-q+1} \log_2 \mathrm{D}.\left[1 - 4Q\left(\sqrt{\frac{3(\rho_{eff} / M) p_i \overline{\lambda_i}}{\mathrm{D}-1}}\right)\right]^{BL} \qquad (5\text{-}27)$$

with $\rho_{eff} = \rho_{\mathrm{d}} \sigma^2_{\hat{\mathbf{H}}_{\mathrm{t}}} / \sigma^2_{\tilde{\mathbf{v}}_{\mathrm{d}}}$. Either non-adaptive or adaptive modulation can be used here. In non-adaptive modulation one type of QAM is used for all eigenchannel irrespective to the SNR on the eigenchannels, whereas in the adaptive modulation the QAM may be chosen for each eigenchannel in a standard optimum way so that the corresponding eigenchannel throughput (practicable capacity) remain as close as possible to the Shannon limit of SISO channel. The 8 dB minimum capacity penalty resulting from using finite block length and uncoded QAM with adaptive modulation over a SISO AWGN channel is reported in [5.17]-[5.18] and it is helpful for the discussion below to include this here (Figure 5-2). Some elaboration is useful here, and the following paragraph is from [5.22]. In the flat, saturated region of the curves, the throughput is insensitive to the block size. However, in the steep dropoff region, the curves are very sensitive to the block size. The dropoff region – especially the corner, where throughput begins to fall away – is critical knowledge for any designer. This is analysed as follows. Defining the corner as 10% BLER (90% chance of the block being receiver error-free), the corner then corresponds to $\lim_{BL\to\infty}(1-\underbrace{SER}_{\equiv x/BL})^{BL} \approx \lim_{BL\to\infty}(1-\frac{x}{BL})^{BL} = e^{-x} = 0.1$ or $x \approx 0.1$. Thus, for any large value of

block length, we need $SER < 0.1/BL$ or $BL < 0.1/SER$. As an example, for SER of 10^{-4}, then the block length can not exceed 1000 symbols.

Table 5-1 illustrates the QAM selection respect to the SNR imposed on each eigenchannel and Figure 5-2 for the block size of $BL = 100$ symbols. This is the most efficient digital technique map for the MIMO system.

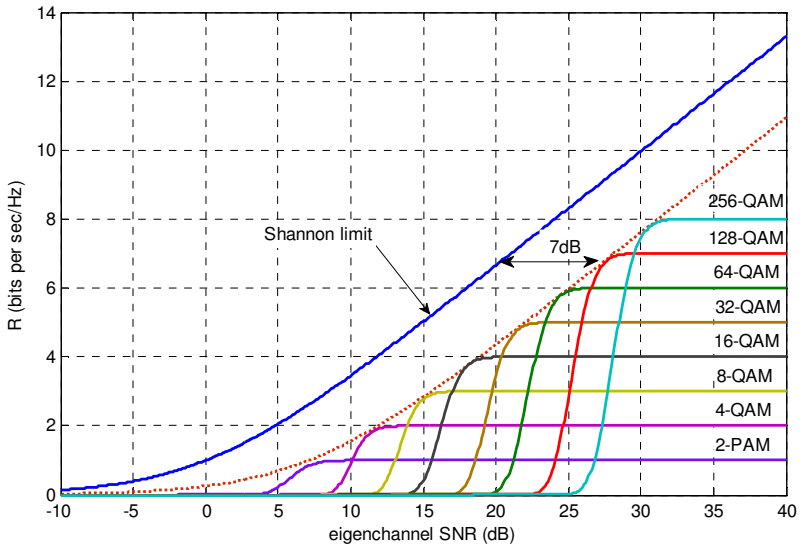

Figure 5-2 Single AWGN channel practicable capacity (throughput) penalties for D-QAM versus SNR for a block size of 100 symbols

Table 5-1 The digital adaptive modulation selection strategy

Modulation Selection (D_i -QAM)	**SNR on *i*th eigenchannel ($SNR_i = (\rho_{eff}/M)p_i\overline{\lambda}_i$**
256-QAM	$29.6\,\text{dB} \le SNR_i$
128-QAM	$26.5\,\text{dB} \le SNR_i < 29.6\,\text{dB}$
64-QAM	$23.4\text{dB} \le SNR_i < 26.5\,\text{dB}$
32-QAM	$20.4\,\text{dB} \le SNR_i < 23.4\,\text{dB}$
16-QAM	$17\,\text{dB} \le SNR_i < 20.4\,\text{dB}$
8-QAM	$13.9\,\text{dB} \le SNR_i < 17\,\text{dB}$
4-QAM	$10.1\,\text{dB} \le SNR_i < 13.9\,\text{dB}$
2-PAM	$SNR_i < 10.1\,\text{dB}$

Therefore, the throughput (practicable capacity) of MIMO system with adaptive modulation (with uniform D_i-QAM signaling on ith eigenchannel) is written as

$$\tilde{R}_{\mathbf{H}} = \sum_{i=1}^{r-q+1} R_i = \frac{T_{\mathrm{d}}}{T} \sum_{i=1}^{r-q+1} \log_2 \mathrm{D}_i \cdot \left[1 - 4Q\left(\sqrt{\frac{3(\rho_{eff}/M) p_i \bar{\lambda}_i}{\mathrm{D}_i - 1}} \right) \right]^{BL} \tag{5-28}$$

Ideally, adaptive modulation allows a wide range of SNRs to bear a good throughput, but it is not easy to implement, and not effective in fast-changing channels. (Of course, in this sense, SVD is similarly unsuited to fast-fading channels.) The reason is that, in practice, it is very difficult (large observation interval required) to estimate the SNR on the eigenchannels with accuracy of decimal points of a dB, nevertheless for limiting capacity calculations ideal estimation of SNR is assumed. On the other hand, non-adaptive modulation has lower complexity compared to the adaptive modulation, but the SNR needs constraining for good throughput. The curves of Figure 5-2 show a zero-capacity region for each modulation, for example, 64-QAM, the zero-capacity region is for a SNR below about 20 dB. This shape of curve also appears in the simulations for the eigen-MIMO channels, below.

All that remains is maximizing (5-27) and (5-28) with respect to various parameters such as ρ_τ, ρ_{d}, T_τ, T_{d}, and T_{fb}. Both adaptive and non-adaptive modulations are considered.

5.2.3 Maximizing the Capacities

The first parameter for optimization is the choice of the training signal $\mathbf{S}_\tau$ which affects the capacities $C_{\mathbf{H}}$ in (5-24), $R_{\mathbf{H}}$ in (5-27), and $\tilde{R}_{\mathbf{H}}$ in (5-28) through the effective SNR, ρ_{eff}. Since all the three capacities are increasing functions of ρ_{eff}, it can be shown, following the same procedure as in [5.14] for open-loop systems, that the optimal training signal is a multiple of a matrix with orthonormal columns. With this choice of training, there results [5.23],

$$\sigma^2_{\tilde{\mathbf{H}}_r} = \frac{1}{1+\frac{\rho_\tau}{M}T_\tau} \qquad \text{and} \qquad \sigma^2_{\hat{\mathbf{H}}_r} = \frac{\frac{\rho_\tau}{M}T_\tau}{1+\frac{\rho_\tau}{M}T_\tau} \tag{5-29}$$

Now, from ((5-6), (5-17) and (5-29), the effective SNR becomes

$$\begin{aligned}\rho_{eff} &= \frac{\rho_{\mathrm{d}}\sigma^2_{\hat{\mathbf{H}}_t}}{\sigma^2_{\tilde{\mathbf{V}}_{\mathrm{d}}}} = \frac{\rho_{\mathrm{d}}\left(1-\sigma^2_{\tilde{\mathbf{H}}_r}+\sigma^2_{\mathrm{fb}}\right)}{1+\rho_{\mathrm{d}}\left(\sigma^2_{\tilde{\mathbf{H}}_r}+\sigma^2_{\mathrm{fb}}\right)} \\ &= \frac{\rho_{\mathrm{d}}\rho_\tau T_\tau + \rho_{\mathrm{d}}\sigma^2_{\mathrm{fb}}\left(M+\rho_\tau T\right)}{M+\rho_{\mathrm{d}}M+\rho_\tau T_\tau+\rho_{\mathrm{d}}\sigma^2_{\mathrm{fb}}\left(M+\rho_\tau T\right)}\end{aligned} \tag{5-30}$$

which makes explicit its dependence on ρ_τ and ρ_{d}.

Since the allocation of power to the training and data affect all the three capacities $C_{\mathbf{H}}$, $R_{\mathbf{H}}$, and $\tilde{R}_{\mathbf{H}}$ only through ρ_{eff}, with capacities an increasing function of ρ_{eff}, the optimal power allocation is chosen to maximize ρ_{eff}. Denoting α as the fraction of the total transmit energy that is devoted to the data, i.e.,

$$\rho_{\mathrm{d}}T_{\mathrm{d}} = \alpha\rho(T-T_{\mathrm{fb}}) \tag{5-31}$$

and

$$\rho_\tau T_\tau = (1-\alpha)\rho(T-T_{\rm fb})\ ;\ 0<\alpha<1\,,$$

giving

$$\rho_{eff} = \frac{-\left(\rho(T-T_{\rm fb})(1+\sigma_{\rm fb}^2)\right)\alpha^2 + \left(\rho(T-T_{\rm fb}) + \sigma_{\rm fb}^2(M+\rho(T-T_{\rm fb}))\right)\alpha}{-\left(\rho(T-T_{\rm fb})\sigma_{\rm fb}^2\right)\alpha^2 + \left(M - T_{\rm d} + \sigma_{\rm fb}^2(M+\rho(T-T_{\rm fb}))\right)\alpha + \dfrac{MT_{\rm d}}{\rho(T-T_{\rm fb})} + T_{\rm d}} \qquad (5\text{-}32)$$

Differentiating ρ_{eff} with respect to α and equating to zero, yields the optimal value for α as

$$\alpha^{\rm opt} = \begin{cases} \gamma - \beta & ;\quad T_{\rm d} > \dfrac{M(1+2\sigma_{\rm fb}^2)}{1+\sigma_{\rm fb}^2} \\ \dfrac{\rho(T-T_{\rm fb})(1+\sigma_{\rm fb}^2) + M\sigma_{\rm fb}^2}{2\rho(T-T_{\rm fb})(1+\sigma_{\rm fb}^2)} & ;\quad T_{\rm d} = \dfrac{M(1+2\sigma_{\rm fb}^2)}{1+\sigma_{\rm fb}^2} \\ \gamma + \beta & ;\quad T_{\rm d} < \dfrac{M(1+2\sigma_{\rm fb}^2)}{1+\sigma_{\rm fb}^2} \end{cases} \qquad (5\text{-}33)$$

with

$$\gamma = \frac{(\rho(T-T_{\rm fb})+M)(1+\sigma_{\rm fb}^2)\,T_{\rm d}}{\rho(T-T_{\rm fb})\,(T_{\rm d} - M + \sigma_{\rm fb}^2(T_{\rm d}-2M))}$$

$$\beta = \frac{\sqrt{MT_{\rm d}(\rho(T-T_{\rm fb})+M)(\rho(T-T_{\rm fb})(1+\sigma_{\rm fb}^2)(1+2\sigma_{\rm fb}^2) + \sigma_{\rm fb}^2(M+T_{\rm d}) + 2\sigma_{\rm fb}^4 T_{\rm d})}}{\rho(T-T_{\rm fb})\,(T_{\rm d} - M + \sigma_{\rm fb}^2(T_{\rm d}-2M))}$$

Thus the optimal training and data power allocation, $\rho_\tau^{\rm opt} = (1-\alpha^{\rm opt})\rho(T-T_{\rm fb})/T_\tau$ and $\rho_{\rm d}^{\rm opt} = \alpha^{opt}\rho(T-T_{\rm fb})/T_{\rm d}$, respectively, are obtained as a functions of various parameters such as T_τ, $T_{\rm d}$, and $T_{\rm fb}$. However, these parameters affect the capacities $C_{\mathbf{H}}$, $R_{\mathbf{H}}$, and $\tilde{R}_{\mathbf{H}}$ in different ways.

For example, on one hand, as the number of training symbols, T_τ, is increased, the estimate of the channel improves. This means that $\sigma^2_{\tilde{\mathbf{H}}_\mathrm{r}}$ decreases (from (5-29)) and ρ_{eff} increases since

$$\rho_{eff} = \frac{\rho_\mathrm{d}(1-\sigma^2_{\tilde{\mathbf{H}}_\mathrm{r}}+\sigma^2_\mathrm{fb})}{1+\rho_\mathrm{d}(\sigma^2_{\tilde{\mathbf{H}}_\mathrm{r}}+\sigma^2_\mathrm{fb})} = \frac{1+\rho_d(1+2\sigma^2_\mathrm{fb})}{1+\rho_\mathrm{d}(\sigma^2_{\tilde{\mathbf{H}}_\mathrm{r}}+\sigma^2_\mathrm{fb})} - 1 \qquad (5\text{-}34)$$

This results in an increase in capacities through ρ_{eff}. On the other hand, as T_τ increases, the time available to transmit data, T_d, decreases. Thus the capacities decrease linearly through the term T_d/T.

In addition, the feedback link parameters (R_fb, T_fb, and σ^2_fb) affect the capacities in different ways. These parameters not only affect the capacities $C_\mathbf{H}$, $R_\mathbf{H}$, and $\tilde{R}_\mathbf{H}$ through pre-log coefficient, but also influence the capacities via SNR. As the feedback delay T_fb increases, the capacities decrease linearly through the term $(T-T_\mathrm{fb}-T_\tau)/T$ for a given T_τ. However, the feedback noise variance decreases since $\sigma^2_{\hat{\mathbf{H}}_\mathrm{r}} \approx 1$ (the channel variance in $\sigma^2_{\hat{\mathbf{H}}_\mathrm{r}} = (\frac{\rho_\tau}{M}T)/(1+\frac{\rho_\tau}{M}T)$ can be approximated to 1 since $T_\tau \geq M$ and $\rho_\tau > 1$ for typical wireless applications, e.g., $\rho_\tau > 5\mathrm{dB}$. Thus, $\rho_\tau T_\tau / M >> 1$), the formula of (5-9) reduces to $\sigma^2_\mathrm{fb} \approx 6/2^{T_\mathrm{fb}R_\mathrm{fb}/MN}$. This increases ρ_{eff}, seen mathematically by rewriting (5-30) as

$$\rho_{eff} = \frac{\rho_\mathrm{d}\rho_\tau T_\tau - M + \rho_\mathrm{d}M + \rho_\tau T_\tau}{M + \rho_\mathrm{d}M + \rho_\tau T_\tau + \rho_\mathrm{d}\sigma^2_\mathrm{fb}(M+\rho_\tau T)} + 1 \qquad (5\text{-}35)$$

and this in turn affects all the three capacities.

By fixing the feedback constraint R_{fb}, the feedback delay can be changed by varying the number of VQ index bits, B, see (5-8) and (5-9). It is shown below that for a moderate to high SNR, the capacities are sensitive to the feedback delay and further increase in capacities are possible by joint optimizing over the number of training symbols and the feedback delay length interval.

However, since analytical solution is not available, the optimal values T_{τ}^{opt} and $T_{\text{fb}}^{\text{opt}}$ for each of the capacities are obtained via Monte-Carlo simulations. This requires searching over $S=\left\{(T_{\tau},T_{\text{fb}})\middle|\, T_{\tau}=M,...,T-T_{\text{fb}}-1;T_{\text{fb}}=1,...,T-M-1\right\}$ with set size $|S|$ to see which pair candidate $(T_{\tau},T_{\text{fb}})^{k}\in S;k=1,...,|S|$ maximizes the capacity.

Denoting $C_{\mathbf{H}}^{k}$, $R_{\mathbf{H}}^{k}$ and $\tilde{R}_{\mathbf{H}}^{k}$ as the corresponding capacities evaluated at $(T_{\tau},T_{\text{fb}})^{k}$, the optimal pair associated with each capacity is obtained as

$$(T_{\tau}^{\text{opt}},T_{\text{fb}}^{\text{opt}})_{C_{\mathbf{H}}}=\arg\max_{(T_{\tau},T_{\text{fb}})^{k}\in S}C_{\mathbf{H}}^{k} \tag{5-36}$$

$$(T_{\tau}^{\text{opt}},T_{\text{fb}}^{\text{opt}})_{R_{\mathbf{H}}}=\arg\max_{(T_{\tau},T_{\text{fb}})^{k}\in S}R_{\mathbf{H}}^{k} \tag{5-37}$$

$$(T_{\tau}^{\text{opt}},T_{\text{fb}}^{\text{opt}})_{\tilde{R}_{\mathbf{H}}}=\arg\max_{(T_{\tau},T_{\text{fb}})^{k}\in S}\tilde{R}_{\mathbf{H}}^{k} \tag{5-38}$$

The Monte Carlo simulations conducted can be represented by the flowchart in Figure 5-3 below.

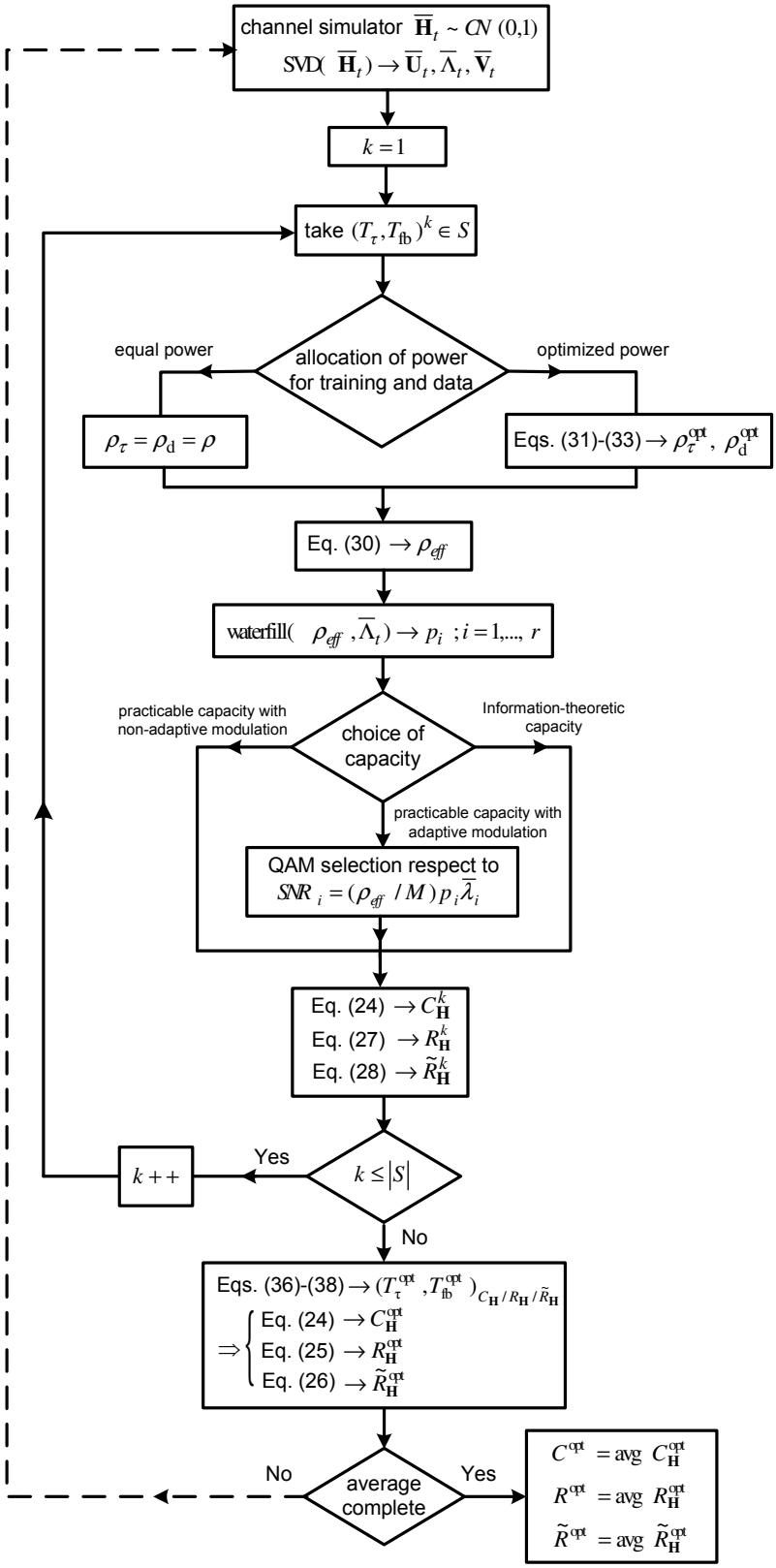

Figure 5-3 Flowchart of the simulation

5.3 Simulation Results

A communication system may not have the luxury of individual power adaptation for training and data phases and dedicate the same power to the training and data symbols for transmission. Thus, two cases are considered: systems with optimized allocation of power between training and data; and the systems with equal training and data power (an arbitrary choice).

The signaling rate is governed by the channel. For a delay spread σ_τ sec., the coherence bandwidth in Hertz is $\sigma_f \approx 1/(x\sigma_\tau)$, where: $x \geq 2\pi$ from the uncertainty principle (with the associated assumptions about the channel); and in practice, $x \approx 8$ (Gans' law); and for conservative design, $x = 10$ is often used. The symbol rate should be no more than this coherence bandwidth. Alternatively stated, the symbol duration should be least ten times the delay spread, in order to avoid intersymbol interference. The example considered in the remainder of the chapter can be related to a pedestrian in a cellular system. The delay spread of an outdoor channel is about $\sigma_\tau = 1\,\mu\text{sec}$, much larger than indoor values which are typically well below $100\,\text{n sec}$. The minimum symbol duration follows as $T_s = 10\sigma_\tau = 10\,\mu\text{sec}$, i.e., the maximum symbol rate is $100\,\text{ksym/sec}$. For a walking speed of a few km/hr and a carrier frequency of about 1 GHz, the channel is slowly varying, having a typical maximum Doppler frequency of $f_\text{D} = 10\,\text{Hz}$. The coherence time is conservatively taken to be a tenth of this maximum rate of change, i.e., $1/(10 f_\text{D})$ sec, and this corresponds to 1000 symbol times. This value

is the time that the channel is essentially constant and it is used for the block length from now on.

In the figures below, the solid lines correspond to capacity results with optimized power allocation and dashed lines correspond to the results with equal power allocation. Also, the lines with a “star” marker are the associated results based on (5-9) with linear quantizer and the lines with “circle” marker are the upper bounds on capacity based on rate distortion method. The simulations are for a 10×10 MIMO system, but the formulation is applicable to any $N\times M$ MIMO system.

At first, it is useful to include an illustration serving as a motivation for changing the feedback delay interval. Figure 5-4 and Figure 5-5 illustrate the ergodic capacities' behavior versus T_{fb} for the example of $R_{\mathrm{fb}}=25$ (i.e., the feedback link requires at least 25 bits/symbol time), $T=1000$ at moderate SNR, $\rho=18\,\mathrm{dB}$, using the optimal training interval. In each system, the capacity Is maximized at a particular feedback delay. For example, with a linear quantizer, the theoretic capacity is maximized at the feedback length of $T_{\mathrm{fb}}^{\mathrm{opt}}=56$ symbol times. Thus the total number of feedback bits is $B=T_{\mathrm{fb}}^{\mathrm{opt}}R_{\mathrm{fb}}=1400$ bits ($b=7$ bits for the real or the imaginary part of one channel coefficient). With rate distortion, the lower bound on the number of beedback bits is $B=T_{\mathrm{fb}}^{\mathrm{opt}}R_{\mathrm{fb}}=40\times25=1000$ bits (or $b=5$, that is less 2 bits per real or imaginary part of a channel coefficient).

The change in the feedback delay interval is rather insensitive at low SNR

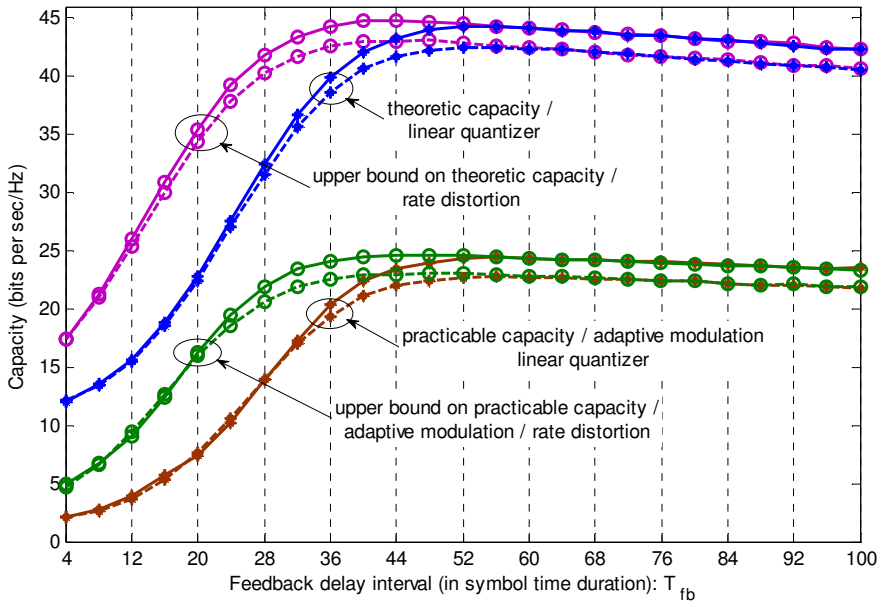

Figure 5-4 The effect of feedback delay T_{fb} on the theoretic/practicable capacities with $R_{\text{fb}} = 25$, and $T = 1000$.

The solid lines correspond to capacity results with optimized power allocation and dashed lines correspond to the results with equal power allocation. The lines with "star" marker are the associated results based on linear quantizer and the lines with "circle" marker are the upper bounds on capacity based on rate distortion method.

(not shown here), however, for a moderate or higher SNR, the capacities are sensitive to the feedback delay. This figure suggests that it is beneficial to change the number of feedback bits, $B = T_{\text{fb}}^{\text{opt}} R_{\text{fb}}$ (or quantization bits *b*), accordingly and motivates a joint optimization over the number of training symbols and the feedback delay length interval.

The resultant ergodic capacities (in bits per sec/Hz) for different values of SNR is illustrated in Figure 5-6 and Figure 5-7 for the example of $R_{\text{fb}} = 25$ and $T = 1000$. Here joint optimization over the training interval and the feedback delay is performed. It is emphasized here that the feedback link, running at 25 bits per

symbol, is not accounted for in the capacity of the forward link, which includes the lesser bit symbol sizes of 4-QAM and 16-QAM.

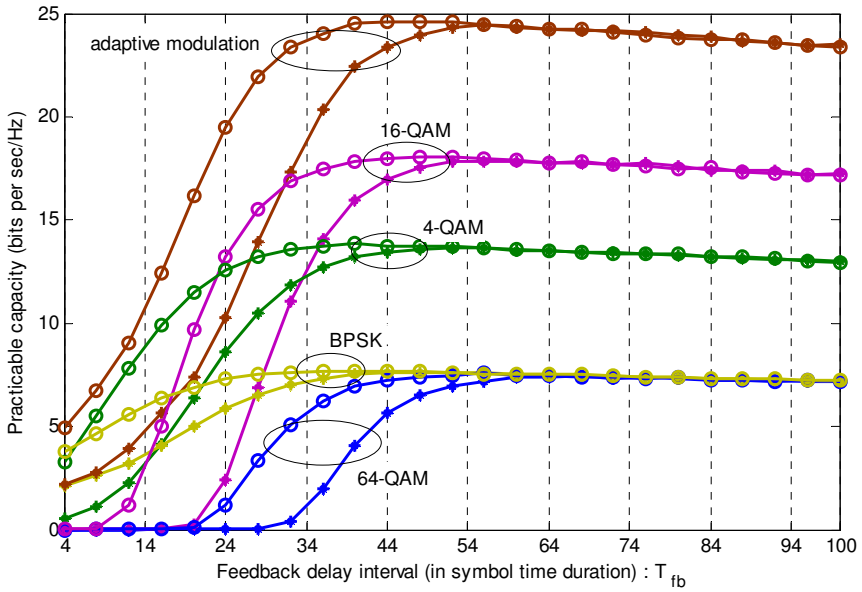

Figure 5-5 The effect of feedback delay T_{fb} on the practicable capacities with non-adaptive modulations and $R_{\text{fb}} = 25$, $T = 1000$.

The lines with "star" marker are the associated results based on linear quantizer and the lines with "circle" marker are the upper bounds on capacity based on rate distortion method. For every block of 1000 symbols using 64-QAM, about 64 symbol times at 25 bits per symbol are required for the feedback.

Figure 5-6 also includes the maximized theoretic capacity of an open-loop MIMO system, as given in [5.14]. For the closed loop case, the total time duration for training and data transmission is reduced to an effective value by the feedback delay. Similarly, the SNR is reduced to an effective SNR (ρ_{eff} - a quantity used only in the formulation) as the result of channel estimation error and feedback noise. Despite these degradations, the theoretic closed-loop capacities are higher than that of the open loop system for low to moderate

values of SNR ρ ~< 13 dB. However, there is no guarantee that a closed-loop system has higher theoretic capacity than an open-loop system for high values of SNR, and this is formulated here through the (generous) limitation of the feedback link. The situation is illustrated in Figure 5-6 for SNR values greater than ρ ~> 13 dB. Nevertheless, for wireless applications the SNR of interest is normally low, and here the closed-loop system has higher capacity.

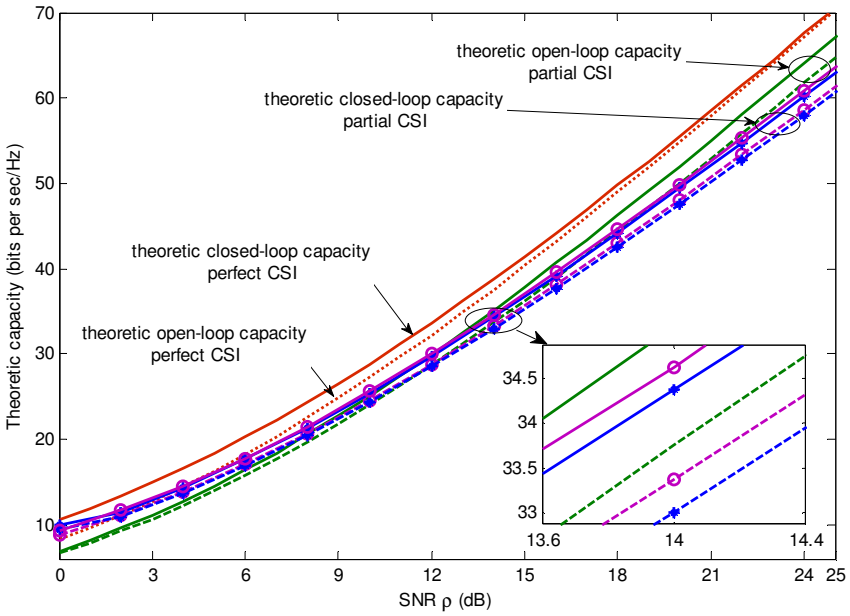

Figure 5-6 The theoretic capacities versus SNR for $R_{\text{fb}} = 25$, $T = 1000$.

The solid lines correspond to capacity results with optimized power allocation and dashed lines correspond to the results with equal power allocation. The lines with "star" marker are the associated results based on linear quantizer and the lines with "circle" marker are the upper bounds on capacity based on rate distortion method.

The practicable capacities with adaptive/non-adaptive modulation are included in Figure 5-7 with adaptive modulation and perfect CSI, the MIMO practicable capacity experiences similar capacity penalty as in SISO channel (the

MIMO practicable capacity is 8 dB away from the optimal theoretic capacity with perfect CSI). Furthermore, the shape of the practicable capacity curves with non-adaptive modulation follows the same trend as the curves of Figure 5-2, with cut-off and limiting capacity behavior for a given modulation.

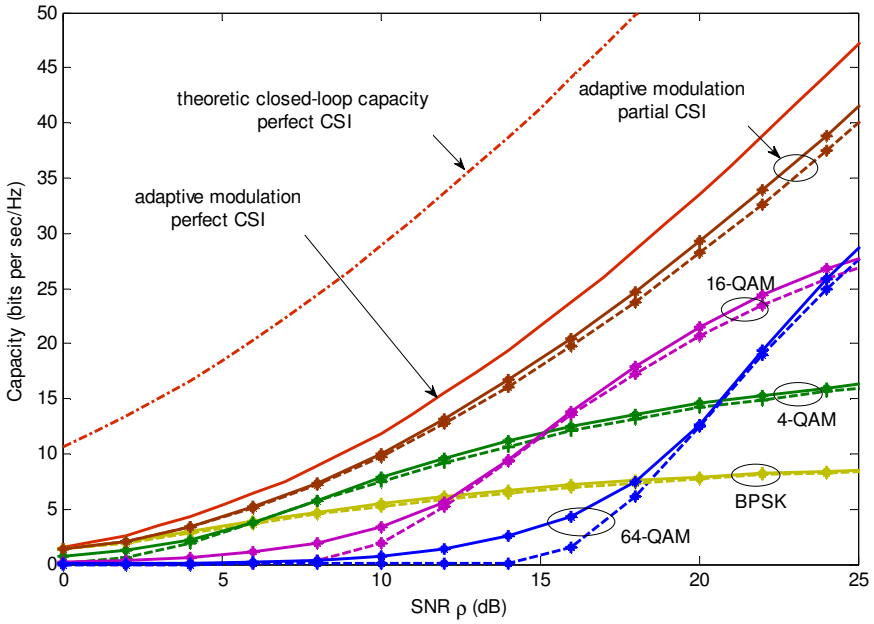

Figure 5-7 The practicable capacities versus SNR for $R_{\text{fb}} = 25$, $T = 1000$.

The solid lines correspond to capacity results with optimized power allocation and dashed lines correspond to the results with equal power allocation. The lines with "star" marker are the associated results based on linear quantizer.

An interesting result is that, with the theoretic capacity, the optimal training interval turns out to follow the same trend as the open-loop system. The optimal training interval is still equal to the number of transmit antennas when the optimal power allocation is used. With the equal power constraint, we get the optimized training time, $T_\tau^{opt} \geq M$. However, this is not the case with the practicable capacities, as laid out in Figure 5-8 and Figure 5-9, for different values of SNR.

Generally, as ρ increases, the T_τ^{opt} decreases to reach M and, for a given ρ, the optimal training interval increases ($\sigma^2_{\tilde{\mathbf{H}}_r}$ decreases) with the constellation size D. In particular, with the adaptive modulation, the optimal training interval is equal or greater than the optimal training interval associated with the theoretic capacity. Also, with the non-adaptive modulation, a very large number of training symbols is suggested specially at low to mid values of SNR and large constellation sizes where most of the eigenchannels are operated in the near-zero capacity region (*cf.*, Figure 5-2). It can be concluded that the logarithmic increase of capacity through ρ_{eff} is more effective than the linear decrease of capacity through T_d/T at low-mid SNRs. In particular, for a given constellation size D, T_τ^{opt} using the optimal power allocation is greater than T_τ^{opt} with the equal power allocation at low-mid SNRs.

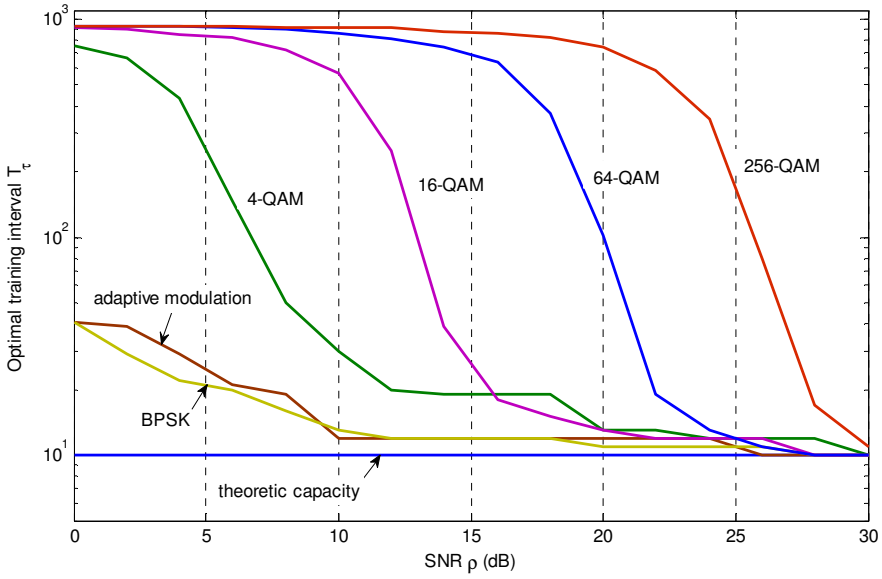

Figure 5-8 The optimal training interval versus SNR (dB) with optimal power allocation using linear quantizer for $R_{\mathrm{fb}} = 25$, $T = 1000$.

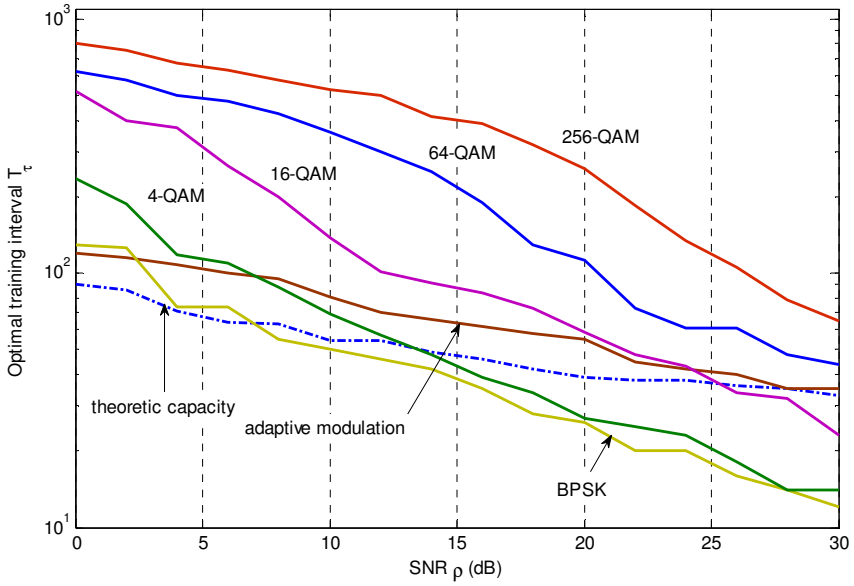

Figure 5-9 The optimal training interval versus SNR (dB) with equal power allocated for training and data transmission using linear quantizer for $R_{\text{fb}} = 25$, $T = 1000$.

However, at higher SNRs (where the eigenchannels are operated in the non-zero capacity region), more time is dedicated to the data payload transmission and T_τ^{opt} using the optimal power allocation would be smaller than that of with the equal power allocation.

From Figures 5-5 and 5-8, for each forward transmission block of 1000 symbols, for the case of say 18dB SNR, 64-QAM and optimal power allocation, there are about 64 symbols (at 25 bits per symbol) needed for the feedback, and in the forward link, about 300 symbols are needed for the training symbols. If the feedback and the forward link training are lumped into in a 1000 symbol block (they are actually treated separately in the above capacity formulation) , it leaves about 636 symbols for the payload.

Finally, there are some interesting points regarding the optimal number of feedback bits or equivalently the required number of bits per real (imaginary) part of one channel coefficient, b. The corresponding results based on the theoretic and the practicable capacities are illustrated in Figure 5-10 and Figure 5-11, respectively with the optimal power allocation. Figure 5-20 also includes the lower bounds on the required number of bits using the rate distortion method. We note that, based on the information given in these figures, the total number of feedback bits B and the feedback length T_{fb} can be calculated easily by $B = 200b$ and $T_{\text{fb}} = 200b / R_{\text{fb}} = 8b$.

In general, the required number of bits increases with SNR. Moreover, for a given SNR, the optimal feedback bits increases with the constellation size D

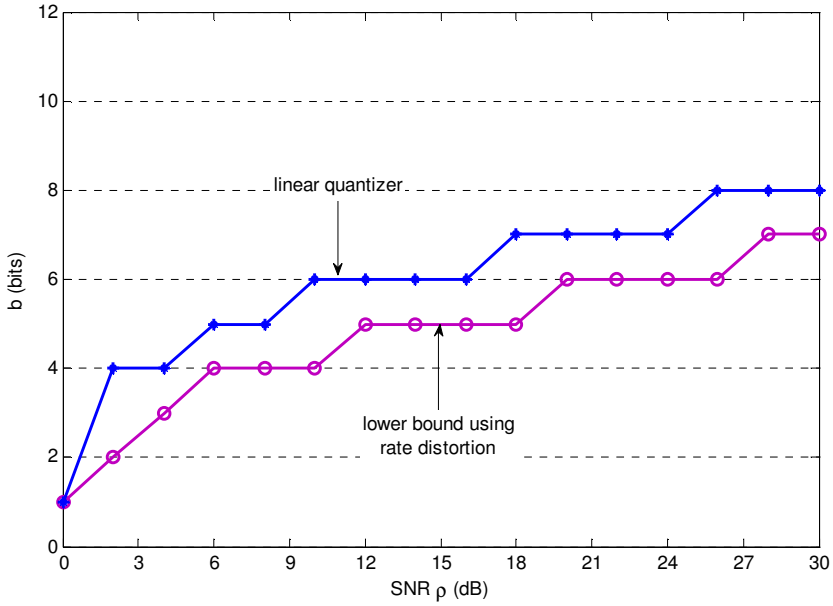

Figure 5-10 The optimal number of bits "*b*" (based on theoretic capacity) for different values of SNR (dB) with optimal power allocation using linear quantizer / rate distortion for $R_{\text{fb}} = 25$, $T = 1000$.

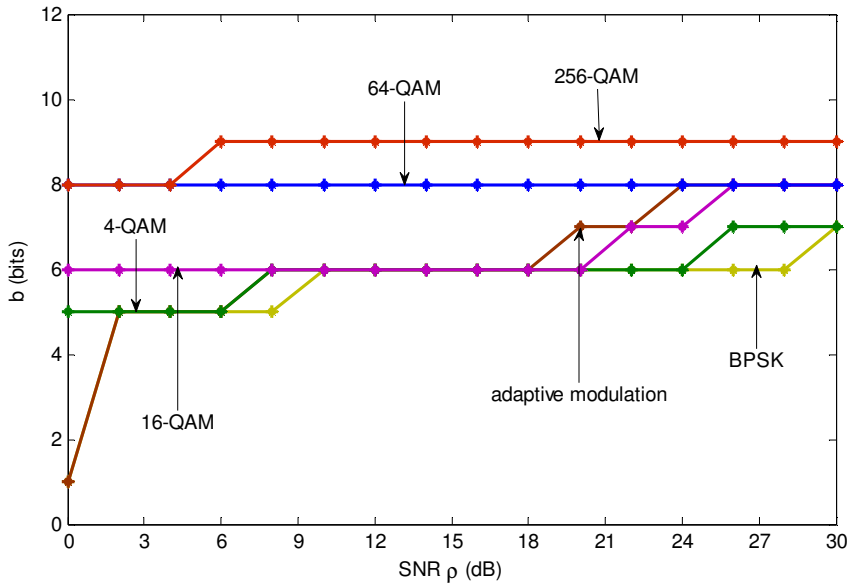

Figure 5-11 The optimal number of bits "*b*" (based on practicable capacity) for different values of SNR (dB) with optimal power allocation and linear quantizer for $R_{\mathrm{fb}} = 25$, $T = 1000$.

(Figure 5-11). This suggests a larger number of feedback bits which results in a decrease in the feedback noise variance (5-9). Since a larger training interval (makes $\sigma^2_{\tilde{\mathbf{H}}_{\mathrm{r}}}$ decrease) is preferred for higher constellation sizes (Figure 5-8, and Figure 5-9), it is evident from (5-7) that increasingly accurate knowledge of the channel at the transmitter is required for increasing constellation sizes.

5.4 Summary and Conclusions

For an eigen-MIMO wireless link, the information-theoretic and the practicable capacities are formulated, and then maximized over the training and data power allocation, the number of training symbols, and the feedback delay interval. The spectral resource required for the feedback is not included in the theoretical or practicable capacity, but its requirement is significant in a changing

channel, and has been quantified for a slowly changing channel example. The imperfection of the feedback is included in the capacity formulation, through the feedback noise variance (quantization noise, but not delay because of the block-fading model) of equation (5-9). Under these conditions, the optimal training interval for maximum theoretic capacity turns out to follow the same trend as in an open-loop system i.e., with the optimal power allocation, the optimal number of training symbols is equal to the number of transmit antennas. However, this is not the case with the practicable capacities: with adaptive modulation, the optimal training interval is equal to or greater than the optimal training interval associated with the theoretic capacity; without adaptive modulation, a large number of training symbols is required especially at low to mid values of SNR where most of the eigenchannels are operated in their near-zero capacity region.

In particular, for a given SNR, the optimal training and feedback delay intervals (number of required quantization bits) both increase with the modulation constellation size. This relates to the need for more accurate CSI at the transmitter for modulations with larger constellations. A comparison is also made with the maximized theoretic capacity of an open-loop MIMO system, and it turns out that as the result of imperfect feedback**,** a closed-loop system may not have a higher theoretic capacity than the open-loop system at high values of SNR. However, for typical wireless mobile communication systems, operation at low SNRs is important, and here the closed-loop system offers a higher information-theoretic capacity.

REFERENCE LIST

[5.1] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *J. Wireless Personal Commun. - Springer*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[5.2] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Trans. Telecommun.,* vol. 10, no. 6, pp. 585-595, Nov.-Dec. 1999.

[5.3] G. Lebrun, J. Gao, and M. Faulkner, “MIMO transmission over a time-varying channel using SVD,” *IEEE Trans. Wireless Commun.,* vol. 4, no. 2, pp. 757-764, Mar. 2005.

[5.4] V. Lau, Y. Liu, and T.-A. Chen, “On the design of MIMO block-fading channels with feedback-link capacity constrainit,” *IEEE Trans. Commun.,* vol. 52, no. 1, pp. 62-70, Jan. 2004.

[5.5] K. N. Lau, Y. J. Liu, and T. A. Chen, “On the role of transmit diversity in wireless communications with partial feedback,” *IEEE Trans. Commun.,* vol. 50, no. 12, pp. 2082-2090, Dec. 2002.

[5.6] S. Zhou and G. B. Giannakis, ”Optimal transmitter eigen-beamforming and space-time block code based on channel mean feedback,” *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2599-2613, Oct. 2002.

[5.7] B. Mielczarek, and W. A. Krzymień, “Influence of CSI feedback errors on capacity of linear multi-user MIMO systems,” in *Proc. IEEE Veh. Technol. Conf.* pp. 2043-2047, Apr. 2007.

[5.8] N. Jindal, “MIMO broadcast channels finite-rate feedback,” *IEEE Trans. Inf. Theory,* vol. 52, no. 11, pp. 5045-5060, Nov. 2006.

[5.9] B. Mielczarek and W. A. Krzymień, “Comparison of partial CSI encoding methods in multi-user MIMO systems,” *J. Wireless Personal Commun.- Springer*, vol. 52, no. 1, Jan. 2010.

[5.10] A. Dana, M. Sharif, and B. Hassibi, “On the capacity region of multi-antenna Gaussian broadcast channels with estimation error,” in *Proc. IEEE Int. Symp. Inf. Theor*y, pp. 1851-1855, Jul. 2006.

[5.11] T. Yoo and A. Goldsmith, “Capacity and power allocation for fading MIMO channels with channel estimation error,” *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203-2214, May. 2006.

[5.12] W. Santipach and M. L. Honig, "Optimization of training and feedback for beamforming over a MIMO channel," in *Proc. IEEE Wireless Commun. and Network. Conf.*, pp. 1139-1143, Mar. 2007

[5.13] M. Kobayashi, N. Jindal, and G. Caire, “Training and feedback optimization for multiuser MIMO downlink," acceped in *IEEE Trans. Commun.*, pp. 1-13, 2011.

[5.14] B. Hassibi, and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links?,” *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951-963, Apr. 2003.

[5.15] E. Visotsky and U. Madhow, “Space-time transmit precoding with imperfect feedback,” *IEEE Trans. Inf. Theory*, vol. 47, pp. 2632–2639, Sep. 2001.

[5.16] S. Ihara, “On the capacity of channles with additive non-Gaussian noise,” *J. of Inf. Contr. - Elsevier*, vol. 37, pp. 34-39, Sep. 1978.

[5.17] S. Catreux, P. F. Driessen and L. J. Greenstein, “Data throughputs using multiple-input multiple-output (MIMO) techniques in a noise-limited cellular environment,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 226 – 235, Apr. 2002.

[5.18] S. A. Mirtaheri, and R. G. Vaughan, “Practicable MIMO capacity in ideal channels,” in *Proc. IEEE Veh. Technol. Conf.* vol. 4, pp. 1620 – 1624, May 2006.

[5.19] S. A. Banani, and R. G. Vaughan, “Practicable capacity in eigen-MIMO with channel estimation and mean feedback,” submitted to *IET Communications,* Jan. 2011. Manuscript ID: COM-2011-0045.

[5.20] S. A. Banani, and R. G. Vaughan, “The effect of training-based channel estimation on the capacity of closed-loop MIMO systems with imperfect CSI feedback,” in *Proc. IEEE Veh. Technol. Conf.*, pp. 1-5, Sep. 2010.

[5.21] J. G. Proakis, *Digital Communications,* 5th ed. New York: McGraw-Hill, 2008.

[5.22] J. K. Cavers, *private communication,* Nov. 2011.

[5.23] A. H. Sayed, *Fundamentals of adaptive filtering*, Hoboken, NJ: Wiley, 2003.

6: THROUGHPUT AND SER TRADE-OFF IN EIGEN-MIMO WITH FIXED MODULATION

Water-filled eigenchannels maximize the multiple-input multiple-output information theoretic capacity, but for realization, adaptive modulation means high complexity. A simplification is to have a fixed number of eigenchannels with a common modulation. For such a system, the uncoded SNR/SER performance trades off with the throughput which depends on the receiver SNR and the power allocation on the eigenchannels. In this chapter, an eigen-MIMO system, simplified by imposing a fixed number of eigenchannels with a fixed, common modulation, is analyzed. The throughput and reliability (the uncoded SER), are determined together. The method is to constrain the eigenchannel SNRs for reliability and seek a power allocation over the eigenchannels for high capacity. A feature of this design approach is that the capacity (and the throughput, which has similar behavior) and the error performance can be managed through the choice of the maximum allowable SNR loss on eigenchannels.

Spatial multiplexing [6.1] and diversity, e.g., [6.2], [6.3] are candidates for achieving high data rates in multiple-input multiple-output (MIMO) wireless links. Both systems have the practical advantage that no channel knowledge is required at the transmitter. But this practical advantage comes with a performance penalty (relative to what is theoretically possible) because joint optimization between the transmitter and receiver is not fully deployed. With

feedback from receiver to transmitter, also called closed loop MIMO, system performance can be improved based on: a desired optimization criterion such as capacity, average error rate, etc.; channel knowledge (full or partial); a transmit power constraint (total power, peak power, etc.); or other considerations such as constraints on the type of signaling, multiple access, and receiver design, etc.

Several linear precoder-decoder (LPD) designs for closed loop operation which are jointly optimal for different criteria are presented in the literature [6.4]-[6.20]. In these, the required channel state information (CSI) at the transmitter is often assumed to be perfect and undelayed. The optimal design for maximizing SNR is to allocate all the power to the eigenchannel with the maximum gain, sometimes called dominant eigenmode transmission [6.6]. The optimum design that maximizes capacity with a total power constraint, is water-filled eigen-MIMO [6.7], [6.8]. In terms of adaptive antennas, the narrowband transmit and receive weights for each parallel eigenchannel are given by the sets of right and left singular vectors, respectively, of the channel matrix. Eigenchannels allow good capacity realization using established, high spectral efficiency communications techniques. For the Rayleigh channels of mobile communications, the large majority of the capacity is available through deploying just one or a few of the available eigenchannels, e.g., [6.9]. In the idealized MIMO channel, the maximum eigenchannel has full diversity order, and the lowest eigenchannel has less capacity than any of the transmit and receive element pairs.

Other examples of optimized LPD designs are as follows. The optimal design that minimizes the sum of output symbol estimation errors, allocates the

available power to the eigenchannels according to an *inverse* water-filling policy [6.5], [6.10]. A different approach based on inverse SNR (or equivalently mean squared estimation error) between the input and output of each mode is presented in [6.5], [6.11]. A general solution is a weighted sum of inverse SNRs and leads to a number of well-known solutions depending on the choice of weights [6.11]. The optimum LPD design that maximizes SNR subject to a zero-forcing constraint on the channel has also been considered in [6.10]. The optimal LPD design for minimizing the pairwise error probability (PEP) with a total transmit power constraint is given in [6.12]. Minimization of the geometric mean square error, defined as the determinant of the error covariance matrix, is the subject in [6.13], [6.14]. There are also power allocation optimizations for different types of communications systems, e.g. [6.15], [6.16], which, like several of the above schemes, optimize the power allocation for minimizing the uncoded error rate. The work in [6.17] considers the problem of information rate maximization with a constraint on the peak power. There are many other specialized schemes (e.g. [6.18]-[6.28] which include the previous investigations of MIMO performance degradation from channel estimation and imperfect feedback), too numerous to include here, but the basic principles for closed loop MIMO have appeared in different forms from different disciplines such as information theory, signal processing, communications theory and techniques, and adaptive antennas.

There remain major technical challenges for optimized LPD designs even when the CSI is assumed perfect. Design challenges include having a system

with good throughput and good reliability at the same time. Because of these challenges, there is ongoing interest in configurations which are simplified, and their associated analysis of throughput and error performance. For the limiting capacity case of perfect water filling, the different eigenchannel gains (and different numbers of eigenchannels, in general, between different channel realizations) contribute to the difficulty of optimally allocating different digital modulations to different eigenchannels [6.29]. Sophisticated hardware is needed (and usually not available in legacy systems) at all the terminals to support variable modulation, and a sophisticated protocol is required to support the modulation adaptation via data exchange and timing handshakes for the blocks of differently modulated data. A major simplification is to have a fixed, common modulation with a common signaling rate, for a fixed number of eigenchannels. With this scheme, the simplification affects both the protocol, and in particular, the hardware and software at all the terminals of a multi-user system. However, the system is still has the complexity of being closed loop because the eigenchannels must be established and maintained, i.e., the adaptive antenna weights (i.e., the precoder-decoder) are still required along with a protocol to set them.

A feature of water-filling with fixed (or adaptive) modulation, is that the uncoded SNR/SER performance of the overall system deteriorates with the weakest eigenchannel. The nature of this deterioration is presented below, for 3×3 and a 4×4 systems. The deterioration is serious because of the large difference between the gains of the eigenchannels, and an associated large

degradation occurs in the overall system SER performance. For example, in a 3×3 system with water-filling, the difference between the average SNR on the strongest and weakest eigenchannel is about 13.5 dB when the average SNR at each receiver is 20 dB. This difference increases to 16.8 dB in a 4×4 system. As a result, full eigen-MIMO systems with water-filling (i.e. with maximized capacity without an error performance constraint) have a weak uncoded SER performance which is not desirable in practice. A preferred system has good reliability and high throughput, simultaneously.

In this work [6-30], an eigen-MIMO system, simplified by imposing a fixed number of eigenchannels with a fixed, common modulation, is analyzed. The throughput and reliability (the uncoded SER), are determined together. The method is to constrain the eigenchannel SNRs for reliability and seek a power allocation over the eigenchannels for high capacity.

For discussing the approach, some elaboration on capacity is helpful. In this chapter, the term *capacity* refers to the ergodic, information-theoretic capacity, typically with units of bits/channel-use. This is also referred to here as the *theoretic capacity* or *information-theoretic capacity* where the context calls for emphasis of this definition. The capacity carries the assumption of error-free transmission of symbols, which in turn implies infinitely long code blocks. The term *throughput*, in bits/sec/Hz, is used for the throughput of bits that would be perfectly detected by an ideal coder. The throughput carries the assumption of error free transmission of finite blocks of symbols. Throughput is also referred to in the literature as practicable capacity (also in in Chapters 5 and 7), data

throughput, goodput, spectral efficiency, etc. Only uncoded data is discussed here, although the choice of SNR/SER constraint in the formulation is useful because it defines the required performance for a designer's choice of coder. The *actual throughput* is therefore not discussed here, and in fact is seldom addressed in the MIMO capacity literature, because it is system-specific. Its calculation requires the implementation detail of a coding architecture (trellis, turbo, LDPC, and associated code rates, etc.) for the system, the actual RF bandwidth used (i.e. the choice of filters and guardbands and guardtimes, etc.), and needs to account for the spectral resources required for any multiple access scheme, and maintaining the eigenchannels (channel sounding, transmission of the CSI and any variable modulation information and associated timing, etc.) for each user. The throughput, in combination with an understanding of the coder requirements from the error performance detail (the error rate for each eigenchannel), is used in this chapter to govern a basic design configuration, preparing the way for system-specific implementation protocol and coding. The allocation of one example of coding - viz., Reed-Solomon (chosen because only RS is the only coding that has a tractable error expression) - to eigen-MIMO, and the optimized throughput that results from such a system, are presented in chapter 7 and [6.31].

In order to determine the trade-offs between the throughput and reliability for the fixed eigen-MIMO system, a maximization of throughput with a constrained SNR is sought. The throughput depends on the constellation size of the modulation and the average SNR at the receivers. In particular, for high

values of receiver SNR, the MIMO throughput reaches its maximum of, for example, $L\log_2 \mathrm{D} = 8$ for uncoded 4-QAM ($\mathrm{D} = 4$) transmission over an idealized $L \times L = 4 \times 4$ system. But for lower values of SNR, the throughput depends on SNR (*cf.,* Figure 6-9 and Figure 6-10, below). In addition, the power allocation over the eigenchannels affects the throughput differently for different values of SNR. In seeking to realize the limits of the throughput, the following issues arise.

- Unlike the capacity, there is no exact analytical formulation for the throughput available. Although some approximate formulations are available, they require specific assumptions. For example, in [6.32], the throughput is expressed in terms of a block error rate (BLER) which requires perfect detection. But there is no error detection or correction, and so the coder performance is not included in the throughput calculation. Also, the throughput calculation assumes that blocks are correctly detected if and only if all decisions within the block are error-free. This is a major assumption. To attempt to realize this throughput, a choice of coding scheme is required of course, and the coder performance would have to be included in order to establish the actual capacity. The expected coder performance in turn requires knowledge of the (uncoded) BER and its distribution across the eigenchannels.

- There has not been a common approach for evaluation of throughput. For example, the results in [6.32], which are obtained with no constraint on BER, are the same as in [6.33] where the throughput (referred to in [6.33]

as spectral efficiency), is obtained by an optimization of variable power transmission for a constrained uncoded BER (referred to as output BER).

- The algorithms presented in [6.23]-[6.34] (and many other related works) are based on adaptive modulation for achieving the highest throughput over single-input, multiple-output channels with variable power at the transmitter. In other words, these are not addressing the power allocation which is addressed here.

Because of the above points, the problem of power allocation is tackled here through maximization of the capacity with an eigenchannel SNR constraint, and then applying the resulting power allocation for the fixed modulation arrangement. Although the power allocation obtained this way is unlikely to be optimum for maximizing the throughput (because the optimization is based on the capacity rather than throughput), it is shown via simulation that the throughput using the presented power allocation shows similar behavior to the capacity. In this sense, the capacity is being maximized under some SNR constraints, in order to determine the tradeoff between the error performance and throughput (at each receiver SNR) for a reduced complexity system which is uncoded and has fixed modulation.

Mathematically, the approach comprises information rate maximization in eigen-MIMO which is constrained by total input power and a maximum allowable SNR loss on the weakest eigenchannel. The maximum allowable SNR loss is the maximum difference between the average SNR on the weakest eigenchannel and the average SNR at each receiver. This is a quantity available to a system

designer which manages the tradeoff between acceptable throughput and acceptable uncoded SER (defines the performance required from the choice of coding arrangement) for the system. The optimum LPD design of eigenchannels requires a power allocation in terms of SNR in each eigenchannel. Three different models for the SNR constraint are considered, each of which is characterized using a parameter (γ, below) which controls the allowable maximum SNR loss on the weakest eigenchannel. The optimal power allocation is obtained via convex optimizations with inequality constraints, using an interior-point method called the logarithmic barrier method [6.35].

Based on the average receiver SNR at each antenna and the maximum SNR loss on the weakest eigenchannel, the best power allocation scheme (along with the best choice for the parameter γ) is selected for maximum capacity. The selection is from the three SNR models (I / II / III) presented below; water-filling applied to different numbers of eigenchannels; and dominant eigenmode transmission. As expected, the resultant capacities are upper and lower bounded by the optimal water-filled capacity and the one obtained via dominant eigenmode transmission, respectively. The position of the capacity between these bounds can be controlled by deploying the appropriately selected SNR model, along with the parameter characterizing it. The throughput (based on the formula provided in [6.32]) and the SER performance, are also evaluated by simulation for a system which uses fixed QAM. The presented approach gives better SER performance for the fixed modulation system than using water-filling, at the expense of a small loss in throughput.

The rest of the chapter is organized as follows. Section 6.1 describes the closed-loop MIMO system model followed by the information rate maximization with SNR constraints in section 6.2. The simulations results are presented in section 6.3, with the conclusions in section 6.4. The sub-script "$\mathbf{H}$" refers to sample realization of channel matrix $\mathbf{H}$.

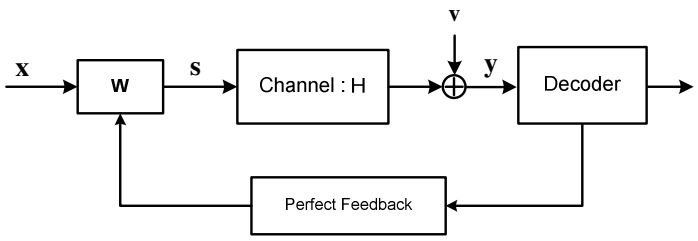

Figure 6-1 A closed-loop MIMO system with a transmit beamformer for using channel knowledge which is perfectly fed back from the receiver

6.1 Closed-Loop MIMO System Model

As introduced in 5.1, and repeated here for convenience, the MIMO channel is quasi-static, flat block-fading with M transmit and N receive. The channel is random but remains static over a fading block cycle (starting from the training transmission, including the channel estimation, quantization, codeword feedback, and calculation of the new beamformer, and detection of the received symbols), but is independent, across different blocks. A beamformer of weights $\mathbf{W}$ derived using channel knowledge at the transmitter, is reproduced for convenience from Figure 5-1, in Figure 6-1. At each symbol time, the $M \times 1$ data vector signal to be transmitted, $\mathbf{x}$, is multiplied by the $M \times M$ weight matrix $\mathbf{W}$, before transmission. The input-output relation for the MIMO system in the data transmission is often written

$$\mathbf{y} = \sqrt{\frac{\rho}{M}}\,\mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{v} = \sqrt{\frac{\rho}{M}}\,\mathbf{H}\mathbf{s} + \mathbf{v} \tag{6-1}$$

where $\mathbf{y}$ is the $N \times 1$ received signal vector and $\mathbf{H}$ is the channel matrix with i.i.d. zero mean complex Gaussian entries each with unity variance; and ρ is the constant (across blocks) average SNR at each receive antenna for the data transmission. The vector $\mathbf{v}$ represents the additive noise with $\mathbf{v} \sim CN(0,1)$. From $\mathbf{x}$ satisfying $E\,\mathbf{x}\mathbf{x}^H = \mathbf{I}_M$ and $\mathbf{s} = \mathbf{W}\mathbf{x}$, the covariance matrix of the transmitted signal $\mathbf{R}_{\mathbf{ss}} = E\,\mathbf{s}\mathbf{s}^H$ is given by

$$\mathbf{R}_{\mathbf{ss}} = \mathbf{W}\mathbf{W}^H . \tag{6-2}$$

Similarly, from the total average transmit power constraint, $\mathbf{W}$ satisfies $\|\mathbf{W}\|_F^2 = \mathrm{tr}\,\mathbf{R}_{\mathbf{ss}} = M$ so with $\mathbf{s} = \mathbf{W}\mathbf{x}$, we have $E\,\mathrm{tr}\,\mathbf{s}\mathbf{s}^H = M$.

6.2 Information-Rate Maximization

This section develops the mathematical basis for the constrained information rate maximization. The instantaneous capacity – that of a sample realization of the channel - is first addressed. This capacity, with a fixed channel $\mathbf{H}$, is defined as [6.8]

$$C_{\mathbf{H}} = \max_{p_{\mathbf{s}}(\mathbf{s})} I(\mathbf{s};\mathbf{y}) = \max_{E[\mathrm{tr}(\mathbf{s}\mathbf{s}^H)]=M} \log_2 \det\left(\mathbf{I}_N + \frac{\rho}{M}\mathbf{H}\mathbf{R}_{\mathbf{ss}}\mathbf{H}^H\right) \tag{6-3}$$

where $p_{\mathbf{s}}(\mathbf{s})$ is the probability density function of the vector $\mathbf{s}$, and $I(\mathbf{s};\mathbf{y})$ is the mutual information between $\mathbf{s}$ and $\mathbf{y}$. The ergodic capacity is the ensemble

average of the capacity achieved when the optimization is performed for each realization of $\mathbf{H}$, i.e. $C = E[C_{\mathbf{H}}]$.

Applying singular value decomposition, the matrix $\mathbf{H}$ with rank $r = \min(M, N)$ can be rewritten as $\mathbf{H} = \mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{V}^H$ where $\mathbf{V}$ and $\mathbf{U}$ are the unitary matrices containing the corresponding input and output singular vectors, respectively, and $\boldsymbol{\Lambda}^{1/2}$ is a non-negative $N \times M$ diagonal matrix with ith diagonal element as $\lambda_i^{1/2}$ (the positive square root of ith eigenvalue of $\mathbf{H}^H\mathbf{H}$). In addition the diagonal elements satisfy $\lambda_i \geq \lambda_{i+1}$. As a result, $\mathbf{H}^H\mathbf{H} = \mathbf{V}\Lambda\mathbf{V}^H$, and

$$C_{\mathbf{H}} = \max_{E[\mathrm{tr}(\mathbf{s}\mathbf{s}^H)]=M} \log_2 \det\left(\mathbf{I}_N + \frac{\rho}{M}\boldsymbol{\Lambda}^{1/2}\mathbf{V}^H\mathbf{R}_{\mathbf{ss}}\mathbf{V}\boldsymbol{\Lambda}^{1/2}\right) \tag{6-4}$$

Also, $\tilde{\mathbf{R}}_{\mathbf{ss}} = \mathbf{V}^H\mathbf{R}_{\mathbf{ss}}\mathbf{V}$ is non-negative definite because $\mathbf{R}_{\mathbf{ss}}$ is non-negative definite, and $\mathrm{tr}(\tilde{\mathbf{R}}_{\mathbf{ss}}) = \mathrm{tr}(\mathbf{R}_{\mathbf{ss}})$. As a result, the maximization over $\mathbf{R}_{\mathbf{ss}}$ with $E[\mathrm{tr}(\mathbf{s}\mathbf{s}^H)] = M$, can be over $\tilde{\mathbf{R}}_{\mathbf{ss}}$ with $E[\mathrm{tr}(\tilde{\mathbf{s}}\tilde{\mathbf{s}}^H)] = M$ [6.6], [6.8]. Since for any non-negative definite matrix $\mathbf{A}$, $\det\mathbf{A} \leq \prod_i [\mathbf{A}]_{i,i}$,

$$\det\left(\mathbf{I}_N + \frac{\rho}{M}\boldsymbol{\Lambda}^{1/2}\tilde{\mathbf{R}}_{\mathbf{ss}}\boldsymbol{\Lambda}^{1/2}\right) \leq \prod_{i=1}^{r}\left(1 + \frac{\rho}{M}[\tilde{\mathbf{R}}_{\mathbf{ss}}]_{i,i}\lambda_i\right) \tag{6-5}$$

with the equality for when $\tilde{\mathbf{R}}_{\mathbf{ss}}$ is diagonal. Thus

$$C_{\mathbf{H}} = \max_{E[\mathrm{tr}(\tilde{\mathbf{s}}\tilde{\mathbf{s}}^H)]=M} \sum_{i=1}^{r}\log_2\left(1 + \frac{\rho}{M}p_i\lambda_i\right) \tag{6-6}$$

where $p_i = [\tilde{\mathbf{R}}_{\mathbf{ss}}]_{i,i}$ (Note p_i is different to $p_{\mathbf{s}}(\mathbf{s})$). We see that the optimum design corresponds to the SNR on the eigenchannels. The above results are known.

The remaining problem is the optimal power allocation under different SNR constraints described below. Once $\tilde{\mathbf{R}}_{\mathbf{ss}}^{\mathrm{opt}}$ is found, an optimal beamformer $\mathbf{W}^{\mathrm{opt}}$ is obtained from (6-2) and $\mathbf{R}_{\mathbf{ss}}^{\mathrm{opt}} = \mathbf{V}\tilde{\mathbf{R}}_{\mathbf{ss}}^{\mathrm{opt}}\mathbf{V}^H$ as

$$\mathbf{W}^{\mathrm{opt}} = \mathbf{V}\left(\tilde{\mathbf{R}}_{\mathbf{ss}}^{\mathrm{opt}}\right)^{1/2}, \tag{6-7}$$

meaning that the data vector $\mathbf{x}$ is multiplied by the matrix $\mathbf{W}^{\mathrm{opt}}$ before transmission occurs. In turn, the receiver multiplies the received data by the matrix $\mathbf{U}^H$. As a result, the overall transmission relationship is

$$\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{y} = \sqrt{\frac{\rho}{M}}\boldsymbol{\Lambda}^{1/2}\tilde{\mathbf{x}} + \tilde{\mathbf{v}} \tag{6-8}$$

where $\tilde{\mathbf{x}} = (\tilde{\mathbf{R}}_{\mathbf{ss}}^{\mathrm{opt}})^{1/2}\mathbf{x}$ and $\tilde{\mathbf{v}} = \mathbf{U}^H\mathbf{v} \sim CN(0,1)$. This denotes the transformation of the MIMO channel into parallel eigenchannels with unequal instantaneous SNRs (the SNR of the *i*th eigenchannel at channel realization **H**, is $SNR_{i,\mathbf{H}} = (\rho/M)\,p_i\lambda_i$)), and the MIMO capacity is obtained by summing the eigenchannel capacities which are governed by their SNRs. Three different configurations for the SNR constraint come to mind, as follows.

6.2.1 SNR Constraint Model I

In the first model, the $2 \le L \le r$ strongest eigenchannels are used and the output SNR of the eigenchannels is constrained to be

$$SNR_{j,\mathbf{H}} = \gamma SNR_{1,\mathbf{H}} \;;\; j = 2,...,L \tag{6-9}$$

where $SNR_{j,\mathbf{H}} = \frac{\rho}{M} p_j \lambda_j$ is the instantaneous output SNR of the *j*th eigenchannel (for fixed channel $\mathbf{H}$) and $0 < \gamma \leq 1$ is a parameter that makes it possible for the user to control the relative SNR in each eigenchannel. From (6-9),

$$p_j \lambda_j = \gamma p_1 \lambda_1 \;\; ; \;\; j = 2,...,L \tag{6-10}$$

From the above $L-1$ constraints and $\sum_{i=1}^{L} p_i = M$, the power allocation is simply obtained through the set of L linear equations

$$\underbrace{\begin{bmatrix} \gamma\lambda_1 & -\lambda_2 & 0 & \cdots & 0 \\ \gamma\lambda_1 & 0 & -\lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma\lambda_1 & 0 & 0 & \cdots & -\lambda_L \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{L-1} \\ p_L \end{bmatrix}}_{\mathbf{p}} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ M \end{bmatrix}}_{\mathbf{d}} \tag{6-11}$$

and is given by $\mathbf{p}^{\text{opt}} = \mathbf{B}^{-1}\mathbf{d}$. As a result, $C_{\mathbf{H}} = \sum_{i=1}^{L} \log_2\left(1 + \frac{\rho}{M} p_i^{\text{opt}} \lambda_i\right)$. for $L = 1$, the configuration is simply the dominant eigenmode transmission.

6.2.2 SNR Constraint Model II

Here, the $2 \leq L \leq r$ strongest eigenchannels are used and are constrained to satisfy

$$SNR_{j,\mathbf{H}} \geq \gamma SNR_{1,\mathbf{H}} \;\; ; \;\; j = 2,...,L \tag{6-12}$$

Thus, in this case, there are $L-1$ inequality constraints,

$p_j\lambda_j \geq \gamma p_1\lambda_1 ; j = 2,...,L$ along with one equality constraint, $\sum_{i=1}^{L} p_i = M$. The mutual information maximization problem is now

$$\begin{aligned} \underset{p_1,p_2,...,p_L}{\text{maximize}} \quad & f_1(p_1,p_2,...,p_L) = \sum_{i=1}^{L} \log_2\left(1+\frac{\rho}{M} p_i\lambda_i\right) \\ \text{subject to} \quad & p_j\lambda_j \geq \gamma p_1\lambda_1 ; j = 2,...,L \\ & \textstyle\sum_{i=1}^{L} p_i = M \end{aligned} \tag{6-13}$$

where the objective for the maximization, i.e. $\sum_{i=1}^{L} \log_2\left(1+\frac{\rho}{M} p_i\lambda_i\right)$, is concave in the variable $p_i ; i = 1,...,L$. Equivalently, (6-13) can be rewritten as

$$\begin{aligned} \underset{p_1,p_2,...,p_L}{\text{minimize}} \quad & f_1(p_1,p_2,...,p_L) = -\sum_{i=1}^{L} \log_2\left(1+\frac{\rho}{M} p_i\lambda_i\right) \\ \text{subject to} \quad & f_j(p_1,p_2,...,p_L) = \gamma p_1\lambda_1 - p_j\lambda_j \leq 0 ; j = 2,...,L \\ & \textstyle\sum_{i=1}^{L} p_i = M \end{aligned} \tag{6-14}$$

where $f_1, f_2,...,f_L$ are convex real functions of $p_i ; i = 1,...,L$. In fact, (6-14) is a convex optimization problem that includes inequality constraints. An interior-point algorithm called the logarithmic barrier method [6.35] is used here to solve (6-14). The first step is to rewrite the problem (6-14), making the inequality constraints implicit in the objective, i.e.,

$$\begin{aligned} \underset{p_1,p_2,...,p_L}{\text{minimize}} \quad & f_1(p_1,p_2,...,p_L) + \sum_{j=2}^{L} g(f_j(p_1,p_2,...,p_L)) \\ \text{subject to} \quad & \textstyle\sum_{i=1}^{L} p_i = M \end{aligned} \tag{6-15}$$

where $g : \Re \to \Re$ is an indicator function for the nonpositive reals, viz.,

$$g(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases} \tag{6-16}$$

The problem (6-15) has no inequality constraint, but its objective function is not differentiable, so descent methods (e.g, gradient descent, Newton's method, etc.) cannot be applied. The basic idea of the logarithmic barrier method is to approximate the indicator function, g, by the function

$$\hat{g}(u) = -(1/t)\log_{10}(-u) \tag{6-17}$$

where $t > 0$ is a parameter that sets the accuracy of the approximation. Like g, the function $\hat{g}$ is convex and nondecreasing, and yields ∞ for $u > 0$. Unlike g, however, $\hat{g}$ is differentiable and closed: it approaches infinity as u approaches zero. As t increases, the approximation becomes more accurate. Substituting $\hat{g}$ for g in (6-15) gives the approximation

$$\begin{aligned} &\underset{p_1,p_2,...,p_L}{\text{minimize}} \quad f_1(p_1,p_2,...,p_L) + (1/t)\sum_{j=2}^{L} -\log_{10}(-f_j(p_1,p_2,...,p_L)) \\ &\text{subject to} \quad \sum_{i=1}^{L} p_i = M \end{aligned} \tag{6-18}$$

The objective here is convex, since $-(1/t)\log_{10}(-u)$ is convex and increasing in u, and differentiable. The function $\phi(p_1,p_2,...,p_L) = -\sum_{j=2}^{L}\log_{10}(-f_j(p_1,p_2,...,p_L))$ is called the *logarithmic barrier* or *log barrier* for the problem (6-18). Its domain is the set of points that strictly satisfies the inequality constraints of (6-14). As a result the objective function in (6-18) is convex only over the region that $f_j(p_1,p_2,...,p_L) \le 0; j = 2,...,L$. No matter what value the positive parameter t, the logarithmic barrier grows without bound if $f_j(p_1,p_2,...,p_L) \to 0; j = 2,...,L$. Denoting $(p_1^{\text{opt}},p_2^{\text{opt}},...,p_L^{\text{opt}})$ as the solution of

(6-18), it is shown [6.35] that $(p_1^{\text{opt}}, p_2^{\text{opt}}, ..., p_L^{\text{opt}})$ is no more than $(L-1)/t$-suboptimal. This suggests a straightforward method for solving the original problem (6-14), with a guaranteed accuracy ε by taking $t = (L-1)/\varepsilon$.

The next step is to solve the equality constraint problem (6-18), by eliminating the equality constraint and then solve the resulting unconstrained problem using methods for unconstrained minimization. We can eliminate p_1 (for example) using the parameterization

$$p_1 = M - \sum_{i=2}^{L} p_i \tag{6-19}$$

The reduced problem is then

$$\begin{aligned} \underset{p_2, p_3, ..., p_L}{\text{minimize}} \quad & -\left[\log_2\left(1 + \frac{\rho}{M}\lambda_1 (M - \textstyle\sum_{i=2}^{L} p_i)\right) + \sum_{i=2}^{L} \log_2\left(1 + \frac{\rho}{M} p_i \lambda_i\right)\right] \\ & -(1/t)\left[\sum_{i=2}^{L} \log_{10}\left(p_i\lambda_i - \gamma\lambda_1 (M - \textstyle\sum_{j=2}^{L} p_j)\right)\right] \end{aligned} \tag{6-20}$$

or equivalently

$$\begin{aligned} \underset{p_2, p_3, ..., p_L}{\text{minimize}} \quad \psi(p_2, p_3, ..., p_L) = & -t\left[\log_2\left(1 + \frac{\rho}{M}\lambda_1 (M - \textstyle\sum_{i=2}^{L} p_i)\right) + \sum_{i=2}^{L} \log_2\left(1 + \frac{\rho}{M} p_i \lambda_i\right)\right] \\ & -\left[\sum_{i=2}^{L} \log_{10}\left(p_i\lambda_i - \gamma\lambda_1 (M - \textstyle\sum_{j=2}^{L} p_j)\right)\right] \end{aligned} \tag{6-21}$$

where we have multiplied the objective by t. Since the composition of a convex function within an affine function is also convex, eliminating equality constraints preserves convexity of a problem. Moreover, the process of eliminating equality constraints and the reconstructing the solution of the original problem from the

solution of the transformed problem, involves standard linear algebra operations. Thus, the problems (6-18) and (6-21) are equivalent.

Now, since $\psi(p_2, p_3, ..., p_L)$ is differentiable and convex, a necessary and sufficient condition for a point $(p_2^{\text{opt}}, p_3^{\text{opt}}, ..., p_L^{\text{opt}})$ to be optimal is

$$\nabla\left(\psi(p_2^{\text{opt}}, p_3^{\text{opt}}, ..., p_L^{\text{opt}})\right) = 0 \tag{6-22}$$

where $\nabla(.)$ denotes the gradient operator. Thus solving the unconstrained minimization problem (6-21) is the same as finding a solution of (6-22), which is a set of $L-1$ equations in the $L-1$ variables $p_2, p_3, ..., p_L$. Since a closed-form solution to the minimization problem (6-21) does not exist, the problem can be solved by any iterative algorithm such as gradient descent method. This is written as

$$\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} - \mu^{(n)} \nabla\psi \tag{6-23}$$

where $\mathbf{p}^{(n)} = \begin{bmatrix} p_2^{(n)} & p_3^{(n)} & \cdots & p_L^{(n)} \end{bmatrix}^T$ is the solution vector point at *n*th iteration and

$$\begin{aligned}\left[\nabla\psi\right]_{\substack{k-1,1 \\ k=2,...,L}} = \frac{\partial\psi}{\partial p_k} = -\frac{t}{\log_{10} 2}\Bigg(&\frac{-(\rho/M)\lambda_1}{1+(\rho/M)\lambda_1(M-\sum_{j=2}^{L} p_j)} + \frac{(\rho/M)\lambda_k}{1+(\rho/M)p_k\lambda_k}\Bigg) \\ &- \frac{\lambda_k + \gamma\lambda_1}{p_k\lambda_k - \gamma\lambda_1(M-\sum_{j=2}^{L} p_j)} - \sum_{\substack{m=2 \\ m\neq k}}^{L} \frac{\gamma\lambda_1}{p_m\lambda_m - \gamma\lambda_1(M-\sum_{j=2}^{L} p_j)}\end{aligned} \tag{6-24}$$

In (6-23), $\mu^{(n)}$ is the step size at iteration *n* that, in general, can be fixed a priori, or chosen via inexact line search. Here, a simple and effective inexact line search is used called backtracking line search [6.35] which depends on two

constants α, β with $0<\alpha<0.5$, $0<\beta<1$. It is called backtracking because it starts with unit step size and then reduces it by the multiplicative factor β until the stopping condition

$$\psi(\mathbf{p}^{(n)}-\mu^{(n)}\nabla\psi)\le\psi(\mathbf{p}^{(n)})-\alpha\mu^{(n)}\nabla\psi^{T}\nabla\psi \tag{6-25}$$

holds. The constant α can be treated as the acceptable fraction of the decrease in ψ predicted by linear extrapolation. In the terminology of iterative methods, the convergence of the gradient descent algorithm using the backtracking line search is at least linear [6.35].

The stopping convergence criteria of such an iterative algorithm is usually of the form $\|\nabla\psi\|\le\eta$, where η is small and positive, as suggested by the suboptimality condition [6.35]. The stopping criterion is often checked immediately after the descent direction, $-\nabla\psi$, is computed.

The method requires a suitable starting point $\mathbf{p}^{(0)}$. We note that the objective function in (6-21) is convex only on the region that the inequality constraints $f_j(p_1,p_2,....,p_L)\le 0; j=2,...,L$ are satisfied. As a result the objective function (6-21) is not convex over the whole work space $S=\left\{(p_2,p_3,...,p_L)\in\Re^{L-1}\middle|\, 0\le p_j\le M;\ p_1+\sum_{j=2}^{L}p_j=M\right\}$ of the iterative gradient descent algorithm in (6-23). Since the objective function $\psi(p_2,p_3,...,p_L)$ may have several local minima, the work space S is divided into D distinct subspaces and the iterative algorithm is run with different starting point candidates chosen from

different subspaces. For example, with uniform gridding of the work space and splitting the interval for p_j into q equal, distinct segments, there are, at most, $D = q^{L-1}$ subspaces. Then, the converged results are compared to see which one is the global minima. Denoting $\mathbf{p}_d^{(c)}$ as the convergent point associated with the starting point chosen from the dth segment, the optimum power allocation is set as

$$\mathbf{p}^{\text{opt}} = [p_1^{\text{opt}} \; p_2^{\text{opt}} \; \cdots \; p_L^{\text{opt}}]^T = [p_1^{\text{opt}} \; \mathbf{p}_{\text{opt}}^{(c)}]^T \tag{6-26}$$

with $$\mathbf{p}_{\text{opt}}^{(c)} = \arg \min_{\substack{\mathbf{p}_d^{(c)} \\ d=1,\ldots,D}} \psi(\mathbf{p}_d^{(c)})$$

In practice, MIMO systems with $N \le 4$ are of immediate interest, and it can be verified via simulations that for any $N \le 4$, the choice $q = N-1$ suffices to avoid local convergence using the above procedure. Having $\mathbf{p}^{\text{opt}}$, there results $C_{\mathbf{H}} = \sum_{i=1}^{L} \log_2\left(1 + \frac{\rho}{M} p_i^{opt} \lambda_i\right)$ and the ergodic capacity is the ensemble average of the capacity achieved when the optimization is for each realization of $\mathbf{H}$, i.e. $C = E[C_{\mathbf{H}}]$. In simulations, this average is approximated with a finite sample size of course.

6.2.3 SNR Constraint Model III

In this case, the SNR constraints are further restricted to follow

$$\begin{cases} SNR_{j,\mathbf{H}} \ge SNR_{j+1,\mathbf{H}} \; ; \; j = 1,\ldots,L-1 \\ SNR_{L,\mathbf{H}} \ge \gamma SNR_{1,\mathbf{H}} \end{cases} \tag{6-27}$$

Thus the optimization problem now becomes

$$\begin{aligned}
&\underset{p_1,p_2,\ldots,p_L}{\text{minimize}} \quad f_1(p_1,p_2,\ldots,p_L) = -\sum_{i=1}^{L}\log_2\left(1+\frac{\rho}{M}\,p_i\lambda_i\right)\\
&\text{subject to} \quad f_j(p_1,p_2,\ldots,p_L) = p_j\lambda_j - p_{j-1}\lambda_{j-1} \le 0\,;\, j=2,\ldots,L\\
&\qquad\qquad\quad f_L(p_1,p_2,\ldots,p_L) = \gamma p_1\lambda_1 - p_L\lambda_L \le 0\\
&\qquad\qquad\quad \textstyle\sum_{i=1}^{L} p_i = M
\end{aligned} \tag{6-28}$$

Following the same procedure described in section 6.2.2, the inequality constraints in (6-28) can be made implicit in the objective as (for large values of t)

$$\begin{aligned}
&\underset{p_1,p_3,\ldots,p_L}{\text{minimize}} \quad -t\sum_{i=1}^{L}\log_2\left(1+\frac{\rho}{M}\,p_i\lambda_i\right) - \left[\sum_{i=2}^{L}\log_{10}\left(p_{i-1}\lambda_{i-1}-p_i\lambda_i\right)+\log_{10}\left(p_L\lambda_L-\gamma p_1\lambda_1\right)\right]\\
&\text{subject to} \quad \textstyle\sum_{i=1}^{L} p_i = M
\end{aligned} \tag{6-29}$$

Next, by eliminating the equality constraint, the problem (6-29) is reduced to

$$\begin{aligned}
\underset{p_2,p_3,\ldots,p_L}{\text{minimize}} \;\; \psi(p_2,p_3,\ldots,p_L) = &-t\left[\log_2\left(1+\frac{\rho}{M}\,\lambda_1(M-\textstyle\sum_{i=2}^{L}p_i)\right)+\sum_{i=2}^{L}\log_2\left(1+\frac{\rho}{M}\,p_i\lambda_i\right)\right]\\
&-\left[\log_{10}\left((M-\textstyle\sum_{i=2}^{L}p_i)\lambda_1-p_2\lambda_2\right)+\sum_{i=3}^{L}\log_{10}\left(p_{i-1}\lambda_{i-1}-p_i\lambda_i\right)+\log_{10}\left(p_L\lambda_L-\gamma\lambda_1(M-\textstyle\sum_{i=2}^{L}p_i)\right)\right]
\end{aligned} \tag{6-30}$$

Finally, the problem is solved using the gradient descent algorithm as in the case with (6-21). The work space here is a subset of the work space in the previous case (section 6.2.2) since the set of constraints (6-27) automatically satisfies the constraints in (6-12). As a result, the optimum solution of (6-28) can be expected to be one of the local minima in (6-14). The structure of the SNR model III is very similar to the SNR architecture in the water-filling scheme, i.e.,

$SNR_{j,\mathbf{H}} \geq SNR_{j+1,\mathbf{H}}$; $j = 1,...,L-1$, except there is no guarantee that $p_j \geq p_{j+1}$; $j = 1,...,L-1$.

6.2.4 Switching Between Different Power Allocation Schemes

In general, none of the three SNR models maximizes the ergodic capacity for all values of system parameters ρ, γ, and SNR_L. (this can be seen in the simulations below.) The designer may choose from different power allocation schemes with the metric of ergodic capacity constrained by a maximum allowable SNR loss (relative to ρ) on the weakest eigenchannel, denoted δ_{SNR} in dB. δ_{SNR} can also be written in terms of the minimum allowable SNR imposed on the weakest eigenchannel, $SNR_L^*(\text{dB})$, as $SNR_L^*(\text{dB}) = \rho(\text{dB}) - \delta_{SNR}(\text{dB})$. The designer picks δ_{SNR} according to the desired capacity and SER system performances, i.e., capacity and SER are traded off through this parameter.

The different power allocation schemes for selection are from: the three SNR models I / II / III; the water-filling schemes applied to different number of eigenchannels, $L \geq 2$; and the dominant eigenmode transmission. The selection procedure is as follows. For each value of δ_{SNR}, the best power allocation scheme, along with the best choice for L and γ that achieves the highest information rate, is selected from the Capacity–SNR_L plane to be to the right of $SNR_L(\text{dB}) \geq SNR_L^*(\text{dB}) = \rho(\text{dB}) - \delta_{SNR}(\text{dB})$. This selection guarantees that the average SNR loss on each eigenchannel would be smaller than δ_{SNR}.

6.3 Simulation Results

6.3.1 Capacity Results

To evaluate the performance of the three proposed SNR models described in section III, a 2×2, a 3×3 and a 4×4 MIMO system are simulated, although the proposed design approach is applicable to any $N\times M$ MIMO system. In addition, the capacity of water-filled eigen-MIMO, and of dominant eigenmode transmission, are included in the graphs as benchmarks.

The capacities of a 4×4 MIMO system and the corresponding SNR imposed on the weakest eigenchannel, $SNR_L = E[SNR_{L,\mathbf{H}}]$, are illustrated in Figure 6-2 and Figure 6-3 respectively, for $0.05 \le \gamma \le 1$ and $\rho = 20$ dB. Using the strongest two eigenchannels (i.e., $L = 2$), the maximized capacity with the SNR models III and II yield the same result as of SNR model I for $\gamma >\sim 0.3$. For small values of $\gamma <\sim 0.2$, the capacity with SNR models III and II become within 0.1 bits/channel-use of the optimal water-filled capacity (typically this is the case in a MIMO system with $r \ge 3$). SNR model II achieves the highest capacity among the three SNR models for fixed system parameters γ and L.

As expected, the capacity results are bounded by those obtained via full eigen-MIMO (water-filling) and dominant eigenmode transmission. These results are sensitive to the choice of γ, as demonstrated in Figure 6-2. The same situation holds for the SNR_L results of Figure 6-3. This sensitivity is beneficial. It makes it possible to select the capacity and the corresponding allowable

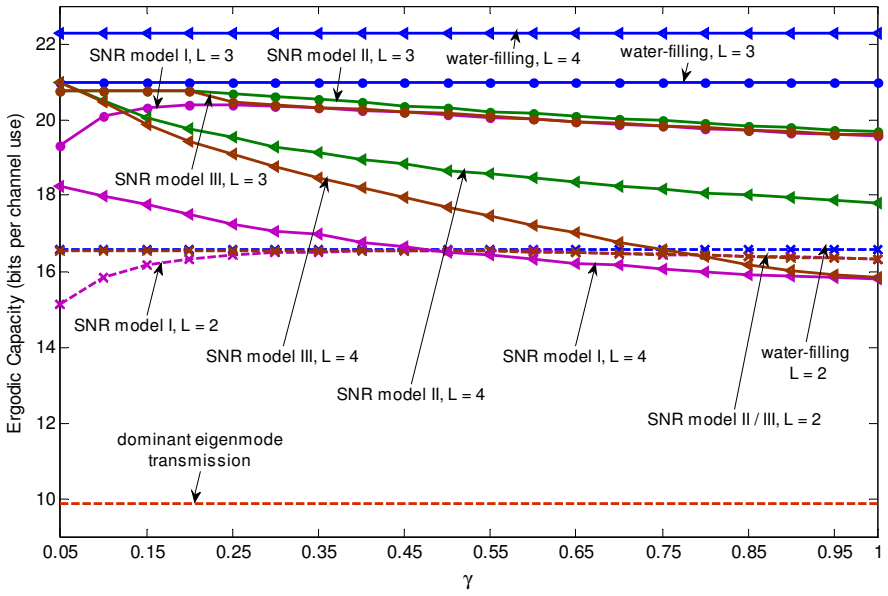

Figure 6-2 The ergodic capacity (in bits per channel use) as a function of γ for a 4×4 MIMO system using different SNR constraint models and $\rho = 20$ dB

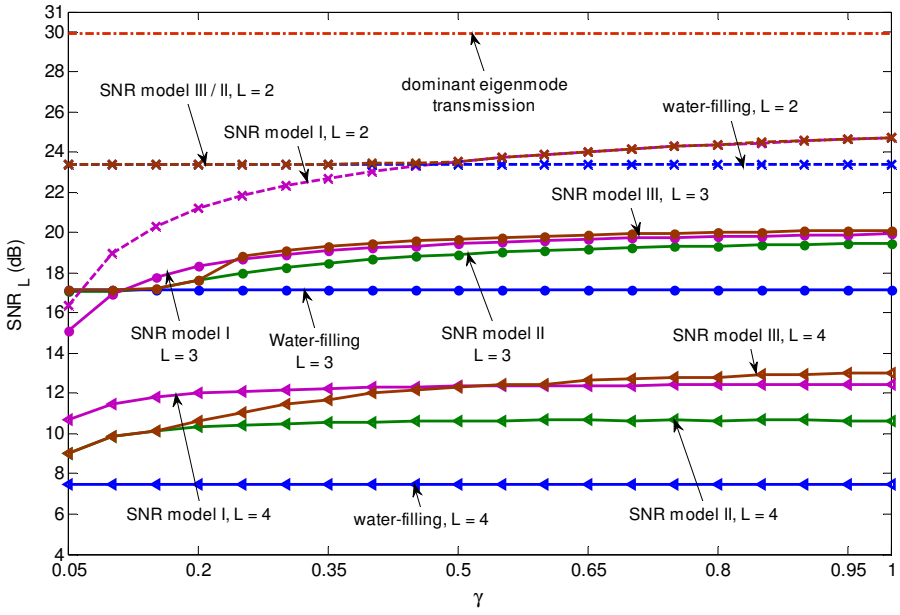

Figure 6-3 The output SNR posed on the weakest eigen-channel, SNR_L in dB, as a function of γ for a 4×4 MIMO system using different SNR constraint models and $\rho = 20$ dB

maximum SNR loss imposed on the weakest eigenchannel, by appropriately choosing γ along with the SNR model.

The information in Figure 6-2 and Figure 6-3 can be represented in another way. In fact, given an SNR model, each γ value constitutes a point in the Capacity– SNR_L plane. Figure 6-4 and Figure 6-5 illustrate the resultant ergodic capacity versus SNR_L for a 4×4 and a 3×3 MIMO systems at a specific receiver SNR, $\rho = 20$ dB, respectively. For a given SNR_L, the capacity obtained under SNR model III outperforms that from SNR models II and I for the same number of used eigenchannels, although capacities with SNR models III and II converge for small values of γ (e.g., $\gamma \le 0.12$ when $L = 4$, and $\gamma \le 0.25$ when $L = 3$ in a 4×4 system). Still, the capacity results of all the SNR models fall below the optimal water-filled eigen-MIMO capacity for the same L. The reason is that,

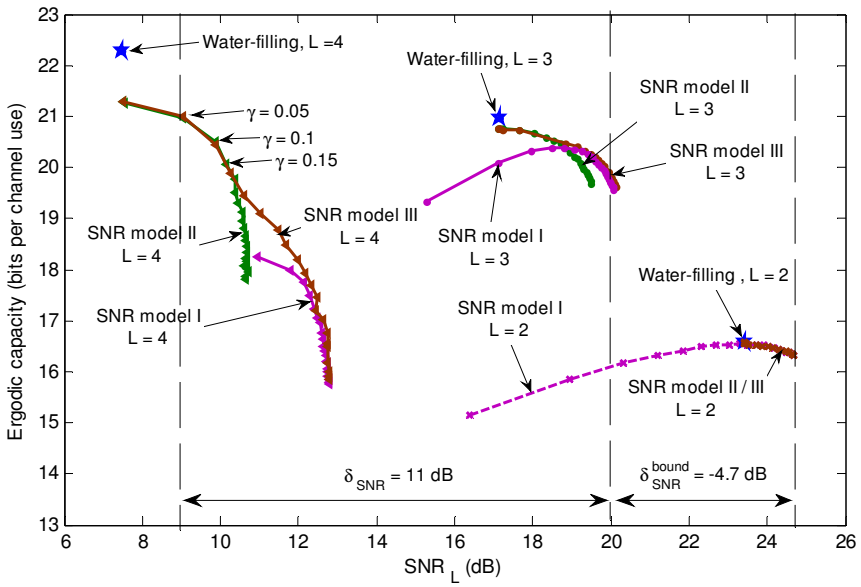

Figure 6-4 The ergodic capacity – SNR_L plot for a 4×4 MIMO system with $\rho = 20$ dB

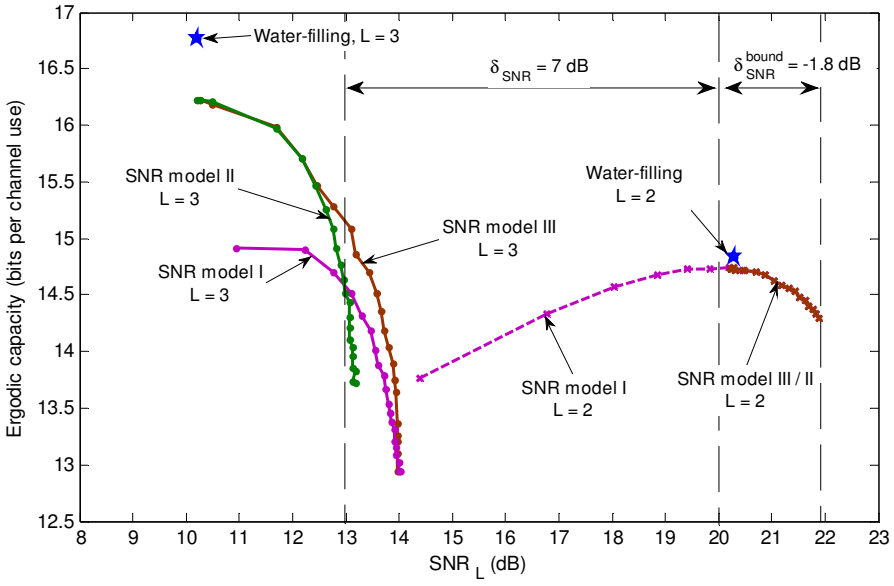

Figure 6-5 The ergodic capacity – SNR_L plot for a 3×3 MIMO system with $\rho = 20$ dB

unlike the water-filling scheme which discards the weakest eigenchannels, all of the available (from the constrained set) L eigenchannels are used within the three presented SNR models at all times (i.e., regardless of the ratio of eigenvalues).

It turns out that none of the power allocation schemes (the three SNR models I / II / III, water-filling with a constrained number of deployed eigenchannels, and dominant eigenmode transmission) maximizes the capacity for all values of ρ and δ_{SNR}. But for a given choice of δ_{SNR}, the maximum capacity is achieved via the presented selection (section 6.2.4) from these different schemes. Some examples are provided in Table 1 for a specific SNR, viz., $\rho = 20$ dB. Details of how the simulations are undertaken for the selection procedure are provided in Appendix 1.

Table 6-1 Selected system configuration for different values of SNR loss δ_{SNR} (dB) for a 4×4 system with $\rho = 20$ dB

δ_{SNR}(dB)	Selected Model	Ergodic capacity (bits per channel use)	SNR_L(dB)	γ
-8	dominant eigenmode trasns.	9.9	29.9	N/A
- 4.7	SNR model I / II / III, $L=2$	16.33	24.7	0.97
- 4	SNR model I / II / III, $L=2$	16.50	24	0.62
-1 – -3	water-filling, $L=2$	16.58	23.42	N/A
0	SNR model III, $L=3$	19.79	20	0.79
1	SNR model III, $L=3$	20.43	19	0.27
2	SNR model III, $L=3$	20.63	18	0.21
3 – 6	water-filling, $L=3$	20.98	17.12	N/A
10	SNR model III, $L=4$	20.98	17.12	N/A
12	SNR model III, $L=4$	21.16	9	0.008
13	water-filling, $L=4$	22.26	7.48	N/A

From Figure 6-4 and Figure 6-5, the dominant eigenmode transmission has the lowest SNR loss, $\delta_{SNR}^{\text{dom}} = -10\log_{10}(E[\lambda_{\max}])$, among all schemes. However, with $L \geq 2$, the SNR loss δ_{SNR} is lower bounded by a value that is achieved by setting $\gamma = 1$ (equal SNRs of all eigenchannels) within any of the SNR models I / II / III, using the first two strongest eigenchannels ($L=2$). This is proved in Appendix 2 and also can be seen from Figure 6-4 and Figure 6-5. Since in this chapter the interest is in using more than one eigenchannel ($L \geq 2$) in order to maximize capacity, so the associated minimum possible value for δ_{SNR} that can be set within the algorithm (selection procedure) is of particular interest and is denoted as $\delta_{SNR}^{\text{bound}}$. The value of $\delta_{SNR}^{\text{bound}}$ depends on the various system parameters such as $r = \min(M,N)$ (e.g., for a 4×4 system, the smallest choice of δ_{SNR} is approximately -4.7dB, whereas in a 3×3 system $\delta_{SNR}^{\text{bound}} \approx -1.8$ dB, and in a 2×2 system $\delta_{SNR}^{\text{bound}} \approx +3.96$ dB). We note that for $\delta_{SNR}^{\text{dom}} \leq \delta_{SNR} \leq \delta_{SNR}^{\text{bound}}$ the only possible scheme is dominant eigenmode transmission.

For different values of the average receiver SNR, ρ, Figure 6-4 and Figure 6-5 would be different set of curves. So for a given δ_{SNR}, the selection depends on ρ. Among the three SNR models, model III yields the highest capacity, as seen in the right half plane given by $SNR_L(\text{dB}) \geq SNR_L^*(\text{dB}) = \rho(\text{dB}) - \delta_{SNR}(\text{dB})$ for the same L. Thus, the selection procedure does not need to include SNR models I and II, and can draw from just: SNR model III; water-filling with an appropriately selected number of deployed eigenchannels; and dominant eigenmode transmission. In addition, with $L = 2$, the highest possible capacity that is achieved to the right of $SNR_L(\text{dB}) \geq SNR_L^*(\text{dB})$, is the same via any of the three SNR models. Therefore, going with SNR model I is preferable when $L = 2$, as it is the simplest.

Figure 6-6, Figure 6-7, and Figure 6-8 are plots of the maximum capacity from selection in a 4×4, a 3×3 and a 2×2 MIMO system, respectively, constrained with different choices of δ_{SNR} for $\rho = 0 - 30$ dB. In general, for larger values of δ_{SNR}, a system with higher capacity is expected. This is because of the trade-off between the capacity and minimum allowable SNR. The described approach offers a configuration that achieves the highest capacity for a maximum SNR loss allowed on the eigenchannels. The position of the capacity between the upper and lower capacity bounds can be controlled by appropriately setting the SNR model along with the parameter characterizing it.

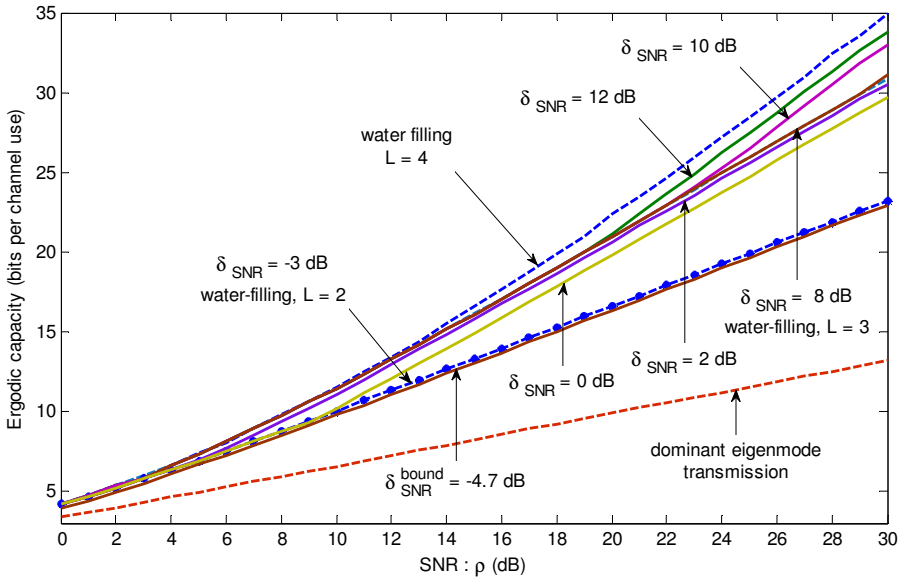

Figure 6-6 The maximum ergodic capacity achieved via selection in a 4×4 MIMO system constrained with different choices of δ_{SNR} for $\rho = 0-30$ dB

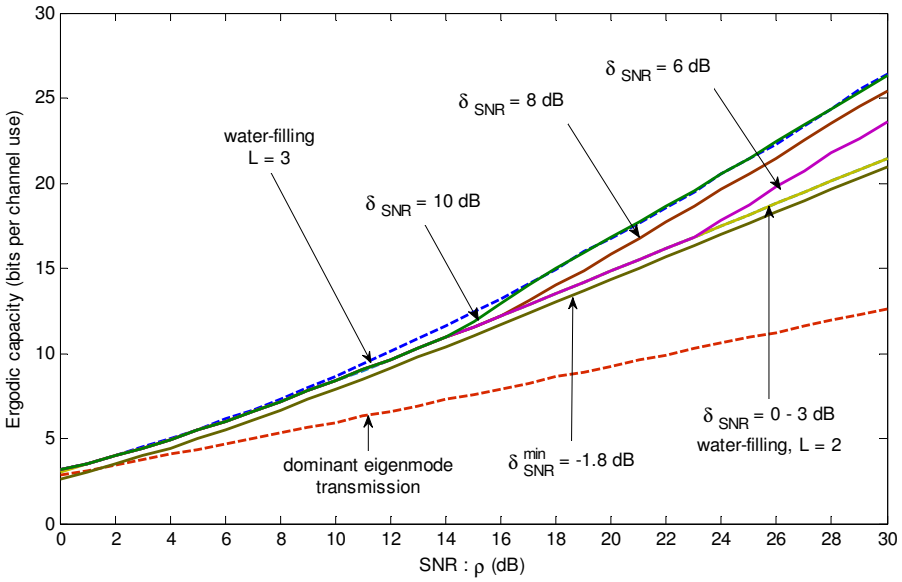

Figure 6-7 The maximum ergodic capacity achieved via selection in a 3×3 MIMO system constrained with different choices of δ_{SNR} for $\rho = 0-30$ dB

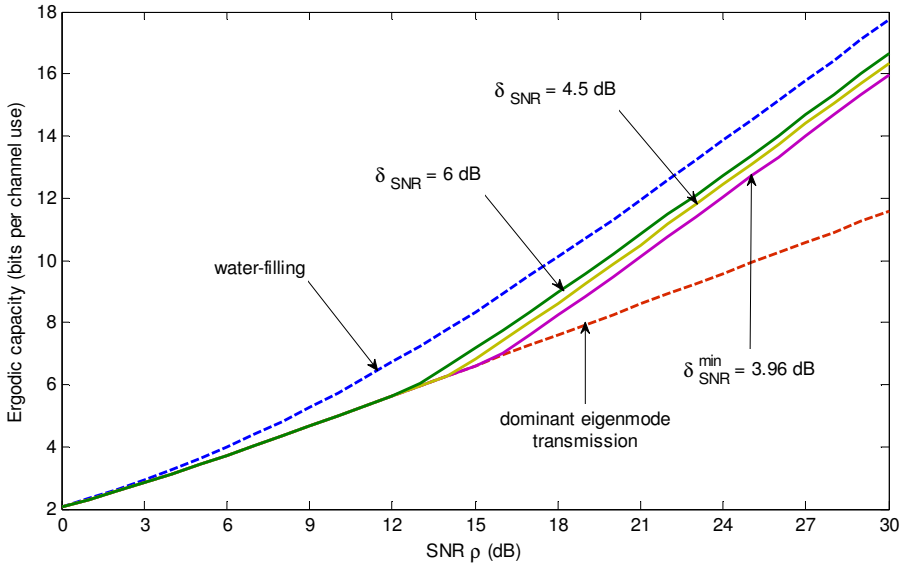

Figure 6-8 The maximum ergodic capacity achieved via selection in a 2×2 MIMO system constrained with different choices of δ_{SNR} for $\rho = 0-30$ dB

Some interesting results from Figure 6-6, Figure 6-7, and Figure 6-8 are:

- In a MIMO system with $r > 2$ and $\delta_{SNR} \geq 0$ dB, the capacity curves obtained via selection (seen from Figure 6-6 for a 4×4 MIMO and Figure 6-7 for a 3×3 MIMO) are upper and lower bounded by the optimal water-filled capacity using all eigenchannels ($L = r$) and by the capacity obtained via water-filling applied to the first two strongest eigenchannels ($L = 2$), respectively. This shows that in MIMO systems with $r > 2$ there is always opportunity to outperform the water-filled eigen-MIMO with $L = 2$, with an allowable SNR loss $\delta_{SNR} \geq 0$ dB.

- For $\delta_{SNR} \geq \delta_{SNR}^{\text{bound}}$ and $r > 2$, the capacity lower bound decreases to the one obtained with the proposed algorithm via any of the SNR models I / II / III with $L = 2$ and $\delta_{SNR} = \delta_{SNR}^{\text{bound}}$.

- In a 2×2 MIMO system (seen from Figure 6-8), the dominant eigenmode transmission is the best scheme for $\delta_{SNR} \le 6$ dB and values of SNR less that $\rho <\sim 13$ dB.

- With no constraint on δ_{SNR}, the dominant eigenmode transmission yields the lower bound on capacity regardless of r (a known result).

6.3.2 Throughput Results

Based on the formula provided in [6.32], the corresponding throughput of a MIMO system with the power allocation from the above approach with fixed modulation can be calculated. For uncoded D-QAM where D denotes the number of points in the signal constellation, the attainable throughput for a single-input single-output channel can be given in terms of the block error rate (BLER) for the block length of *BL* symbols as [6.32]

$$R_{\mathbf{H}} = \log_2 \mathrm{D}.\left[1 - BLER(SNR)\right] = \log_2 \mathrm{D}.\left[1 - SER(SNR)\right]^{BL} \tag{6-31}$$

Here, SER is the probability of symbol error for an AWGN channel with D-QAM modulation and coherent detection, and SNR is the signal-to-noise ratio per symbol. For example, for a uniform, square D-QAM, the SER is approximated as [6.36]

$$SER(SNR) \approx 4(1-\frac{1}{\sqrt{\mathrm{D}}})Q\left(\sqrt{\frac{3SNR}{\mathrm{D}-1}}\right) \tag{6-32}$$

Figure 6-9 illustrates the single AWGN channel throughput penalties for D-QAM versus SNR with coherent detection. The associated throughput for a

single Rayleigh fading channel are also depicted for the examples of 4-QAM, 8-QAM and 16-QAM. For each constellation size, there is a zero-capacity region and a capacity-saturated region. From Figure 6-9, the ergodic capacity-saturated region for 4-QAM occurs for an SNR above about 25 dB. In other words, for values of SNR above 25 dB, the single channel ergodic throughput is 2 bits/sec/Hz, while at low to moderate values of SNR <~ 25 dB, the single channel throughout is lower than 2 bits/sec/Hz and for values of SNR <~ 5 dB, the throughput is almost zero.

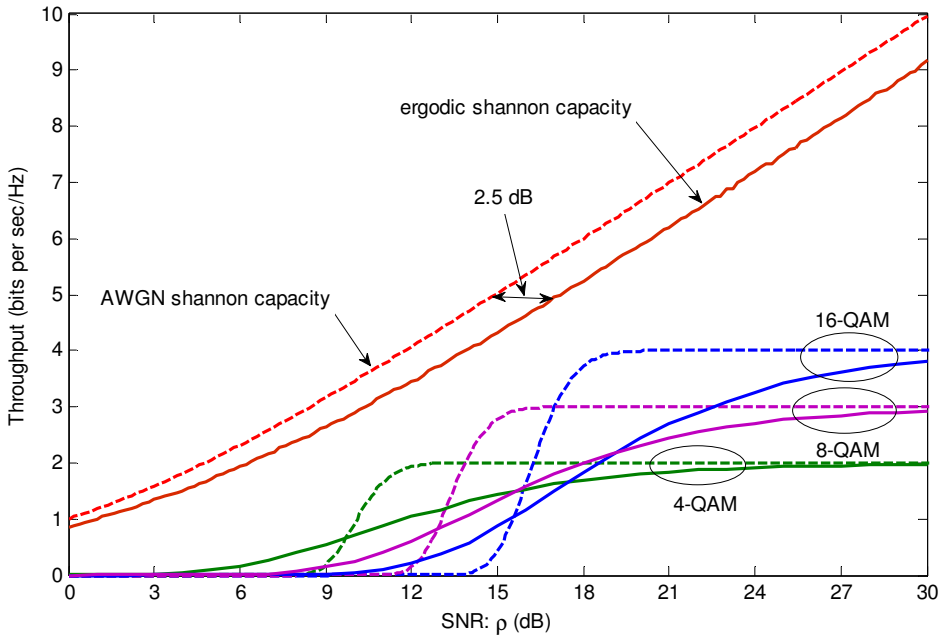

Figure 6-9 The throughput (solid lines) for uncoded data transmission over a single Rayleigh fading channel for the examples of 4-QAM, 8-QAM and 16-QAM. The associated AWGN throughputs (dotted lines) are included for comparision.

Thus in general, the throughput depends on the constellation size of modulation as well as the average SNR of the channel. The same situation holds for MIMO throughput. The MIMO throughput is obtained by summing the

eigenchannel throughputs which are governed by their SNRs. For example, for uniform square D-QAM, the throughput of a MIMO system (with a fixed channel $\mathbf{H}$) is the sum of the throughputs of the eigenchannels:-

$$R_{\mathbf{H}} = \sum_{i=1}^{L} R_i = \sum_{i=1}^{L} \log_2 \mathrm{D}.\left[1 - SER(SNR_i)\right]^{BL}$$
$$= \sum_{i=1}^{L} \log_2 \mathrm{D}.\left[1 - 4(1 - \frac{1}{\sqrt{\mathrm{D}}})Q\left(\sqrt{\frac{3(\rho/M)p_i^{opt}\lambda_i}{\mathrm{D}-1}}\right)\right]^{BL} \quad (6\text{-}33)$$

where $SNR_i = (\rho/M)p_i^{\mathrm{opt}}\lambda_i$ is the output SNR per symbol on the ith eigenchannel with p_i^{opt} obtained from the presented algorithm. Figure 6-10 presents the resulting throughput for a 2×2 MIMO system with 4-QAM and the different power allocation schemes. Here, the SNR region is for the throughput being lower than the limiting $2\log_2 4 = 4$ bits/sec/Hz, but approaching this limit for high SNR.

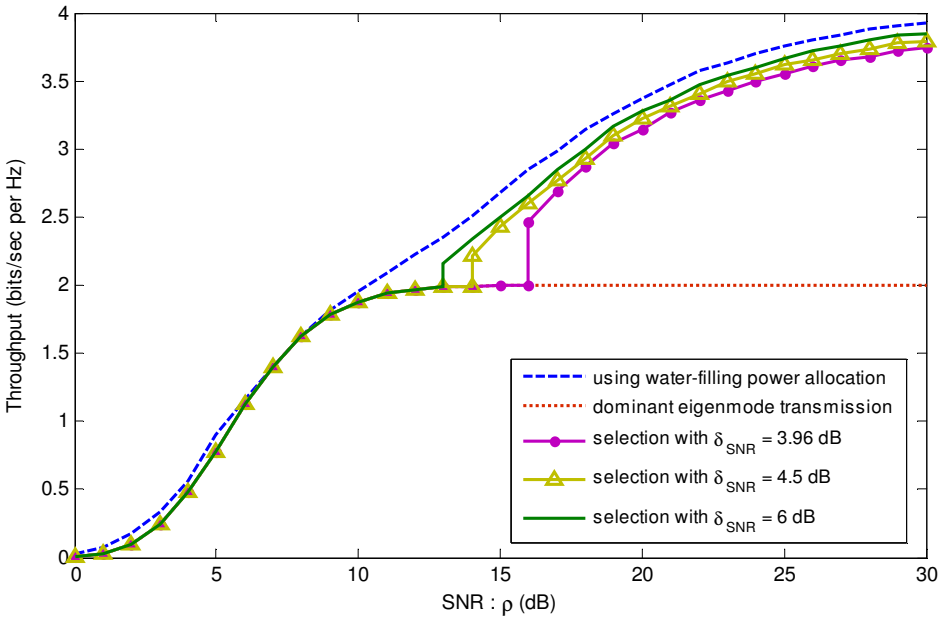

Figure 6-10 The throughput achieved via selection in a 2×2 MIMO system with the power allocation from the presented algorithm and QAM

The power allocation obtained via the presented optimization problem is not optimum for the throughput maximization because the optimization is based on the information-theoretic capacity formula rather than throughput. However, it turns out that the resulting throughput shows similar behavior to the capacity (see Figure 6-6 - Figure 6-8) in the sense that lower throughput (but better SER performance) is realized with a smaller δ_{SNR}. The discontinuities in the throughput curves result from the selection nature of the design approach.

6.3.3 Overall System SER Performance

An important aspect of the above design approach is that the resulting SER performance is better than that of water-filling and using digital modulation (QAM). This is expected (because the water filling does not account for error performance and so the SER is "out of control") and this error performance improvement is quantified below. The design is simulated with independent QAM signals transmitted over each of the eigenchannels. The overall SER can be defined as

$$\mathrm{SER} \overset{\Delta}{=} \frac{\sum_{i=1}^{r} \#\,\text{of correctly detected symbols over eigenchannel } i}{\sum_{i=1}^{r} \#\,\text{of transmitted symbols over eigenchannel } i} \tag{6-34}$$

which is general in the sense that the number of deployed eigenchannels in the summations may be smaller than r. The actual number depends on the type of power allocation scheme used. For example, with the dominant eigenmode transmission, only the first eigenchannel contributes and $r = 1$.

The overall SER results for a 2×2 and a 3×3 system are illustrated in Figure 6-11 and Figure 6-12. Also depicted are the SER curves associated with each of the eigenchannels when water-filling is used. Including these individual eigenchannel performance curves is useful since they illustrate how the overall SER performance of the system deteriorates with the weakest eigenchannel. (With adaptive modulation, using higher modulations for the stronger eigenchannels, the individual SER performance of the higher eigenchannels is worse than using fixed modulation and as a result the overall system SER performance with adaptive modulation is worse than that of using fixed modulation. But the capacity with adaptive modulation is higher than for fixed modulation of course.) Based on the above overall SER metric, water-filling has

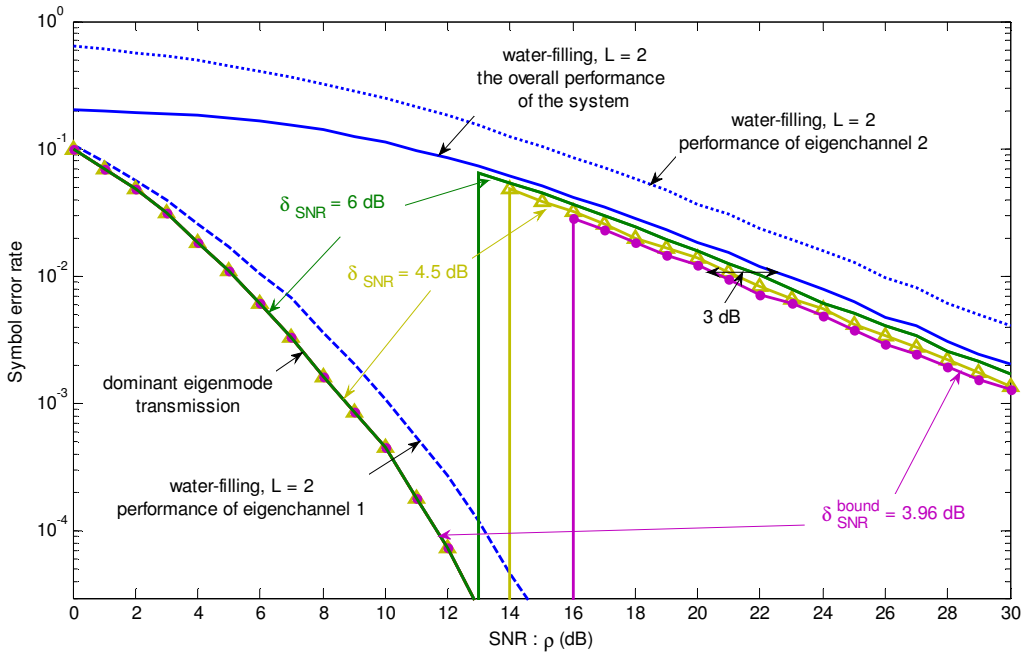

Figure 6-11 The overall SER results of a 2×2 system using different power allocation schemes and QAM

the worst overall SER performance and the dominant eigenmode transmission is best.

As with the throughput, the SER curves have discontinuities. Because of the selection process, several SER curves for different values of δ_{SNR} may lie together for some SNR regions. For example, in Figure 6-11, the curves tagged with $\delta_{SNR}=6\,\mathrm{dB}$, $\delta_{SNR}=4.5\,\mathrm{dB}$, and $\delta_{SNR}=3.96\,\mathrm{dB}$ lie on the SER result of dominant eigenmode transmission for low to moderate SNR (in particular, $\rho<\sim 13\,\mathrm{dB}$ for $\delta_{SNR}=6\,\mathrm{dB}$, and $\rho<\sim 16\,\mathrm{dB}$ for $\delta_{SNR}=3.96\,\mathrm{dB}$). A similar case is also seen from Fig 12 where the curves tagged with various values of $\delta_{SNR}\geq 0\,\mathrm{dB}$, partially lie on

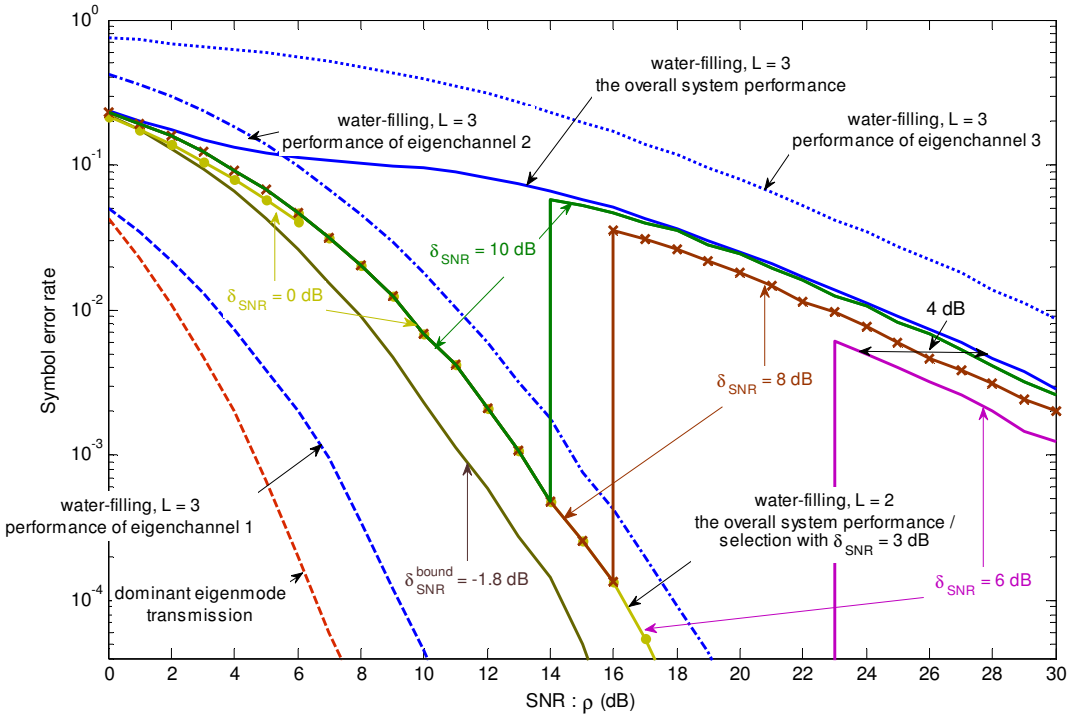

Figure 6-12 The overall SER results of a 3×3 system using different power allocation schemes and QAM

the result of water-filling scheme with $L = 2$ (e.g. $\rho >\sim 6\,\text{dB}$ for $\delta_{SNR} = 0\,\text{dB}$; $\rho \geq 0\,\text{dB}$ for $\delta_{SNR} = 3\,\text{dB}$; and $\rho <\sim 23\,\text{dB}$ for $\delta_{SNR} = 6\,\text{dB}$).

The SNR-constrained design is the reason that the overall SER performance is better than that of eigen-MIMO water-filling. For example, in a 2×2 system with $\delta_{SNR} \approx 4\,\text{dB}$, the difference between SER curves (the error performance curve using the design approach) is almost 3 dB for moderate to high values of SNR ($\rho \geq\sim 16\,\text{dB}$). This improvement is at the expense of a loss of just 0.3 bits/s/Hz in throughput compared to that of water-filled eigen-MIMO (seen from Figure 6-10). The improvement of the SER performance at low values of SNR is clear since the selection scheme chooses the dominant eigenmode transmission as the best scheme among others for $\rho <\sim 16\,\text{dB}$ and $\delta_{SNR} \approx 4\,\text{dB}$ (this holds for all the three cases $\delta_{SNR} \approx 4\,\text{dB}$, $\delta_{SNR} = 4.5\,\text{dB}$ and $\delta_{SNR} \approx 6\,\text{dB}$ when $\rho <\sim 13\,\text{dB}$). Here, there is only a very small loss in throughput (smaller than 0.1 bits/s/Hz for $\rho <\sim 10\ \text{dB}$) relative to that with water-filling power allocation. It turns out that the SNR/SER performance of the overall eigen-MIMO system can be improved while maintaining its high throughput.

As another example, in a 3×3 system with $\delta_{SNR} = 6\,\text{dB}$, the resultant SER with the design approach outperforms that of full eigen-MIMO water-filling by 4 dB at high SNR, $\rho \geq\sim 23\,\text{dB}$. This improvement in SER is at the expense of a loss of 2.5 bits per channel use in the capacity (seen from Figure 6-7) compared to that of water filling (with no error performance constraint). At low to moderate values of SNR, the advantage of the design approach is highlighted since a large

improvement in SER – equivalent to more than 15 dB change in SNR for SER~10^{-3} - is observed. Here the water-filling with $L=2$ is selected as the best scheme and there is only a modest loss in capacity (smaller than 1 bit per channel use for $\rho <\sim 15\,\mathrm{dB}$) relative to that of optimal eigen-MIMO.

The presented design approach makes it possible to configure a reliable system with high capacity (i.e., optimal capacity, suboptimum throughput, and actual capacity not addressed) and manage the trade-off between the capacity (and the throughput which has similar behavior) and error performance. The SNR loss on the eigenchannels can be associated with the overall SER performance. Some form of error correcting coding would be expected in a practical system, but has not been included here. The use and allocation of FEC complicates the optimization of actual capacity, and such an optimization also depends on the implementation detail of the protocol, etc. By using the above design approach, a robust basic configuration for the fixed modulation system can be established. For example, with knowledge of the receiver SNR, desired throughput and SER figures such as Figure 6-10 (for throughput) and Figure 6-11 and Figure 6-12 (for SER) define a basic design configuration with power allocation.

6.4 Summary and Conclusions

In this chapter, an eigen-MIMO system with fixed modulation on a fixed number of eigenchannels is described and analyzed. The configuration allows control of the tradeoff between the throughput (*viz.*, throughput of correctly detected bits that could be achieved with an idealized coding system) and the reliability in the form of uncoded SER performance. A design approach has been

presented. The dependence of the throughput on the average receiver SNR (ρ) and the power allocation over the eigenchannels is determined by simulation. Since there is no exact analytical formulation for throughput, the design basis is the maximization of constrained capacity followed by selection between different power allocation schemes. These schemes comprise the power allocation form: the three presented SNR constraint models; water-filling with an appropriately fixed number of deployed eigenchannels; and dominant eigenmode transmission. The approach yields optimal capacity, but suboptimum throughput. However, the throughput turns out to be high in the sense that its behavior follows the trend of the capacity.

Small MIMO systems are evaluated, having the following performance tie-points. With more than two eigenchannels (e.g., in a 3×3 or a 4×4 system, etc.), the resulting optimum capacity is upper bounded by the full eigen-MIMO (water-filled) capacity, and lower bounded by the capacity obtained via the capacity optimization for the strongest two eigenchannels with an equal SNR constraint, an intuitively expected result. In particular, for a system with a minimum eigenchannel SNR less than ρ (in the notation of the chapter, the maximum allowable SNR loss is $\delta_{SNR} \geq 0\,\mathrm{dB}$), the capacity from the three SNR models can outperform the water-filled capacity applied to the strongest two eigenchannels. For a 2×2 system, the dominant eigenmode transmission is the best choice for an SNR loss $\delta_{SNR} \leq 6\,\mathrm{dB}$ and for a receiver SNR of $\rho <\sim 13\,\mathrm{dB}$. The throughput follows similar behavior to the capacity in the sense that lower throughput (but better SER performance) is realized with a smaller δ_{SNR}. For example, in a 2×2

system with $\delta_{SNR} \approx 4\,\mathrm{dB}$, the difference between SER curves is almost 3 dB for moderate to high values of receiver SNR, viz., $\rho \geq\sim 16\,\mathrm{dB}$. At lower SNRs, the advantage of the design approach is more emphasized since a large improvement in SER is possible with only a very small loss in throughput (smaller than 0.1 bits/s/Hz for $\rho <\sim 10\ \mathrm{dB}$) relative to that with water-filling power allocation. This demonstrates that the SNR/SER performance of the overall eigen-MIMO system can be improved while maintaining its high throughput, by using the presented power allocations.

APPENDICES

Appendix 1

The simulation for the design approach is depicted by the flowchart in Figure 6-13. It has two phases, and both of these are undertaken in an offline mode, so the calculations do not need to be made online, i.e., for maintaining the link operation.

1) In the first phase, the ergodic capacity along with the average SNR on the lowest eigenchannel, SNR_L, associated with each of the power allocation schemes (three SNR model I / II / III; water-filling; and dominant eigenmode transmission) are evaluated for different values of $0 < \gamma \leq 1$ and $L = 2,...,r$. This requires discretizing the interval $(0\ 1]$ into $|\Gamma|$ equal distinct segments (here, Γ denotes the discrete set consisting of $\gamma_k \in \Gamma$ – the value chosen from the *k*th segment – with set size $|\Gamma|$) and calculating the instantaneous capacities and the instantaneous SNRs on the lowest eigenchannel for various discrete values γ_k and L. As an example, $C_{\mathbf{H}}^{\text{model II},\gamma_k,L}$ and $SNR_{L,\mathbf{H}}^{\text{model II},\gamma_k,L}$ (seen from flowchart in Figure 6-13) stands for the instantaneous capacity and the instantaneous SNR on the lowest eigenchannel respectively, associated with the SNR model II using the first L eigenchannels and parameter value γ_k. The ergodic capacities and

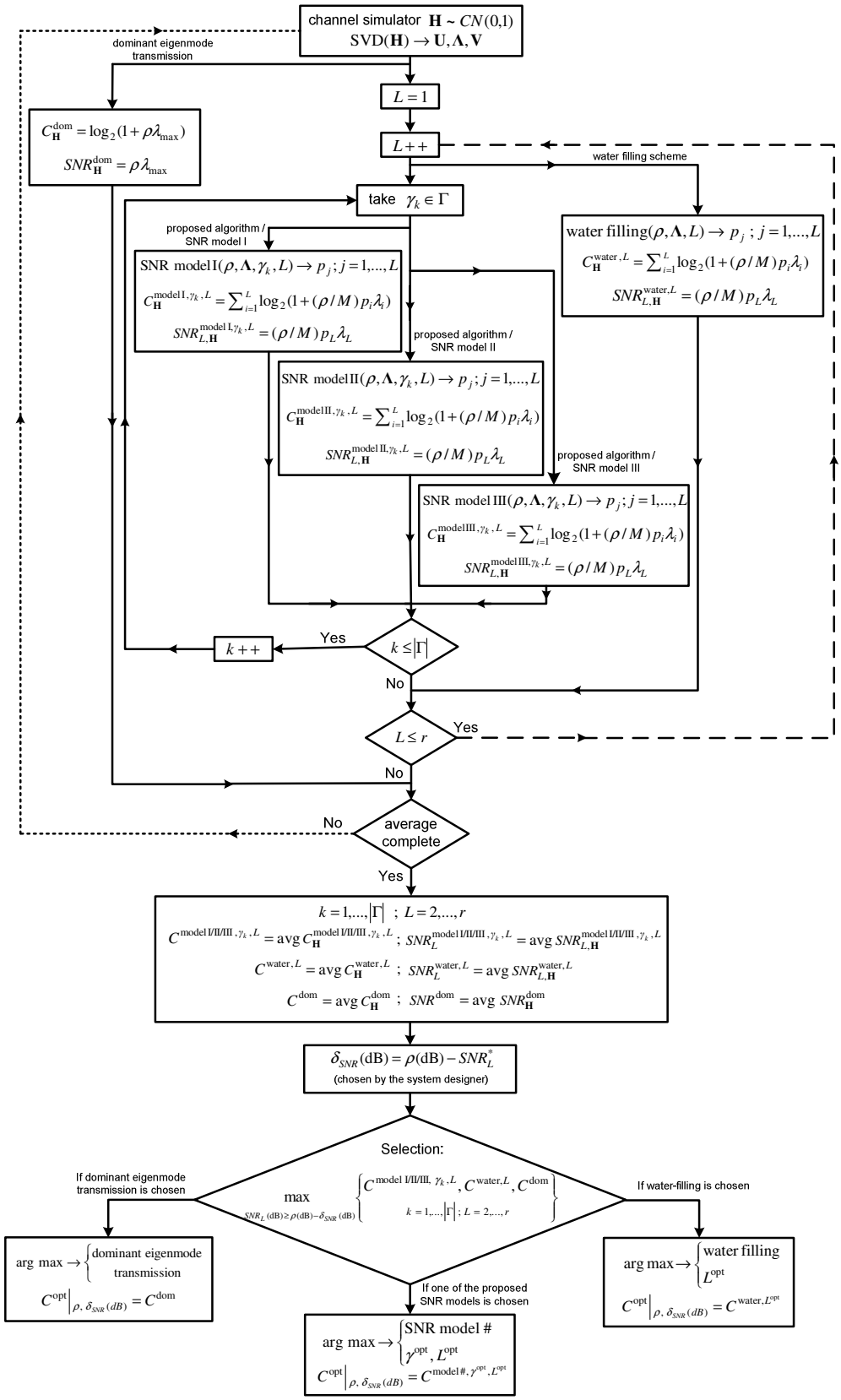

Figure 6-13 Flowchart of simulation

the average SNRs on the lowest eigenchannel are calculated by averaging the instantaneous capacities, e.g. $C^{\text{model II},\gamma_k,L} = \text{avg}\, C_{\mathbf{H}}^{\text{model II},\gamma_k,L}$; $SNR_L^{\text{model II},\gamma_k,L} = \text{avg}\, SNR_{L,\mathbf{H}}^{\text{model II},\gamma_k,L}$. The Capacity– SNR_L plots (Figure 6-4 for a 4×4 MIMO; and Figure 6-5 for a 3×3 MIMO) are the output of the simulations in phase 1. These simulation results are used in the second phase.

2) The designer first chooses a value for the maximum allowable SNR loss on eigenchannels, δ_{SNR}. The choice is according to SER system performance which is required by a coding scheme, and this is traded off against the desired throughput performance. Next, given ρ and δ_{SNR}, the best power allocation scheme is selected using the metric of highest capacity. Based on the type of the selected scheme, the best value for γ_k (γ^{opt}) and/or the optimum number of eigenchannels, L^{opt}, are provided as the output of second phase of simulation. Once the best scheme (along with parameters γ^{opt} and L^{opt}) is identified, the designer can deploy the power allocation for online operation from the selected best scheme.

Appendix 2

Here it is shown that δ_{SNR} (in dB) cannot be smaller than a specific value that is achieved by setting $\gamma = 1$ (equal SNRs of all eigenchannels) within any of the presented SNR models applied to the first two strongest eigenchannels.

Since $SNR_L^* = \rho(\text{dB}) - \delta_{SNR}(\text{dB})$, the case that maximizes SNR_L^* is sought. For an $N \times M$ MIMO system with $r = \min(M, N) \geq 2$, it is possible to choose from a different number of eigenchannels, $2 \leq L \leq r$. The maximum of $SNR_{L,\mathbf{H}} = (\rho / M) p_L \lambda_L$ is achieved for the largest possible values of p_L and λ_L. Since $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$, the largest possible value for λ_L is λ_2. It is likely that p_L is the largest when the total power is distributed between only the first two strongest eigenchannels. This means that the largest SNR_L is associated with the case with $L = 2$, i.e. $\arg\max_L SNR_{L,\mathbf{H}} = 2$. Now the $SNR_{L=2,\mathbf{H}}$ maximization problem can be written as

$$\begin{aligned} &\underset{p_1, p_2}{\text{maximize}}\ SNR_{2,\mathbf{H}} = \frac{\rho}{M} p_2 \lambda_2 \\ &\text{subject to}\ \ SNR_{2,\mathbf{H}} \leq SNR_{1,\mathbf{H}} \\ &\qquad p_1 + p_2 = M \end{aligned} \tag{6-35}$$

Equivalently, (6-35) can be rewritten as

$$\begin{aligned} &\underset{p_1, p_2}{\text{minimize}}\ f_1(p_1, p_2) = -\frac{\rho}{M} p_2 \lambda_2 \\ &\text{subject to}\ \ f_2(p_1, p_2) = \frac{\rho}{M} p_2 \lambda_2 - \frac{\rho}{M} p_1 \lambda_1 \leq 0 \\ &\qquad p_1 + p_2 = M \end{aligned} \tag{6-36}$$

where f_1, f_2 are convex real functions of p_1 and p_2. Following the same procedure described in section 6.2.2, the inequality constraints in (6-36) can be made implicit in the objective as ($t \to \infty$)

$$\begin{aligned} &\underset{p_1, p_2}{\text{minimize}} -t\frac{\rho}{M} p_2\lambda_2 - \log_{10}\left(\frac{\rho}{M} p_1\lambda_1 - \frac{\rho}{M} p_2\lambda_2\right) \\ &\text{subject to} \quad p_1 + p_2 = M \end{aligned} \tag{6-37}$$

Next, by eliminating the equality constraint, the problem (6-34) is reduced to

$$\underset{p_2}{\text{minimize}}\, \varphi(p_2) = -t\frac{\rho}{M} p_2\lambda_2 - \log_{10}\left(\frac{\rho}{M}(M - p_2)\lambda_1 - \frac{\rho}{M} p_2\lambda_2\right) \tag{6-38}$$

Finally the value of p_2 that minimizes $\varphi(p_2)$ is obtained by setting $\lim_{t\to\infty} \frac{\partial \varphi(p_2)}{\partial p_2} = 0$. As a result

$$p_2 = \lim_{t\to\infty}\left(\frac{M\lambda_1}{\lambda_1 + \lambda_2} - \frac{1}{t(\rho/M)\lambda_2}\right) = \frac{M\lambda_1}{\lambda_1 + \lambda_2} \quad , \quad p_1 = \frac{M\lambda_2}{\lambda_1 + \lambda_2} \tag{6-39}$$

With these values of p_1 and p_2, $SNR_{1,\mathbf{H}} = SNR_{2,\mathbf{H}}$ which is satisfied by setting $\gamma = 1$ within any of the presented SNR models I / II / III.

REFERENCE LIST

[6.1] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *J. of Wireless Personal Commun. - Springer*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[6.2] V. Tarokh, A.F. Naguib, N. Seshadri and A.R.Calderbank, “Combined array processing and space-time coding," *IEEE Trans. Inform. Theory*, Vol. 45, No.4, pp. 1121-1128, May 1999.

[6.3] G. Lebrun, J. Gao, and M. Faulkner, “MIMO transmission over a time-varying channel using SVD,” *IEEE Trans. Wireless Commun.,* Vol. 4, No. 2, pp. 757-764, Mar. 2005.

[6.4] J. Yang and S. Roy, “On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, Vol 42. No 12, pp. 3221-3231, Dec. 1994.

[6.5] W. Jang, B. R. Vojcic, R. L. Pickholtz, “Joint transmitter-receiver optimization in synchronous multiuser communications over multi path channels," *IEEE Trans. Commun.,* Vol.46, No 2. pp. 269-278, Feb. 1998.

[6.6] A. Paulraj, R. Nabar and D. Gore, *Introduction to space-time wireless communications*, Cambridge University press, 2003.

[6.7] G. Raleigh and J. Cioffi, “Spatio-temporal coding for wireless communications," *IEEE Trans. Commun.*, Vol. 46. No. 3, pp. 357-366, Mar. 1998.

[6.8] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Trans. Telecommun.*, Vol. 10, No. 6, pp. 585-595, Nov.-Dec. 1999.

[6.9] R. Vaughan and J. Bach Andersen, *Channels, Propagation and Antennas for Mobile Communications*, The Institution of Electrical Engineers, 2003.

[6.10] A. Scaglione, G.B. Giannakis and S. Barbarossa, “Redundant filter bank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, Vol. 47, No. 7, pp. 1988-2006, Jul. 1999.

[6.11] H. Sampath, P. Stoica, and A. Paulraj, “Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion,” *IEEE Trans. Commun.,* Vol. 49, No. 12, pp. 2198-2206, Dec. 2001.

[6.12] V. Tarokh, N. Seshadri and A. R. Calderbank, “Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, Vol 44, No.2, pp. 744-765, Mar. 1998.

[6.13] J. Yang and S. Roy, “Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Info. Theory*, Vol. 40, No. 5, pp. 1334-1347, Sep. 1994.

[6.14] L. Vandendorpe, J. Louveaux, B. Maison and A. Chevreuil, “About the asymptotic performance of MMSE MIMO DFE for filter-bank based multicarrier transmission," *IEEE Trans. Commun*, Vol 47, No. 10, pp. 1472-1475, Oct. 1999.

[6.15] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, “Optimal designs for space-time linear precoders and decoders,*” IEEE Trans. Signal Process.*, vol. 50, pp. 1051-1064, May 2002.

[6.16] N. Wang and S. D. Blostein, “Power loading for CP-OFDM over frequency-selective fading channels,” in Proc*. GLOBECOM*, Vol. 4, pp. 2305-2309, Dec. 2003.

[6.17] A. Scaglione, S. Barbarossa and G.B. Giannakis, “Filter bank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inform. Theory*, Vol. 45, No. 3, pp. 1019-1032, Apr. 1999.

[6.18] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, “Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization,” *IEEE Trans. Signal Process.*, Vol. 51, No. 9, pp. 2381–2401, Sep. 2003.

[6.19] Y. Ding, T. N. Davidson, Z. Q. Luo, and K. M. Wong, “Minimum BER block precoders for zero-forcing equalization,” *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2410–2423, Sep. 2003.

[6.20] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, “Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints,” *IEEE Trans. Signal Processing*, Vol. 52, No. 5, pp. 1179-1197, May 2004.

[6.21] M. A. Khojastepour, W. Xiaodong, and M. Madihian, “Design of multiuser downlink linear MIMO precoding systems with quantized feedback,” *IEEE Trans. Veh. Technol.,* Vol. 58, No. 9, pp. 4828-4836, Nov. 2009.

[6.22] B. Mielczarek, and W. A. Krzymień, “Influence of CSI feedback errors on capacity of linear multi-user MIMO systems,” in Proc. *IEEE Veh. Technol. Conf.* pp. 2043-2047, Apr. 2007.

[6.23] N. Jindal, “MIMO broadcast channels finite-rate feedback,” *IEEE Trans. Inform. Theory,* Vol. 52, No. 11, pp. 5045-5060, Nov. 2006.

[6.24] B. Mielczarek and W. A. Krzymień, “Comparison of partial CSI encoding methods in multi-user MIMO systems,” *J. of Wireless Personal Commun.- Springer,* Vol. 52, No. 1, Jan. 2010.

[6.25] S. P. Shenoy, L. Ghauri, and D. T. M. Slock, “Optimal precoding and MMSE receiver designs for MIMO WCDMA,” in Proc. *IEEE Veh. Technol. Conf.* pp. 893-897, May 2008.

[6.26] T. Yoo and A. Goldsmith, “Capacity and power allocation for fading MIMO channels with channel estimation error,” *IEEE Trans. Inform. Theory*, Vol. 52, No. 5, pp. 2203-2214, May. 2006.

[6.27] Z. Jian-Kang, K. M. Wong, J. Wenwen, and A. B. Gershman, “Joint design of transceivers for multiple-access channels using MMSE decicion feedback detection,” *IEEE Trans. Veh. Technol.,* Vol. 60, No. 8, pp. 3792 – 3804, Oct. 2011.

[6.28] S. A. Banani, and R. G. Vaughan, “The effect of training-based channel estimation on the capacity of closed-loop MIMO systems with imperfect CSI feedback,” in Proc. *IEEE Veh. Technol. Conf.*, pp. 1-5, Sep. 2010.

[6.29] A. Svensson, "An introduction to adaptive QAM modulation schemes for known and predicted channels," *IEEE Proceedings,* Vol. 95, No. 12, pp. 2322-2336, Dec. 2007.

[6.30] S. A. Banani, and R. G. Vaughan, “Throughput and SER trade-off in eigen-MIMO with fixed modulation,” submittied to *IEEE Trans. Veh. Technol.,* Nov. 2011.

[6.31] S. A. Banani, and R. G. Vaughan, “Adaptive RS coded modulation for practicable capacity maximization in eigen-MIMO,” submitted to *IET Communications,* Sep. 2011.

[6.32] S. Catreux, P. F. Driessen and L. J. Greenstein, “Data throughputs using multiple-input multiple-output (MIMO) techniques in a noise-limited cellular environment,” *IEEE Trans. Wireless Commun.*, Vol. 1, No. 2, pp. 226 – 235, Apr 2002.

[6.33] A. J. Goldsmith and S.-G Chua, “Variable-rate variable-power MQAM for fading channels,” *IEEE Trans. Commun.*, Vol. 45, pp. 1218–1230, Oct. 1997.

[6.34] L. Teen Ong, M. Shikh-Bahaei, and J. A. Chambers, “Variable rate and variable power MQAM system based on Bayesian bit error rate and channel estimation techniques,” *IEEE Trans. Commun.,* Vol. 56, No. 2, Feb. 2008.

[6.35] S. Boyd, and L. Vandenberghe, *Convex optimization,* Cambridge University Press, 2004.

[6.36] J. G. Proakis, Digital communications, 4th ed. New York: McGraw-Hill, 2001.

7: ADAPTIVE RS CODED MODULATION FOR PRACTICABLE CAPACITY MAXIMIZATION IN EIGEN-MIMO

Adaptive coded modulation (ACM) is for spectral efficiency and robustness in digital transmission schemes, e.g., [7.1]-[7.16], and works by adapting the transmission parameters (transmitted power, modulation rate, coding rate, spreading factor, etc.) to match a transmission rate to the current channel conditions. Some examples are as follows. In [7.2]-[7.3], trellis coded adaptive MQAM is considered where the system is optimized for maximal average spectral efficiency while maintaining the instantaneous bit-error-rate (BER) below a target value, BER_0. With perfect CSI, this design shows an effective coding gain of 3 dB relative to uncoded adaptive MQAM, for a simple four-state trellis code (higher coding gains are possible with a more complex trellis) in SISO fading channels. In [7.4]-[7.5], pilot symbol assisted modulation (PSAM) is used for channel estimation, and the impact of the imperfect CSI on trellis coded modulation is investigated. In [7.6] a MIMO system is analyzed where perfect but delayed CSI (i.e., outdated) is applied. No pilots are used, and the BER constraint is subject to instantaneous BER and average BER, respectively.

Examples of ACM schemes based on low-density parity-check (LDPC) codes can be found in, e.g., [7.7]-[7.12]. The asymptotic performance of LDPC codes, optimized for multi-level coding, is analyzed in [7.9]. LDPC codes are

promising candidates for channel codes because of their capacity-approaching performance [7.10]. However, ACM based on LDPC typically leads to high system complexity because many encoders/decoders (CODECs) for dedicated LDPC codes are needed to support a wide range of transmission rates. As a result, achieving flexible throughput is not currently attractive for the large required chip size. Complexity reduction through using one CODEC is also considered in [7.11]-[7.12].

There are many other specialized schemes (e.g. [7.13]-[7.15] based on turbo and convolutional coding), too numerous to include here. Despite the large body of literature on MIMO capacity, there are practical aspects which justify further investigation since they can lead to new capacity realization techniques. It appears that no previous work has formally considered adaptive Reed-Solomon (RS) coding for improving the spectral efficiency [7.16]. RS-ACM schemes are instead typically presented for reliability enhancement, e.g. [7.17]-[7.19]. RS codes are widely applied in digital communications, digital media, and memory storage systems. They also maintain their presence in newly developed communications systems such as WiMAX and DVB. Owing to their ability to correct burst errors, RS codes are still preferred where low delay and robust communication are high priorities [7.20].

In this work [7.21], an adaptive coded 2^D-QAM scheme based on RS codes is presented for spectral efficiency improvement in eigen-MIMO. The RS codes closed form error formula is convenient for optimizing the modulation constellation size(s), code rate(s) and the power allocation over eigenchannels,

in order to achieve the highest instantaneous practicable capacity (data throughput). The optimization is here constrained by the total transmitter power. The formulation is new, reveals interesting capacity behavior, and allows comparison between inner and outer coding quantifying the improvement in practicable capacity from the inner-coding architecture. In the simplest, but topical example of a 2×2 system, inner coding shows an effective coding gain of 4.5 dB over uncoded adaptive modulation with water-filling power allocation.

This chapter extends the work in [7.16] which considers a system with RS coding and fixed common modulation over eigenchannels. Accurate approximations are also given here for the overall BER using the presented ACM scheme for a single Rayleigh fading channel and a 2×2 system, via nonlinear least squares curve-fitting. These error curve expressions are useful for quick assessment of the system error performance. Here, use is made of these error curves for selecting between different system configurations (different combinations of CODECs constrained with fixed number of deployed eigenchannels, and power allocation schemes. This includes the presented ACM applied to dominant eigenmode transmission) so as to achieve the highest practicable capacity for a given average or instantaneous output BER. This selection procedure is presented in the simulation section below.

The problem considered here is different to previous works on adaptive modulation, e.g., [7.1]-[7.6], as:

1) The adaptive schemes in [7.1]-[7.6] are to maximize the average spectral efficiency (weighted sum of the throughput associated with different

modulations), whereas the algorithm here maximizes the instantaneous eigen-MIMO throughput.

2) Here, the total power at the transmitter is fixed for each MIMO channel realization, which favors regulatory arrangements; whereas in [7.1]-[7.6], variable power transmission is considered (at each channel realization, the power at the transmitter changes according to the channel conditions although the average transmit power is fixed). Up to now, the variable transmit power schemes are mostly designed for single antenna systems channels as in [7.1]-[7.3] and single-input multiple-output systems as in [7.4]-[7.5]. With eigen-MIMO, the challenging problem of power allocation over eigenchannels arises. Perhaps because of this, previous treatments of power allocation in MIMO have been constrained by fixed total transmit power for each channel realization [7.6], [7.22]-[7.23].

3) Unlike in [7.1]-[7.6], the algorithm presented here is similar to [7.24]-[7.25] in the sense the capacity maximization problem is initially formulated (that is, optimization excluding selection between different system configurations presented in the simulation section) without constraint on the output BER. However, including such a constraint is a straightforward extension of the algorithm presented here. The selection procedure presented below addresses capacity maximization with a target (i.e., a constrained) output BER. An average or an instantaneous BER can be targeted, both of which are treated here. It is worth noting that, as an alternative to the presented selection procedure, it is possible to initially reformulate the capacity maximization problem to be subject to an instantaneous BER, a topic for future work.

The rest of the chapter is organized as follows. Section 7.1 describes the closed-loop MIMO system model. The problem of adaptive RS coded 2^D-QAM is addressed in section 7.2, with the simulation results presented in section 7.3 and concluded in section 7.4.

7.1 Closed-Loop MIMO System Model

As in Figures 5.1. and 6.1, and explained again here for convenience, the MIMO channel is a quasi-static, flat block-fading channel with M transmit and N receive antennas. The channel is modeled by a random (fading) distribution which remains static over a fading block cycle (starting from the training transmission, including the channel estimation, quantization, codeword feedback, calculation of the new beamformer, and detection of the received symbols), but is independent across different blocks. With perfect channel knowledge at the transmitter, improvement of the throughput (practicable capacity) is the focus. A transmit filter (beamformer) comprising weights $\mathbf{W}$ is shown in Figure 7-1. At each block, the $M \times 1$ data vector (with independent elements) to be transmitted, $\mathbf{x}$, is multiplied by the $M \times M$ weights $\mathbf{W}$, before transmission. Assuming perfect timing, etc., in the usual way, the MIMO model is

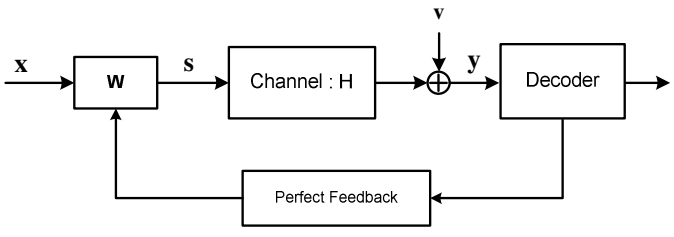

Figure 7-1 A closed-loop MIMO system with a transmit beamformer for using channel knowledge which is perfectly fed back from the receiver

$$\mathbf{y} = \sqrt{\frac{\rho}{M}}\mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{n} = \sqrt{\frac{\rho}{M}}\mathbf{H}\mathbf{s} + \mathbf{n} \tag{7-1}$$

where $\mathbf{y}$ is the $N \times 1$ received signal vector, ρ is the known average SNR at each receive antenna branch, $\mathbf{H}$ is the channel matrix with $(\mathbf{H})_{i,j} \sim CN(0,1)$, and $\mathbf{n} \sim CN(0,1)$ is the additive noise. The total transmit power constraint is expressed as $\mathrm{tr}\,\mathbf{R}_{\mathbf{ss}} = E\,\mathrm{tr}\,\mathbf{s}\mathbf{s}^H = M$ (note this is different to the variable power transmissions of cellular system terminals). Following singular value decomposition, the channel matrix $\mathbf{H}$ with rank $r = \min(M, N)$ is written as $\mathbf{H} = \mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{V}^H$ where $\mathbf{V}$ and $\mathbf{U}$ are the unitary matrices containing the corresponding input and output singular vectors, respectively, and $\boldsymbol{\Lambda}^{1/2}$ is a non-negative $N \times M$ diagonal matrix with ith diagonal element as $\lambda_i^{1/2}$ (the positive square root of ith eigenvalue of the gram matrix of $\mathbf{H}$), with eigenchannel gains $\lambda_i \geq \lambda_{i+1}$. So $\mathbf{W} = \mathbf{V}$ is the set of transmit weights, and $\mathbf{U}^H$ contains the receiver weights, and the overall transmission relationship is

$$\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{y} = \sqrt{\frac{\rho}{M}}\boldsymbol{\Lambda}^{1/2}\mathbf{x} + \tilde{\mathbf{n}} \tag{7-2}$$

where $\tilde{\mathbf{n}} = \mathbf{U}^H\mathbf{n} \sim CN(0,1)$. Since $\mathbf{s} = \mathbf{W}\mathbf{x} = \mathbf{V}\mathbf{x}$ and $E\,\mathrm{tr}\,\mathbf{s}\mathbf{s}^H = M$, then $\mathrm{tr}\,\mathbf{R}_{\mathbf{xx}} = \mathrm{tr}\,E\,\mathbf{x}\mathbf{x}^H = M$. Therefore, the MIMO channel is transformed into r parallel eigenchannels with unequal gains and the capacity is the sum of eigenchannel capacities governed by their SNRs (for the ith eigenchannel, the SNR is $(\rho/M)\,p_i\lambda_i$, where $p_i = [\mathbf{R}_{\mathbf{xx}}]_{i,i}$). The above formulation is written in a general

format in the sense that the data to be transmitted over each of the eigenchannels can be either uncoded or selected from a modulation and a coding set comprising a set of RS codes with different code rates, in order to match a transmission rate and to current channel conditions. Furthermore, only the first $1 \le \bar{r} \le r$ strongest eigenchannels are deployed and the remaining $r - \bar{r}$ weakest eigenchannels are unused. In the formulation here, $\bar{r}$ is chosen and left fixed, and so the number of deployed eigenchannels is not part of the optimization.

The different arrangements of the CODECs are depicted in Figure 7-2. For the outer coding design (Figure 7-2 (b)), one RS code is assigned to the overall

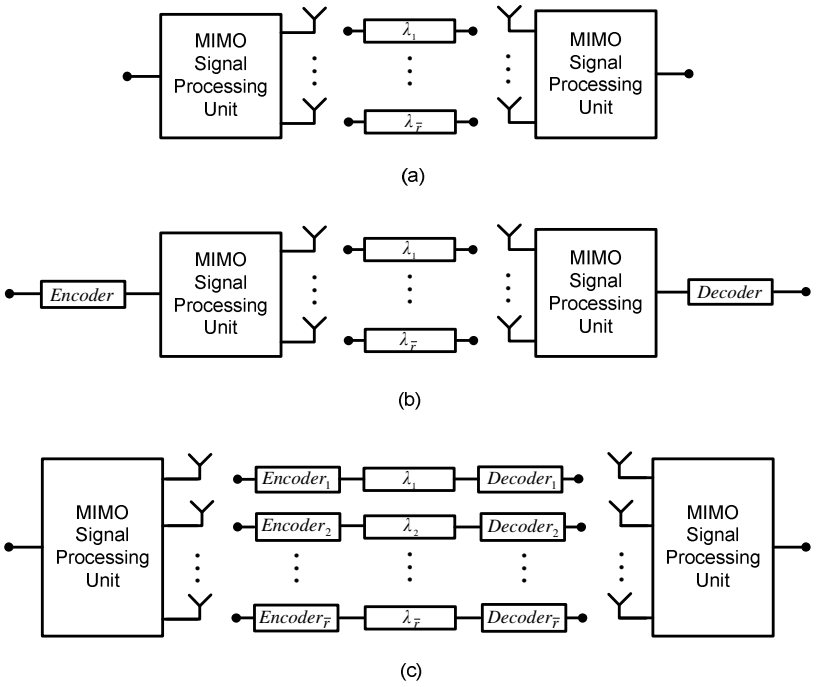

Figure 7-2 The use of the first $\bar{r}$ strongest eigenchannels with (a) the uncoded transmission in MIMO system; (b) the outer coding arrangement of encoder/decoder in MIMO system; (c) the inner coding arrangement of encoders/decoders in MIMO system

serial data, and here a common constellation size and code rate is assigned for all of the $\bar{r}$ eigenchannels. For the inner coding design (Figure 7-2 (c)), a constellation and code rate is assigned to each eigenchannel for each channel realization so as to achieve the highest possible MIMO throughput for a given receive SNR, ρ. The inner coding has high complexity since at each channel realization, the SNR for each eigenchannel is required in order to select the constellation sizes and code rates, and the deployment requires a capacity overhead. Nevertheless, an "eigen set" of RS CODECs are relatively straightforward to implement, considering the inherent high complexity of eigen-MIMO.

RS (n_c, k_c) codes are nonbinary cyclic codes with symbols of $\mathrm{D} > 2$ bits for which $0 < k_c < n_c < 2^{\mathrm{D}} + 2$, where k_c is the number of data symbols being encoded and n_c is the total number of symbols in the encoded block. Here, conventional RS (n_c, k_c) codes are considered, as [7.20]

$$(n_c, k_c) = (2^{\mathrm{D}} - 1, 2^{\mathrm{D}} - 1 - 2t_c) \tag{7-3}$$

where t_c is the symbol-error correcting capability of the code with $n_c - k_c = 2t_c$ the number of parity symbols. Requiring correction of at least one symbol error, the code can provide $(n_c - 1)/2$ different rates

$$R_c = \frac{k_c}{n_c} \;\; ; \;\; k_c = 1,3,...,n_c - 2 \tag{7-4}$$

Denoting $R_c^{(i)}$ as the code rate associated with the selected RS code on the ith eigenchannel, it is possible to select between uncoded data or RS coded data associated with different constellation sizes to obtain the optimal code rates $(R_c^{(1)}, R_c^{(2)}, ..., R_c^{(\bar{r})})^{\text{opt}}$ in order to maximize the instantaneous practicable capacity.

The power allocation over eigenchannels also needs addressing. It is constrained by the total of the powers allocated to all the deployed eigenchannels, recalled as $\sum_{i=1}^{\bar{r}} p_i = \text{tr}\,\mathbf{R}_{\mathbf{xx}} = M$, where $p_i = [\mathbf{R}_{\mathbf{xx}}]_{i,i}$. The information-theoretic capacity for a given channel realization, $C_{\mathbf{H}} = \sum \log_2(1 + (\rho / M) p_i \lambda_i)$, is maximized by the water-filled power allocated to the ith eigenchannel. However, for implementation, where digital techniques such as QAM (instead of Gaussian signals), and finite block lengths (instead of infinitely long codes), etc., will degrade the capacity from the Shannon limit to the practicable possibilities of a digital link, the optimality of the water-filling scheme is no longer guaranteed [7.26]. This motivates the problem of joint power allocation and code rate adaptation for the practicable capacity maximization in an eigen-MIMO with the total input power constraint. The optimal outer coding design is also the special case of inner coding design where $R_c^{(1)} = R_c^{(2)} = ... = R_c^{(\bar{r})}$.

7.2 Adaptive Reed-Solomon Coded Modulation

The capacity of a sample realization of the channel is first addressed. The ergodic capacity is the ensemble average of the capacity achieved when the

optimization is performed for each realization of $\mathbf{H}$, denoted $R = E\,R_{\mathbf{H}}$. (Note that the symbol R here is different to the symbol $\mathbf{R}$ which is a covariance matrix.)

The attainable normalized throughput (practicable capacity) in bits/sec/Hz for the RS-coded data on ith eigenchannel can be given in terms of the block error rate (BLER) as

$$R_{\mathbf{H}}^{(i)} = R_c^{(i)}\,\mathrm{D}_i.\left[1 - BLER_c(SNR^{(i)})\right] \tag{7-5}$$

where $BLER_c^{(i)}$ is the probability of block errors (block-error rate) associated with RS coded $2^{\mathrm{D}_i}-\mathrm{QAM}$ data on the ith eigenchannel with coherent detection. With hard decision decoder, $BLER_c^{(i)}$ (codeword error probability) can be given as

$$BLER_c^{(i)}(P_{\mathrm{D}_i}(SNR^{(i)})) \approx \sum_{j=t_c^{(i)}+1}^{n_c^{(i)}} \binom{n_c^{(i)}}{j} (P_{\mathrm{D}_i})^j (1-P_{\mathrm{D}_i})^{n_c^{(i)}-j} \tag{7-6}$$

where P_{D_i} is the probability of symbol error of uncoded $2^{\mathrm{D}_i}-\mathrm{QAM}$ modulation. For example, for uniform $2^{\mathrm{D}_i}-\mathrm{QAM}$, $P_{\mathrm{D}_i}(SNR^{(i)})$ is approximated as

$$P_{\mathrm{D}_i}(SNR^{(i)}) \approx 4Q\left(\sqrt{\frac{3(\rho/M)p_i\lambda_i}{2^{\mathrm{D}_i}-1}}\right) \tag{7-7}$$

with $SNR^{(i)} = (\rho/M)p_i\lambda_i$ being the SNR per symbol in the ith eigenchannel. Equation (7-5) assumes perfect error detection with no error correction, wherein blocks are correctly detected if and only if codeword is error free.

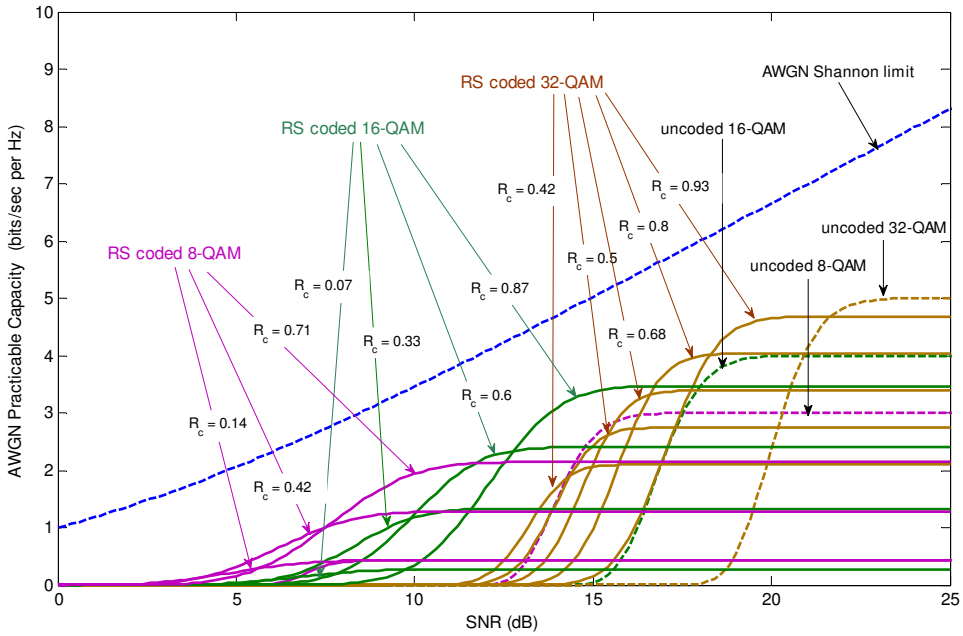

Figure 7-3 Single AWGN channel practicable capacity penalties for uncoded and RS coded data transmission with different code rates for 8-QAM, 16-QAM and 32-QAM

Only the capacity results for some of the code rates are included.

Figure 7-3 illustrates the single AWGN channel practicable capacity (throughput) penalties for uncoded and RS coded data transmission with different code rates for the examples of 8-QAM, 16-QAM and 32-QAM. There is a zero-capacity and a capacity-saturated region for each modulation with a specific coding rate. The impact of increased coding strength (larger t_c) is to decrease the saturated value of the curve and shift its transition to the left. The figure quantifies how stronger codes give better performance at low SNR and poorer performance at high SNR.

With a predefined power allocation scheme such as water-filling or equal power, a look-up-table (LUT) can be used with the adaptive coding. In these cases, for each eigenchannel, selection would be performed among different

candidates of transmission schemes (uncoded and coded data with different code rates associated with different constellation sizes) in a standard optimum way so that the corresponding eigenchannel practicable capacity remains as close as possible to the Shannon limit of the single AWGN channel. The LUT is set up by identifying which transmission scheme achieves the highest practicable capacity for a given SNR and during operation, the LUT would use of the known SNR over each eigenchannel. However, this LUT approach is not applicable for the architecture here because the SNRs on different eigenchannels are unknown until the optimal power allocation is obtained via the optimization problem described below. Maximization of the practicable capacity is from joint optimization of power allocation and code rate over the eigenchannels.

Since the RS (n_c, k_c) code can provide $(n_c - 1)/2 = (2^{\mathrm{D}-1} - 1)$ different code rates, including uncoded data transmission, there are $(n_c - 1)/2 + 1 = (n_c + 1)/2 = 2^{\mathrm{D}-1}$ different candidates for $2^{\mathrm{D}} - \mathrm{QAM}; \mathrm{D} \geq 3$. Confining the modulation set to $\left\{ 2^{\mathrm{D}} - \mathrm{QAM} \middle| 2 \leq \mathrm{D} \leq \mathrm{D}^* = 7 \right\}$ (e.g., the modulation set consists of 4-QAM, 8-QAM, 16-QAM, 32-QAM, 64-QAM and 128-QAM), the total number of transmission candidates over each eigenchannel would be $\sum_{\mathrm{D}=3}^{\mathrm{D}^*} 2^{\mathrm{D}-1}$, plus uncoded 4-QAM. It follows that there are $\xi = \left(1 + \sum_{\mathrm{D}=3}^{\mathrm{D}^*} 2^{\mathrm{D}-1} \right)^{\bar{r}}$ candidates for a MIMO system with $\bar{r}$ deployed eigenchannels when inner coding design is used. For outer coding design, there are only $\xi = 1 + \sum_{\mathrm{D}=3}^{\mathrm{D}^*} 2^{\mathrm{D}-1}$ candidates.

Denoting $R_{\mathbf{H},k}$ and $\mathbf{p}_k^{\text{opt}}$ as the instantaneous MIMO practicable capacity and the optimal power allocation associated with the *k*th candidate of the transmission scheme (note *k* is different to k_c which denotes the number of data symbols being encoded by a RS code), the best transmission scheme index along with the associated optimal power allocation is obtained as

$$(k^{\text{opt}}, \mathbf{p}^{\text{opt}}) = \arg \max_{\substack{\mathbf{p}_k^{\text{opt}} \\ k=1,\ldots,\xi}} R_{\mathbf{H},k}(\mathbf{p}_k^{\text{opt}}) \tag{7-8}$$

where $\mathbf{p}_k^{\text{opt}}$ denotes the optimum power allocation associated with *k*th transmission scheme candidate.

The complexity in the above search can be reduced by omitting those transmission candidates which are always worse than that of others in practicable capacity. For example: the single AWGN channel capacity of uncoded 4-QAM as well as the resultant capacities of RS coded 16-QAM with $k_c = 1,3,5,7$ are fully covered by those of RS coded 8-QAM (seen from Figure 7-3); and the capacity of uncoded 16-QAM transmission is worse than those of RS coded 32-QAM transmission for all values of SNR. As a result, the RS $(15, k_c); k_c = 9,11,13$ codes based on 16-QAM are the only three RS codes that achieve the highest single AWGN channel throughput for a range of SNR.

A similar trend holds for higher constellation sizes at well. The RS $(31, k_c); k_c = 1,3,...,21$ codes based on 32-QAM yield lower single channel practicable capacities than those obtained with RS coded 16-QAM for all values of SNR. Also, the uncoded 32-QAM single channel capacity is less than that of

Table 7-1 RS coded $2^{\mathrm{D}}-\mathrm{QAM}$ candidates for transmission over each eigenchannel in the efficient search procedure

	transmission schemes based on $2^{\mathrm{D}}-\mathrm{QAM}$	# of trans. schemes based on $2^{\mathrm{D}}-\mathrm{QAM}$
8-QAM	RS $(7,k_c)$; $k_c=1,3,5$	$m_8=3$
16-QAM	RS $(15,k_c)$; $k_c=9,11,13$	$m_{16}=3$
32-QAM	RS $(31,k_c)$; $k_c=23,25,27,29$	$m_{32}=4$
64-QAM	RS $(63,k_c)$; $k_c=49,51,...,59,61$	$m_{64}=7$
128-QAM	RS $(127,k_c)$; $k_c=105,107,...,123,125$ **+ uncoded 128-QAM**	$m_{128}=12$

RS coded 64-QAM. It follows that the four RS $(31,k_c)$; $k_c=23,25,27,29$ codes (based on 32-QAM) are the only transmission scheme candidates. Table 7-1 lists the RS codes for the more efficient search for this example, and is based on the assumption that the modulation set is confined to $\left\{2^{\mathrm{D}}-\mathrm{QAM}\,\middle|\,3\le \mathrm{D}\le \mathrm{D}^*=7\right\}$ with 8-QAM and 128-QAM being the lowest and largest constellation sizes, respectively. Denoting m_{D} as the efficient number of transmission schemes candidates based on $2^{\mathrm{D}}-\mathrm{QAM}$, there are $\sum_{\mathrm{D}=3}^{\mathrm{D}^*} m_{\mathrm{D}}$ transmission candidates for each eigenchannel. Consequently, the total number of transmission candidates for a MIMO system with $\bar{r}$ deployed eigenchannels and inner coding, is $\xi_{\text{eff}}=\left(\sum_{\mathrm{D}=3}^{\mathrm{D}^*} m_{\mathrm{D}}\right)^{\bar{r}}$. With outer coding, this number reduces to $\xi_{\text{eff}}=\sum_{\mathrm{D}=3}^{\mathrm{D}^*} m_{\mathrm{D}}$. Relative to the full search with $\mathrm{D}^*=7$, this more efficient search has its complexity reduced by $\frac{\xi_{\text{eff}}}{\xi}\approx\frac{1}{20}$ with inner coding design ($\frac{\xi_{\text{eff}}}{\xi}\approx\frac{3}{13}$ with outer coding).

The last step is to show how the optimal power allocation can be found for each of the transmission scheme candidates. In the following, as an example, the problem is formulated of finding the optimal power allocation when fixed RS coded $2^{\mathrm{D}_i}-\mathrm{QAM}$; $i=1,...,\bar{r}$ with code rates $(R_c^{(1)},R_c^{(2)},...,R_c^{(\bar{r})})$ are assigned to the eigenchannels. This configuration is assumed to be the *k*th power allocation candidate, $\mathbf{p}_k^{\mathrm{opt}}$ (to be used in (7-8)) so the subscript *k* is added in the formulas below. For a chosen number of deployed eigenchannels, $\bar{r}$, the MIMO practicable capacity (with a fixed channel $\mathbf{H}$) is the sum of the throughputs

$$\begin{aligned} R_{\mathbf{H},k} &= \sum_{i=1}^{\bar{r}} R_{\mathbf{H},k}^{(i)} = \max_{\sum p_i=M} \sum_{i=1}^{\bar{r}} R_c^{(i)}\,\mathrm{D}_i \cdot\left[1-BLER_c^{(i)}(P_{\mathrm{D}_i}(SNR^{(i)}))\right] \\ &= \max_{\sum p_i=M} \sum_{i=1}^{\bar{r}} R_c^{(i)}\,\mathrm{D}_i \times\left[1-\sum_{j=t_c^{(i)}+1}^{n_c^{(i)}}\binom{n_c^{(i)}}{j}(P_{\mathrm{D}_i}(SNR^{(i)}))^j(1-P_{\mathrm{D}_i}(SNR^{(i)}))^{n_c^{(i)}-j}\right] \end{aligned} \tag{7-9}$$

with $P_{\mathrm{D}_i}(SNR^{(i)})$ given in (7-7). In (7-9) there is one equality constraint, $\sum_{i=1}^{\bar{r}} p_i = M$. This needs to be eliminated so that resulting unconstrained problem can be solved. We can eliminate p_1 (for example) using the parameterization $p_1 = M - \sum_{i=2}^{\bar{r}} p_i$. The reduced problem is then

$$\begin{aligned} &\underset{p_2,p_3,...,p_{\bar{r}}}{\text{maximize}}\ \psi(p_2,p_3,...,p_{\bar{r}}) = \\ &R_c^{(1)}\,\mathrm{D}_1\left[1-\sum_{j=t_c^{(1)}+1}^{n_c^{(1)}}\binom{n_c^{(1)}}{j}\left(4Q\left(\sqrt{\frac{3(\rho/M)(M-\sum_{i=2}^{\bar{r}}p_i)\lambda_1}{2^{\mathrm{D}_1}-1}}\right)\right)^j\times\left(1-4Q\left(\sqrt{\frac{3(\rho/M)(M-\sum_{i=2}^{\bar{r}}p_i)\lambda_1}{2^{\mathrm{D}_1}-1}}\right)\right)^{n_c^{(1)}-j}\right] \\ &\quad+\sum_{i=2}^{\bar{r}} R_c^{(i)}\,\mathrm{D}_i\left[1-\sum_{j=t_c^{(i)}+1}^{n_c^{(i)}}\binom{n_c^{(i)}}{j}\left(4Q\left(\sqrt{\frac{3(\rho/M)p_i\lambda_i}{2^{\mathrm{D}_i}-1}}\right)\right)^j\left(1-4Q\left(\sqrt{\frac{3(\rho/M)p_i\lambda_i}{2^{\mathrm{D}_i}-1}}\right)\right)^{n_c^{(i)}-j}\right] \end{aligned} \tag{7-10}$$

The elimination of the equality constraints (and reconstructing the solution of the original problem from the solution of the transformed problem) involves standard linear algebraic operations. Thus the problems (7-9) and (7-10) are equivalent. A condition for a point $(p_2^{\text{opt}}, p_3^{\text{opt}}, ..., p_{\bar{r}}^{\text{opt}})$ to be optimal is

$$\nabla\left(\psi(p_2^{\text{opt}}, p_3^{\text{opt}}, ..., p_{\bar{r}}^{\text{opt}})\right) = 0 \tag{7-11}$$

where $\nabla(.)$ denotes the gradient operator. Thus solving the unconstrained maximization problem (7-10) is the same as finding the solution of (7-11), which is a set of $\bar{r}-1$ equations in the $\bar{r}-1$ variables $p_2, p_3, ..., p_{\bar{r}}$. There is no analytical solution to the optimality problem (7-11) and so an iterative technique is used, for example using the gradient descent method

$$\mathbf{p}_k^{(n+1)} = \mathbf{p}_k^{(n)} - \mu^{(n)} \nabla\psi \tag{7-12}$$

is used, where $\mathbf{p}_k^{(n)} = [p_2^{(n)} \quad p_3^{(n)} \quad \cdots \quad p_{\bar{r}}^{(n)}]^T$ is the estimated solution at the *n*th iteration. The required derivative uses an approximation of $Q(x)$ given by [7.27]

$$Q(x) \approx e^{-x^2/2} \sum_{u=1}^{U} a_u x^{u-1} \tag{7-13}$$

where

$$a_u = \frac{(-1)^{u+1} A^u}{B\sqrt{\pi}(\sqrt{2})^{u+1} u!} \tag{7-14}$$

with $A = 1.98$ and $B = 1.135$, and for the simulations $U = 10$ is sufficient. Let

$$P_{\mathrm{D}_i}(x) \approx f(x) = 4e^{-x^2/2} \sum_{u=1}^{U} a_u x^{u-1} \tag{7-15}$$

then from (7-6),

$$BLER_c^{(i)}(P_{\mathrm{D}_i}(x)) \approx g(f(x))$$
$$= \sum_{j=t_c^{(i)}+1}^{n_c^{(i)}} \binom{n_c^{(i)}}{j} (4e^{-x^2/2} \sum_{u=1}^{U} a_u x^{u-1})^j (1-4e^{-x^2/2} \sum_{u=1}^{U} a_u x^{u-1})^{n_c^{(i)}-j} \tag{7-16}$$

So

$$[\nabla \psi]_{\substack{i-1,1 \\ i=2,\ldots,\bar{r}}} = \frac{\partial \psi}{\partial p_i} = R_c^{(1)} \mathrm{D}_1 \left(\frac{3(\rho/M)\lambda_1}{(2^{\mathrm{D}_1}-1)x} \cdot \frac{\partial f(x)}{\partial x} \cdot \frac{\partial g(f(x))}{\partial f} \right) \Bigg|_{x=\sqrt{\frac{3(\rho/M)(M-\sum_{j=2}^{r} p_j)\lambda_1}{2^{\mathrm{D}_1}-1}}}$$
$$- R_c^{(i)} \mathrm{D}_i \left(\frac{3(\rho/M)\lambda_i}{(2^{\mathrm{D}_i}-1)x} \cdot \frac{\partial f(x)}{\partial x} \cdot \frac{\partial g(f(x))}{\partial f} \right) \Bigg|_{x=\sqrt{\frac{3(\rho/M)p_i\lambda_i}{2^{\mathrm{D}_i}-1}}} \tag{7-17}$$

In (7-12), $\mu^{(n)}$ is the step size at iteration *n*, chosen via a backtracking line search. In the terminology of iterative methods, the convergence of the gradient descent algorithm using the backtracking line search is at least linear [7.28]. The stopping convergence criteria of an iterative algorithm such as (7-12) is usually of the form $\|\nabla \psi\| \le \varepsilon$, where ε is small and positive, following the suboptimality condition [7.28]. The method also requires a suitable starting point $\mathbf{p}_k^{(0)}$. Since the objective function $\psi(p_2, p_3, \ldots, p_{\bar{r}})$ is not convex, it may have several maxima in the work space $S = \{(p_2, p_3, \ldots, p_{\bar{r}}) \in \Re^{r-1} \mid 0 \le p_j \le M;\ p_1 - \sum_{j=2}^{\bar{r}} p_j = M\}$ of the iterative gradient descent algorithm (7-12). The work space S is divided into *G* subspaces and the algorithm is run within these. For example, with uniform gridding of the work space and splitting the interval for p_j into *q* equal, distinct segments, there results, at most, $G = q^{\bar{r}-1}$. The subspace solutions are compared to see which one is the global maxima. Denoting $\mathbf{p}_k^{(c)} \mid_{\mathbf{p}_k^{(0)} \in S_g}$ as the convergent point associated with the *g*th segment, the optimum point is

$$\mathbf{p}_k^{\text{opt}} = [\, p_1^{\text{opt}} \;\; p_2^{\text{opt}} \;\; \cdots \;\; p_{\bar{r}}^{\text{opt}} \,]^T = [\, p_1^{\text{opt}} \;\; \mathbf{p}_k^{(c)_{\text{opt}}} \,]^T \tag{7-18}$$

where

$$\mathbf{p}_k^{(c)_{\text{opt}}} = \arg \max_{\substack{\mathbf{p}_k^{(c)} | \mathbf{p}_k^{(0)} \in S_g \\ g=1,\ldots,G}} \psi(\mathbf{p}_k^{(c)} | \mathbf{p}_k^{(0)} \in S_g) \tag{7-19}$$

In practice, MIMO systems with $r = \min(M, N) \le 4$ are of immediate interest, and simulations suggest that for any $\bar{r} \le 4$, setting $q = \bar{r} - 1$ is suitable for the above procedure. Finally, for the *k*th transmission scheme candidate,

$$R_{\mathbf{H},k} = \sum_{i=1}^{\bar{r}} R_c^{(i)} \, \mathrm{D}_i . \left[1 - \sum_{j=t_c^{(i)}+1}^{n_c^{(i)}} \binom{n_c^{(i)}}{j} \left(Q\left(\sqrt{\frac{3(\rho/M) p_i^{\text{opt}} \lambda_i}{2^{\mathrm{D}_i} - 1}} \right) \right)^j \left(1 - Q\left(\sqrt{\frac{3(\rho/M) p_i^{\text{opt}} \lambda_i}{2^{\mathrm{D}_i} - 1}} \right) \right)^{n_c^{(i)} - j} \right] \tag{7-20}$$

We note that for those candidates of transmission schemes which include uncoded data transmission over at least one of the eigenchannels, the formulation in (7-9)-(7-17) has to be modified (R_c cannot simply be set to 1) according to $R_{\mathbf{H}}^{(i)} = R_c^{(i)} \, \mathrm{D}_i . \left[1 - P_{\mathrm{D_i}}(SNR^{(i)}) \right]^{BL}$ where BL is the block length in uncoded symbols (this capacity formulation is presented in chapters 5 and 6 in detail) in order to account for the uncoded throughput where it is appropriate.

7.3 Simulation Results

For simulations, the modulation set is confined to $\left\{ 2^{\mathrm{D}} - \mathrm{QAM} \middle| \, 2 \le \mathrm{D} \le \mathrm{D}^* = 7 \right\}$ (the set consists of 4-QAM, 8-QAM, 16-QAM, 32-QAM, 64-QAM and 128-QAM) and the presented adaptive RS-coded modulation is applied to the examples of a single Rayleigh fading channel, and a 2×2 MIMO system with different system configurations (different arrangements of CODECs

with different number of deployed eigenchannels $\bar{r} \leq 2$ and power allocation scheme); although the algorithm is applicable to any $N \times M$ MIMO system. The deployment of different numbers of eigenchannels in MIMO includes the case of ACM applied to only the eigenchannel with the maximum gain (so $\bar{r} = 1$), referred to as dominant eigenmode transmission. In this case, as well as transmission over a single Rayleigh fading channel, the LUT-based approach can be used as an alternative to the presented ACM algorithm. This is because at each channel realization, all the fixed available power at the transmitter is dedicated to a single channel and the SNR on the channel is known before searching for the best transmission candidate with the highest capacity. Therefore, the power allocation problem is discarded, and the presented algorithm and the LUT-based approach yield the same result. However, when the number of deployed eigenchannels in MIMO is greater than one, the use of an LUT-based approach for the cases with a predefined power allocation scheme such as water-filling, would not yield the same result as with the presented algorithm (shown below).

Figure 7-4 illustrates the resulting ergodic practicable capacities for a single Rayleigh fading channel. The presented ACM scheme shows an effective coding gain of 2 dB relative to adaptive uncoded modulation in a SISO system for moderate values of SNR. It is emphasized that the results presented here are for the case where the total transmit power at the transmitter is fixed at each channel realization. With variable power transmission, higher coding gain is expected since the power at the transmitter changes according to channel

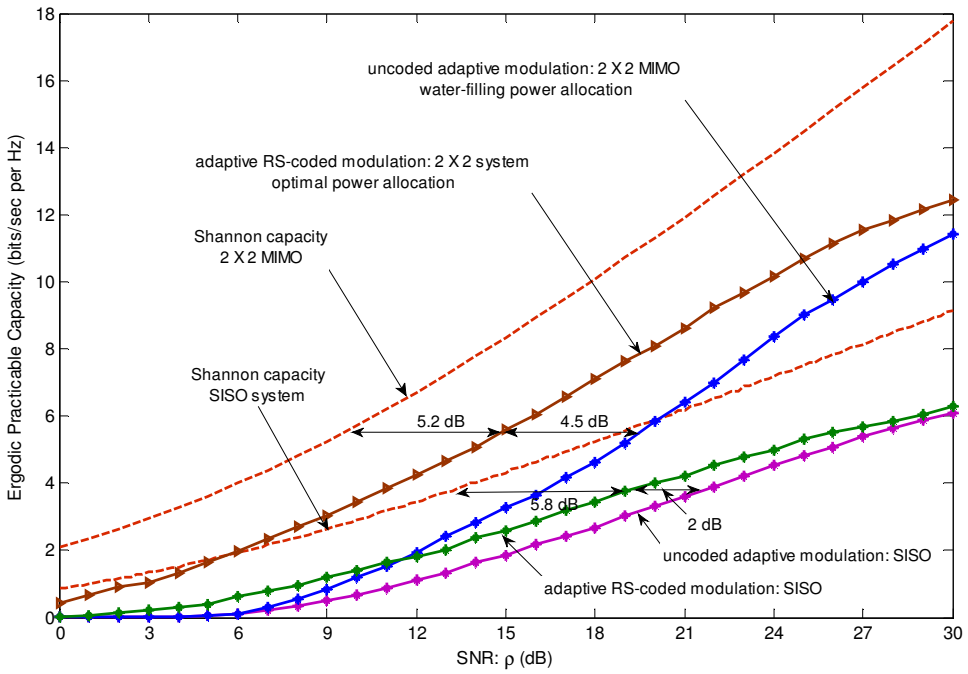

Figure 7-4 The ergodic practicable capacities resulting from the presented adaptive RS coded modulation applied to a single Rayleigh fading channel and a 2×2 MIMO system with inner coding design.

The practicable capacities of the uncoded systems are also included. The curves use 5000 realizations to approximate the ergodic capacity

conditions. The resulting enhancement in the capacity is from the variable power transmission and associated ACM. As noted in the chapter introduction, variable power transmission is not considered here since the focus is on power allocation with fixed total power at the transmitter.

In the fixed total transmit power (for each channel realization) case treated here, the power allocation over eigenchannels contributes to capacity enhancement. The resulting practicable capacities for a 2×2 MIMO system with outer/inner coding design are presented in Figure 7-4 and Figure 7-5, and the associated practicable capacities with water-filling power allocation are also

included. With either of outer and inner coding designs, the maximized capacity (with the optimal power allocation) is higher than that from using water-filling power allocation. As an example, the difference in the respective capacities reaches 0.7 bits/sec/Hz at moderate values of SNR. Alternatively stated, ACM with the optimal power allocation increases the capacity by 30% over the case where water-filling is used. The results in Figure 7-5 quantify the non-optimality of the practicable capacity from using water-filling for the power allocation, for this system configuration.

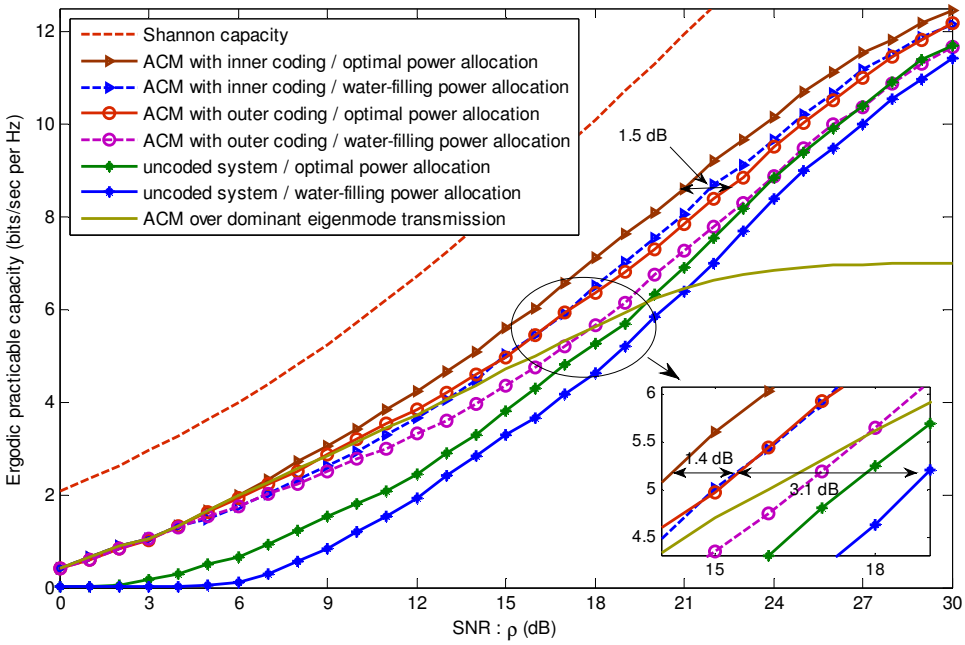

Figure 7-5 Comparison of the resulting ergodic practicable capacities with optimal power allocation for outer/inner adaptive RS coded modulation with the corresponding capacities using water-filling in a 2×2 MIMO system

The inner coding design outperforms the outer coding design irrespective of the power allocation used. For example, for a rate example of 8 bits/sec/Hz, the difference in the average SNR requirement is about 1.5 dB. In general, it

turns out that the difference between the capacity curves depends on many factors such as average SNR at the receiver ρ, power allocation scheme, arrangement of CODECs, minimum number of transmit and receive antennas r, and number of deployed eigenchannels $\bar{r}$. In a 2×2 MIMO system, ACM with $\bar{r}=2$ and inner coding achieves the highest practicable capacity among other configurations. It shows an effective coding gain of about 4.5 dB relative to uncoded adaptive modulation with water-filling, for moderate values of SNR (above 6 dB) typical of wireless systems. Although in the presented algorithm the power allocation part cannot be separated from the coding and modulation assignment over eigenchannels, the coding gain of 4.5 dB may be intuitively viewed as the addition of two gains resulting from: the effect of the ACM scheme with non-optimal power allocation applied to uncoded system (3.1 dB gain - the difference between the inner coded ACM with water-filling and the uncoded system with water-filling); and the correction of the optimal power allocation (1.4 dB gain - the difference between the inner coded ACM with optimal power allocation and the inner coded ACM with water-filling power allocation).

Although the presented ACM with $\bar{r}=2$ and inner coding offers the highest practicable capacity in a 2×2 system, the error performance of the overall system deteriorates with the weakest eigenchannel. Because of the sharply decaying distribution of eigenvalues over eigenchannels [7.23], [7.29], there is a large difference between the average SNR of the strongest and the lowest eigenchannel. For example, in a 2×2 system, the difference between the

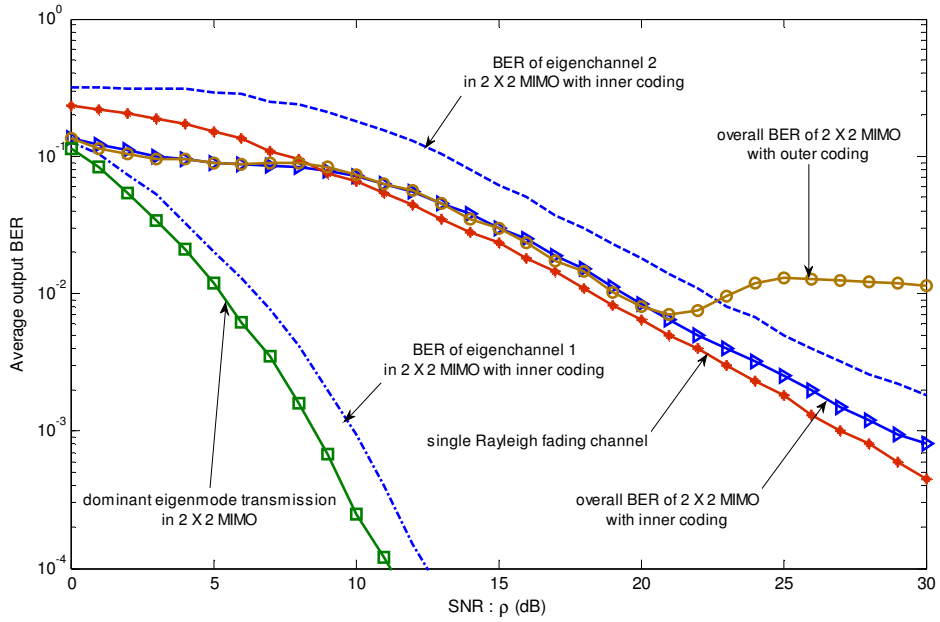

Figure 7-6 The overall BER performance of the presented adaptive RS coded modulation applied to a single Rayleigh fading channel and a 2×2 MIMO with different system configurations.

The individual BER curves associated with each of the eigenchannels are also included.

average SNRs on eigenchannels is approximately 8.5 dB when the average receive SNR is 20 dB, and this difference increases to 13.5 dB in a 3×3 system [7.30]. As a result, a large degradation occurs in the overall system BER performance. The overall BER performance of the system (referring to the BER of the overall serial data in MIMO) is a measure for comparison among different system configurations. The overall BER for a single Rayleigh fading channel and a 2×2 MIMO using the presented ACM are illustrated in Figure 7-6. The BER curves associated with each of the eigenchannels are included since they help to understand how the overall BER performance of the system deteriorates. Based on the overall BER performance, the use of ACM with dominant eigenmode

transmission is the best. This was expected since it is the optimum design that maximizes the output SNR.

The presented ACM yields nearly the same performance with either of inner or outer coding design for small to mid values of SNR. However, for large values of SNR, the system with outer coding experiences a large degradation in overall BER performance. This is because of the assignment of only one transmission scheme (constellation size and code rate) to all of the eigenchannels with the metric of highest MIMO practicable capacity. Since the uncoded 128-QAM transmission yields the highest practicable capacity (of the schemes considered) for large values of SNR, it is selected as the transmission scheme for all eigenchannels in the outer coding design. As a result, the overall BER performance of the system severely deteriorates with the weakest eigenchannel transmitting uncoded 128-QAM.

Accurate polynomial approximations for the overall BER of system using the proposed ACM scheme can be found from a nonlinear least-squares curve-fitting in the form

$$BER(\rho) = \sum_{i=1}^{\Gamma} a(i) e^{b(i)\rho} \tag{7-21}$$

where $a(i)$ and $b(i)$; $i = 1,...,\Gamma$ are the (complex-valued) coefficients that can be found curve fitting to the simulated overall BER performance of the ACM. The approximations are illustrated in Figure 7-7 for $\Gamma = 4$, showing the good accuracy. The coefficients are given in Table 7-2. These BER expressions are

useful for quick assessment of the systems' error performance. In particular, for each system configuration, the minimum SNR can be found at which the average

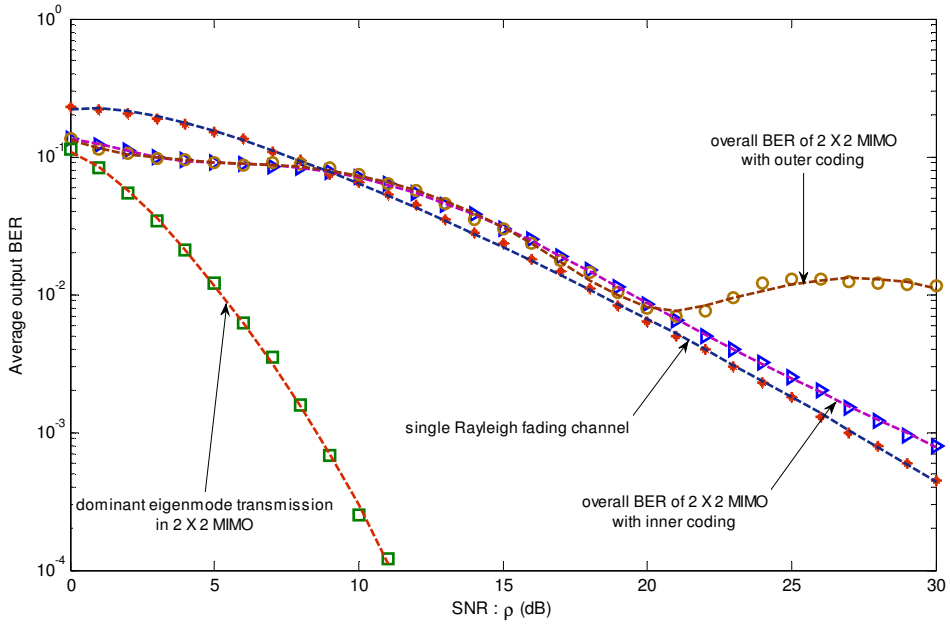

Figure 7-7 The BER approximations for the presented ACM applied to a single Rayleigh fading channel and a 2×2 MIMO with different system configurations.

The dotted lines denote the approximations obtained via nonlinear LS curve fitting, whereas the markers represents the results from simulations

Table 7-2 The coefficients of the BER approximation expression (7-21) associated with the presented adaptive RS-coded modulation applied to a single Rayleigh fading channel and a 2×2 MIMO with different system configurations

	single Rayleigh fading channel	dominant eigenmode transmission in a 2×2 MIMO	overall BER of a 2×2 MIMO with inner coding	overall BER of a 2×2 MIMO with outer coding
$a(1)$	0.1095 - 0.5154i	0.0545 + 0.1258i	0.0261 + 0.8417i	0.0015 + 0.0002i
$b(1)$	-0.2531 + 0.0660i	-0.6317 - 0.2091i	-0.2577 - 0.0560i	-0.1255 + 1.6298i
$a(2)$	0.1095 + 0.5154i	0.0545 - 0.1258i	0.0404 - 0.0946i	0.1647 - 0.0002i
$b(2)$	-0.2531 - 0.0660i	-0.6317 + 0.2091i	-0.2701 - 0.3495i	-0.1108 + 0.0001i
$a(3)$	0	0	0.0276 - 0.8417i	-0.0166 + 0.0034i
$b(3)$	0	0	-0.2576 + 0.0559i	-0.0617 + 0.3306i
$a(4)$	0	0	0.0403 + 0.0946i	-0.0168 - 0.0035i
$b(4)$	0	0	-0.2701 + 0.3494i	-0.0623 - 0.3308i

BER falls below a targeted value. For example, for a target average BER of 0.01, denoted as $\mathrm{BER}_0 = 0.01$, the minimum SNR for the presented ACM over dominant eigenmode transmission is $\mathrm{SNR}_{\min}^{\mathrm{dom}}\big|_{\mathrm{BER}_0} = 5\,\mathrm{dB}$. This value increases to 19 dB for ACM with inner coding. In general, none of the system configurations simultaneously maximizes the practicable capacity and satisfies a targeted BER constrained, for all values of SNR. For a given SNR, the choice is from different system configurations with the metric of highest practicable capacity subject to an average BER value. Different system configurations for selection are different combinations of CODECs (inner and outer coding) with different number of deployed eigenchannels $\bar{r} \le 2$, and power allocation scheme (optimal power allocation and water-filling). The selection is from those system configurations that have the required minimum SNR for the BER constraint satisfaction. For example, for $\mathrm{BER}_0 = 0.01$ and $5\,\mathrm{dB} \le \rho \le 17\,\mathrm{dB}$, the dominant eigenmode transmission is the only available scheme, whereas for $\rho = 20\,\mathrm{dB}$, the selection is from multiple system configurations. This is depicted in Figure 8 which illustrates the capacity results with selection for different values of target average BER.

With the presented selection, the capacity curves are piece-wise discontinuous. The discontinuities are the result of selection part of the algorithm. This is not restrictive since at each SNR value the best system configuration is selected irrespective to other SNR values. We note that, as the result of selection, several capacity curves for different values of BER_0 may lie together

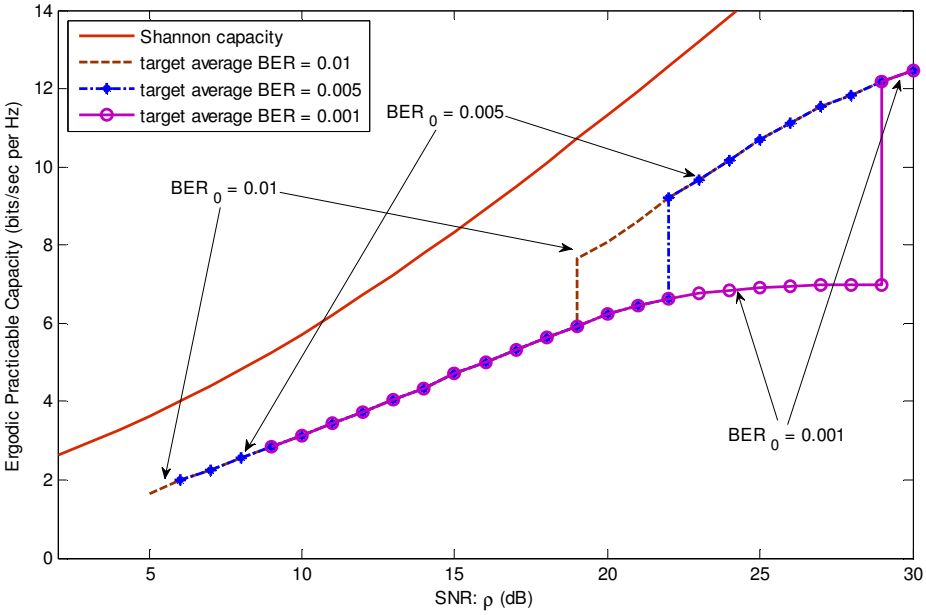

Figure 7-8 The practicable capacities resulted from selection between different system configurations subject to different values of average output BER for the 2×2 MIMO example.

Dominant eigenmode transmission (with the presented ACM) gives the best capacity (constrained by BER_o=0.01) for an SNR of 5 to 19dB, and for larger SNRs, the inner coded system has the highest capacity.

for some SNR regions. For example, in Figure 7-8, the curves tagged with $\mathrm{BER}_0 = 0.01$, $\mathrm{BER}_0 = 0.005$, and $\mathrm{BER}_0 = 0.001$ partially lie on the capacity result of dominant eigenmode transmission and ACM with inner coding for moderate and high values of SNR, respectively. In particular, with $\mathrm{BER}_0 = 0.005$, for $\sim 6\,\mathrm{dB} < \rho < \sim 22\,\mathrm{dB}$ the curve lies on the result of dominant eigenmode transmission and for $\sim 22\,\mathrm{dB} < \rho$ the capacity curve jumps to the capacity of the ACM with inner coding. This shows how a target average output BER, say $\mathrm{BER}_0 = 10^{-3}$, in a MIMO system constrained with a fixed total input transmit

power, forces the data to be transmitted over just one or a few eigenchannels having the strongest eigenvalues.

Targeted with an average BER, the selection procedure is undertaken in an off-line mode. Once the best system configuration is identified for a given SNR, the designer deploys the selected system configuration for on-line operation.

The selection procedure for the highest practicable capacity can also be subject to a target instantaneous BER. In this case, at each channel realization, the instantaneous error probabilities are computed for different system configurations and the one with the highest instantaneous practicable capacity is selected. This selection has to be carried out online. The associated ergodic

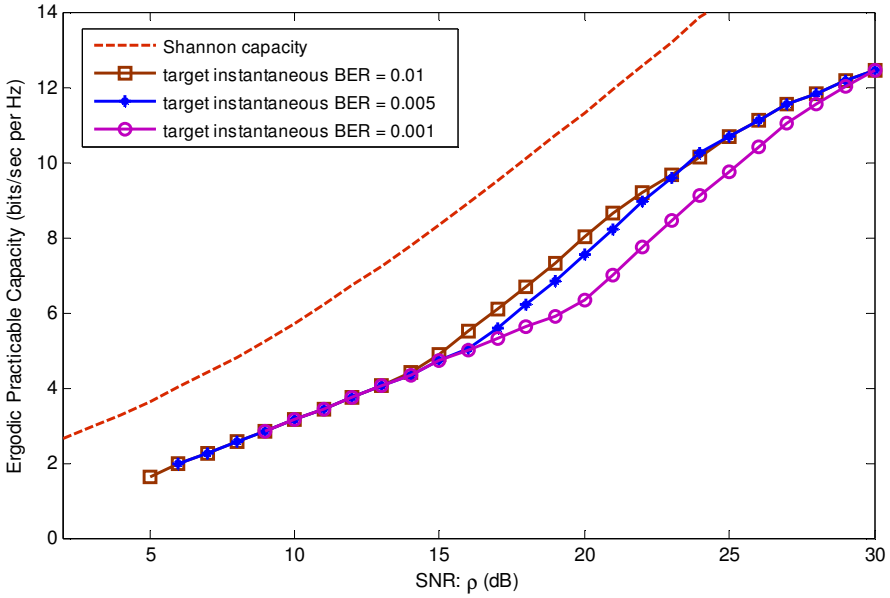

Figure 7-9 The practicable capacities resulted from selection between different system configurations subject to different values of instantaneous output BER for the 2×2 MIMO example.

practicable capacity is the ensemble average of the instantaneous capacities achieved via selection at different channel realizations. The results are depicted in Figure 7-9 for different values of target instantaneous BER. As the result of ensemble averaging, continuous curves are obtained here.

Regardless of the type of output BER constraint (average or instantaneous), the resultant capacity curves via selection are upper and lower bounded by the capacities obtained from the presented ACM applied to inner-coded full-eigen MIMO, and dominant eigenmode transmission, respectively. In general, for smaller values of BER_0, the system with lower capacity is expected. Nevertheless, the described selection procedure makes it possible to design a system that achieves the highest practicable capacity with a desired output BER_0. In other words, the position of the capacity between the upper and lower capacity bounds can be controlled by appropriately setting the BER_0 and the output BER constraint type.

7.4 Summary and Conclusions

The problem is addressed for joint power allocation and code rate adaptation maximizing the practicable capacity in an eigen-MIMO digital link with Reed-Solomon coded adaptive modulation. As can be intuitively expected, inner coding performs better than outer coding, but at the expense of higher complexity, and here the performances are quantified. Also, it turns out that water-filling is not the optimum power allocation for the practicable capacity maximization for given system configuration. The approach allows the capacity to

be compared between uncoded and adaptively RS-coded modulation schemes with respect to: the CODEC architecture (inner or outer coding); average SNR; the power allocation scheme; the number of antennas; and the number of deployed eigenchannels. For example, with inner coding, the presented ACM shows an effective coding gain of 4.5 dB for moderate to high SNRs compared to that of uncoded adaptive modulation with water-filling in a 2×2 system.

Comparison of the error curves of the ACM algorithm with different system configurations including ACM applied to dominant eigenmode transmission, allow an optimal capacity selection for a target average or instantaneous output BER, BER_0. None of the system configurations maximizes the capacity for all values of SNR and BER_0. Nevertheless, with a given choice of BER_0, the maximum capacity is achieved by selection from these different system configurations. The capacity curves via selection are upper bounded by the inner coded full eigen-MIMO ACM capacity, and lower bounded by the capacity obtained via the presented ACM applied to the dominant eigenmode transmission. Furthermore, the position of the capacity between the upper and lower capacity bounds can be controlled by appropriately setting the BER_0 and the type of the output BER constraint.

REFERENCE LIST

[7.1] G. Ungerboeck, “Channel coding with multilevel/phase signals,” *IEEE Trans. Inf. Theory,* vol. IT-28, pp. 55–67, Jan. 1982.

[7.2] J. Goldsmith and S. Chua, “Adaptive coded modulation for fading channels," *IEEE Trans. Commun.,* vol. 46, no. 5, pp. 595-602, May 1998.

[7.3] S. M. Alamouti and S. Kallel, “Adaptive trellis-coded multiple-phasedshift keying for Rayleigh fading channels,” *IEEE Trans. Commun.,* vol. 42, no. 6, pp. 2305–2314, Jun 1994.

[7.4] D. V. Duong, G. E. Øien, and K. J. Hole, “Adaptive coded modulation with receive antenna diversity and imperfect channel knowledge at receiver and transmitter,” *IEEE Trans. Veh. Technol.,* vol. 55, no. 2, Mar. 2006.

[7.5] X. Cai and G. B. Giannakis, “Adaptive PSAM accounting for channel estimation and prediction errors,” *IEEE Trans. Wireless Commun.,* vol. 4, no. 1, pp. 246–256, Jan. 2005.

[7.6] S. Zhou and G. B. Giannakis, “Adaptive modulation for multi-antenna transmissions with channel mean feedback,” *IEEE Trans. Wireless Commun.,* vol. 3, no. 5, pp. 1626–1636, Sep. 2004.

[7.7] H. Sankar, N. Sindhushayana, and K. R. Narayanan, “Design of low density parity-check (LDPC) codes for high order constellations," in *Proc. IEEE Globecom*, pp. 3113-3117. 2004.

[7.8] G. Masera, F. Quaglio, and F. Vacca, “Implementation of a flexible LDPC decoder," *IEEE Trans. Circuits Syst. II*, vol. 54, no. 6, pp. 542-546, Jun. 2007.

[7.9] J. Hou, P. H. Siegel, L. B. Milstein, D. Pfister, “Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2141-2155, Sep. 2003.

[7.10] G. Caire, and K. R. Kumar, “Information theoretic foundations of adaptive coded modulation,” *IEEE Proceedings,* vol. 95, no. 12, pp. 2274-2298, Dec. 2007.

[7.11] Seok-Ki Ahn and K. Yang, “Adaptive modulation and coding schemes based on LDPC codes with irregular modulation,” *IEEE Trans. Commun.,* vol. 58, no. 9, pp. 2465-2470, Sep. 2010.

[7.12] J. Ha, J. Kim, and S. W. McLaughlin, “Rate-compatible puncturing of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2824-2836, Nov. 2004.

[7.13] S. Vishwanath, A. Goldsmith, “Adaptive turbo-coded modulation for flat-fading channels,” *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 964-972, Jun. 2003.

[7.14] S. Falahati, A. Svensson, M. Sternad, and H. Mei, “Adaptive trellis-coded modulation over predicted flat fading channels,” in *Proc. IEEE Veh. Technol. Conf.*, pp. 1532–1536, Oct. 2003.

[7.15] A. I. Phelps, and R. M. Buehrer, “A simplified technique for adaptive MIMO, modulation, and coding,” in *Proc. Radio and Wireless Symp.,* pp. 566-569, 2009.

[7.16] S. A. Banani, and R. G. Vaughan, “Adaptive Reed-Solomon coding in eigen-MIMO with non-adaptive modulation,” Accepted for presentation in *IEEE Veh. Technol. Conf. (VTC2011-Fall)*, San Francisco, USA, Sep. 2011.

[7.17] L. Chaari, M. Fourati, and L. Kamoun, “Image transmission quality analysis over adaptive Reed-Solomon coding,” in *Proc. IEEE MELECON Conf.,* pp. 409-414, 2010.

[7.18] T. A. Tran, N. Chehrazi, Tho Le-Ngoc, “Adaptive Reed-Solomon coding scheme for OFDM systems over frequency-selective fading channels,” in *Proc. IEEE Veh. Technol. Conf.,* pp. 1-5, 2006.

[7.19] Y. L. Grushevsky, and G. F. Elmasry, “Adaptive RS codes for message delivery over an encrypted mobile network,” *IET Commun.,* vol. 3, no. 6, pp. 1041-1049, 2009.

[7.20] J. G. Proakis, and M. Salehi, *Digital communication*, 5th ed. New York: McGraw-Hill, 2008.

[7.21] S. A. Banani, and R. G. Vaughan, “Capacity Maximization in Eigen-MIMO from Adaptive Modulation and Reed-Solomon Coding,” *IET Commun.,* Jul. 2011, Manuscript ID: COM-2011-0730.

[7.22] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” Wireless Personal Commun., Vol. 6, No. 3, pp. 311-335, March 1998

[7.23] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” European Trans. Telecommun., Vol. 10, No. 6, pp. 585-595, Nov.-Dec. 1999

[7.24] S. Catreux, P. F. Driessen and L. J. Greenstein, “Data throughputs using multiple-input multiple-output (MIMO) techniques in a noise-limited cellular environment,” *IEEE Trans. Wireless Commun.,* vol. 1, no. 2, pp. 226 – 235, Apr 2002.

[7.25] M. V. Eyuboglu, “Detection of coded modulation signals on linear, severely distorted channels using decision feedback noise prediction with interleaving,” *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 401–409, Apr. 1988

[7.26] S. A. Banani, amd R. Vaughan, “Power allocation for practicable capacity maximization in eigen-MIMO,” in *Proc. IEEE Veh. Technol. Conf.,* Budapest, Hungary, May, 2011

[7.27] Y. Isukapalli, and B. Rao, “An Analytically tractable approximation for the Gaussian Q-function,” *IEEE Commun. Letters,* vol. 12, no. 9, pp. 669-671, Sep. 2008.

[7.28] S. Boyd, and L. Vandenberghe, *Convex optimization*, Cambridge , UK: Cambridge University Press, 2004.

[7.29] R. Vaughan, and J. Bach Anderson, *Channels, propagation and antennas for mobile communications*, London, UK: IEE, 2003.

[7.30] S. A. Banani, and R. G. Vaughan, “Information-rate maximization in eigen-MIMO with output SNR constraint,” submitted to the *IEEE Trans. Wireless Commun,* Sep. 2010.

8: APPENDIX: LIST OF IDEAS FOR FUTURE WORKS

Some future research projects, that would follow from this thesis, is as follows:

- Consideration of decision-directed channel estimation for systems with cooperative communications and relaying.
- Decision-directed channel estimation for closed-loop MIMO system. In this regard, one have to tackle the challenges regarding the practical communications techniques such as the choice of modulation, coding, the allowed transmit power, and especially the required feedback which consumes bandwidth. The required channel resource for the feedback link can be relatively small in terms of bandwidth consumption, but its inclusion requires considerable processing and protocol resource.
- Adaptive Reed-Solomon coded modulation in eigen-MIMO with channel estimation and imperfect feedback.
- Adaptive trellis-coded modulation in eigen-MIMO with variable power transmission (instead of fixed total transmit power) at the transmitter. This includes the challenging problem of optimal power allocation over eigenchannels.