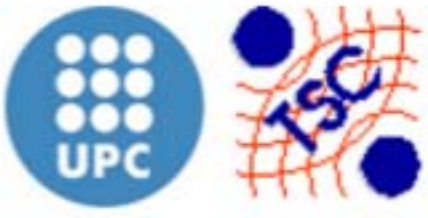


Antenna Selection Techniques in Single- and Multi-user Systems: A Cross-layer Approach

TESI

per a l'obtenció del títol de

Doctor per la Universitat Politècnica de Catalunya

per

José López Vicario
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
E-mail: vicario@ieee.org

Dirigida per

Dr. Carles Antón Haro
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
E-mail: carles.anton@cttc.es

i tutoritzada per

Prof. Ana I. Pérez Neira
Universitat Politècnica de Catalunya (UPC)
E-mail: anuska@gps.tsc.upc.edu

20 de Juny de 2006

Universitat Politècnica de Catalunya (UPC)

Departament de Teoria del Senyal i Comunicacions (TSC)

A Mayte,

"Such a call would understandably upset most people"

Charles Bronson

Abstract

Over the last decade, the massive demand for high data-rate wireless applications has motivated the study and design of new communication technologies. Among all of them, multi-antenna schemes have been shown to provide remarkable benefits in terms of spectral efficiency. In order to achieve channel capacity bounds, some sort of pre-processing on the transmit side must be encompassed. Unless reciprocity between the forward and reverse links can be assumed, a feedback channel is required to convey channel state information. In such a context, transmit antenna selection emerges as an effective alternative requiring a low amount of information in the feedback channel.

Antenna selection algorithms are usually designed in order to maximize physical layer performance but, recently, optimization criteria based on cross-layer designs are gaining popularity. This is mainly motivated by the inefficiency observed in the direct application of the Open Systems Interconnection (OSI) protocol stack in wireless networks. Basically, these new cross-layer criteria are based on the combination of functionalities and information between different layers of the protocol stack in order to attain the highest end-to-end performance.

In this PhD dissertation, we address the study of antenna selection algorithms in single- and multi-users systems from a cross-layer perspective. In particular, we focus on the lowest two layers of the OSI protocol stack (physical and data link layer) and our designs are devoted to maximize data link layer performance.

First, antenna selection is encompassed in a single user scenario, where a Vertical Bell-Labs LAyered Space-Time (V-BLAST) architecture and a Hybrid-Automatic Repeat-reQuest (H-ARQ) strategy are adopted in the physical and data link layers, respectively. We address a cross-layer methodology in the sense that the *criterion* for the selection of antenna subsets is the maximization of *data link layer throughput*. This cross-layer approach is shown to exhibit superior performance in comparison with conventional physical layer criteria since all the system characteristics in both the physical and data link layers are taken into consideration. In order to enhance system performance, adaptive modulation is included to jointly perform antenna selection and rate adaptation. In addition, reduced-complexity versions are derived along with a detailed complexity analysis, showing that the proposed cross-layer approach attains significant gains with an affordable computational complexity.

Next, the study is extended to a multi-user scenario, where only one user is scheduled for transmission in each time-slot by means of an opportunistic scheduling strategy. In this case, antenna selection mechanisms are considered in order to efficiently exploit both spatial and multi-user diversity gains. The analysis is conducted in scenarios where different impairments are introduced in the channel state information available at the scheduler. Robustness against channel state information degradation is then improved by combining Orthogonal Space-Time

Block-Coding (OSTBC) with antenna selection algorithms.

Finally, Orthogonal Random Beamforming (ORB) is considered in order to serve several users in the same time-slot. This strategy is adopted due to its applicability in wireless networks as only partial channel state information is required at the transmitter. However, significant gains are only obtained in scenarios with a high number of users. In order to improve performance in low populated environments, we adopt the idea of antenna selection to develop a beam selection approach. Besides, sub-optimum beam selection procedures are also derived and trade-offs in terms of system performance vs. amount of information in the feedback channel are addressed. In particular, we derive a beam selection approach capable of extracting most of the gains with a few bits in the feedback channel.

Resumen

En la última década, la demanda masiva de aplicaciones inalámbricas con altas velocidades de transmisión ha motivado el estudio y diseño de nuevas tecnologías de comunicación. Entre todas ellas, los esquemas con múltiples antenas han mostrado obtener considerables ventajas en términos de eficiencia espectral. Para explotar la capacidad de estos sistemas, deben llevarse a cabo técnicas de procesado de señal las cuales requieren el conocimiento del estado canal. No obstante, al menos que los enlaces ascendentes y descendentes presenten algún tipo de reciprocidad, un canal de *feedback* es necesario para informar del estado del canal al transmisor. En ese contexto, los sistemas basados en la selección de antenas en transmisión han emergido como una alternativa efectiva dado que requieren una baja cantidad de información en el canal de *feedback*.

Los algoritmos de selección de antenas se suelen diseñar con el propósito de maximizar parámetros en la capa física pero, recientemente, criterios de optimización basados en diseños *cross-layer* han estado ganando popularidad. Esto es principalmente debido a la ineficiencia observada cuando una pila de protocolos *Open Systems Interconnection* (OSI) se aplica directamente a una red inalámbrica. Básicamente, estos nuevos criterios *cross-layer* se basan en combinar funcionalidades e informaciones entre las diferentes capas de la pila de protocolos para conseguir las máximas prestaciones *end-to-end*.

En esta tesis, nos centramos en el estudio, desde un punto de vista *cross-layer*, de algoritmos de selección de antenas tanto en escenarios con un único usuario como en entornos multi-usuario. En concreto, consideramos las dos capas más bajas de la pila de protocolos OSI (capas física y de enlace de datos) y nuestros diseños se basan en maximizar las prestaciones en la capa de enlace.

En primer lugar, la selección de antenas se lleva a cabo en un entorno con un único usuario, en el cual se adoptan una arquitectura *Vertical Bell-Labs LAyered Space-Time* (V-BLAST) y una estrategia *Hybrid-Automatic Repeat-reQuest* (H-ARQ) en las capas física y de enlace, respectivamente. Llevamos a cabo una metodología *cross-layer* en el sentido que el criterio usado para seleccionar antenas se basa en la maximización del *throughput* de la capa de enlace. Este método *cross-layer* presenta mejores prestaciones respecto a criterios convencionales basados en la capa física debido a que todas las características del sistema, tanto en la capa física como en la de enlace, son tomadas en consideración. Con el propósito de mejorar las prestaciones del sistema, se incluyen esquemas de modulación adaptativa con la finalidad de seleccionar de forma conjunta las antenas y la velocidad de transmisión. Además, se derivan versiones más eficientes del algoritmo de selección en términos de coste computacional, con lo cual se muestra que el método propuesto es capaz de conseguir una mejora considerable con unos niveles tolerables de coste computacional.

A continuación, el estudio se extiende al caso multi-usuario, donde en cada *slot* temporal sólo se permite la transmisión a un usuario por medio de una estrategia de *scheduling* oportunista. En este caso, los mecanismos de selección de antenas son considerados para explotar de forma eficiente las ganancias de diversidad tanto espacial como multi-usuario. El análisis se lleva a cabo en escenarios con diferentes anomalías en la información del estado del canal disponible en el *scheduler*, siendo la robustez ante tales anomalías mejorada mediante la combinación de los algoritmos de selección de antenas con *Orthogonal Space-Time Block-Coding* (OSTBC).

Finalmente, para servir varios usuarios en el mismo *slot* temporal, se considera la estrategia *Orthogonal Random Beamforming* (ORB) debido a su aplicabilidad en redes inalámbricas ya que únicamente se necesita en el transmisor el conocimiento parcial del estado del canal. Además, se pueden obtener considerables ganancias en escenarios con un número alto de usuarios. Con el propósito de mejorar las prestaciones del sistema en entornos con pocos usuarios, recurrimos a la idea de selección de antenas y la adaptamos al nuevo contexto proponiendo un esquema de selección de haces. Aparte de eso, se proponen esquemas de selección sub-óptimos y se llevan a cabo estudios de compromiso en términos de prestaciones del sistema vs. cantidad de información en el canal de *feedback*. En concreto, derivamos un esquema de selección capaz de extraer gran parte de la ganancia del sistema con unos pocos *bits* en el canal de *feedback*.

Acknowledgements

Bueno, aquí estamos, todo se acaba y esto por fin ha llegado a su fin. Hacer la tesis es una cosa que está muy bien pero siempre diré que debería durar solo tres años. Ahora, es momento de agradecer el apoyo de las personas recibido en este tiempo.

En primer lugar, debo agradecer la ayuda de mi director de tesis Carles Antón. Siempre ha estado disponible para discutir los trabajos realizados y para solucionar cualquier problema. A mi tutora de tesis Ana I. Pérez también debo darle gracias por haberme guiado durante mi proyecto final de carrera y haberme dado la oportunidad de iniciar el doctorado. No obstante, gran parte de mi motivación para iniciar la tesis es debida a mi asistencia a una clase de *Arrays* impartida por Miguel Ángel Lagunas. En aquella época no tenía muy claro que hacer, y dicha clase hizo que naciera el investigador que llevo dentro.

En el CTTC he conocido muy buena gente y es por ello que voy a darle las gracias por aguantarme durante todo este tiempo. Empezaré con los chicos del *Ghetto*, con los cuales he vivido muchas situaciones divertidas, entrañables y raras, raras, raras (como diría Papuchi, que en paz descanse). Así que, gracias a Miquel Payaró (el Cañita Brava), Ricardo Martínez (el Richal), Toni Morell (el Antonio), Diego Bartolomé (el Picasol), Toni Pascual (el Pascualín) y Francisco Rubio (el Johny). Aparte de esta gente, he tenido muy buena relación con otras personas fuera del *Ghetto* pero dentro del bar, como Pavel Miskovsky, Jesús Gómez y Javi Matamoros, a ver si te apuntas Aitor del Coso. También ha habido colegas de *running* como Christian Ibars, al cual le debo agradecer su ayuda en la revisión de mi tesis y su gran forma de ser, y a Jordi Cebriá, quien me ha ayudado con el tema informático y televisivo. Finalmente, debo agradecer a Josep Mangues, Jesús Alonso, Marc Realp, Raul Muñoz, Christos Verikoukis, Stephan Pfletschinger, Mónica Navarro, Aurora Anguita, Xavi Nieto, Iván Martínez, Lluis Berenguer, Lluis Ventura, Joan Anton Leyva, y a todo el resto del CTTC por esas conversaciones que van saliendo y reconfortan día a día.

Also, I would like to thank my international colleagues and collaborators: Roberto Bosisio, Umberto Spagnolini, Christoph Mecklenbräuker, Stavros Toumpis and Hongji Xu (el Chungo).

Y como se suele decir, en último lugar, pero no por menos importante, debo agradecer el apoyo incondicional de toda mi familia. Especialmente a Mayte, con quien inicié este largo camino y espero que me acompañe en todo lo que me queda por hacer. Te quiero mucho.

Y ahora, a otra cosa mariposa.

José López Vicario (el Paperman)

Castelldefells, junio 2006.

Contents

Notation

Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, upper-case italics denote sets, and lower-case italics denote scalars.

$\mathbb{N}, \mathbb{R}, \mathbb{C}$	The set of natural, real and complex numbers, respectively.
$\mathbb{R}^{n\times m}, \mathbb{C}^{n\times m}$	The set of $n \times m$ matrices with real- and complex-valued entries, respectively.
$\mathbf{X}^*$	Complex conjugate of the matrix $\mathbf{X}$.
$\mathbf{X}^T$	Transpose of the matrix $\mathbf{X}$.
$\mathbf{X}^H$	Complex conjugate and transpose (Hermitian) of the matrix $\mathbf{X}$.
$[\mathbf{X}]_{i,j}$	(i,j)th component of the matrix $\mathbf{X}$.
$\mathrm{Tr}(\mathbf{X})$	Trace of the matrix $\mathbf{X}$.
$\lvert\mathbf{X}\rvert$ or $\det(\mathbf{X})$	Determinant of the matrix $\mathbf{X}$.
$\mathrm{diag}(\mathbf{A})$	Vector constructed with the elements in the diagonal of matrix $\mathbf{A}$.
$\mathbf{I}_n$	Identity matrix of dimensions $n \times n$.
arg	Argument.
$\max, \min$	Maximum and minimum.
$(x)^+$	Positive part of the real scalar x, i.e., $(x)^+ = \max\{0, x\}$.
$\lceil x \rceil$	The lowest integer higher or equal to the real scalar x.
$\lvert x\rvert$	Modulus of the complex scalar x.
$\lVert\mathbf{x}\rVert$	Euclidean norm of the vector $\mathbf{x}$: $\lVert\mathbf{x}\rVert = \sqrt{\mathbf{x}^H\mathbf{x}}$.
$\mathbf{X}^{-1}$	Inverse of the matrix $\mathbf{X}$.
$\mathbf{X}^{\#}$	Moore-Penrose pseudo-inverse of the matrix $\mathbf{X}$.
$\mathbf{X}^{1/2}$	Hermitian square root of the positive semidefinite matrix $\mathbf{X}$, i.e., $\mathbf{X}^{1/2}\mathbf{X}^{1/2} = \mathbf{X}$.
$\sim$	Distributed according to.

$\text{Prob}(\cdot)$	Probability.
$\mathbb{E}[\cdot]$	Mathematical expectation.
$\mathcal{N}(\mathbf{m}, \mathbf{C})$	Real Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	Complex circularly symmetric Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
χ^2_n	Chi-square distribution with n degrees of freedom.
$\text{card}(\mathcal{A})$	Cardinality of the set $\mathcal{A}$, i.e., number of elements in $\mathcal{A}$.
$\text{Re}\{\cdot\}$	Real part.
$\text{Im}\{\cdot\}$	Imaginary part.
$\propto$	Equal up to a scaling factor (proportional).
$\triangleq$	Defined as.
$\simeq$	Approximately equal.
$\lim$	Limit.
$\log(\cdot)$	Natural logarithm.
$\log_a(\cdot)$	Base-a logarithm.
$E_i(\cdot)$	Exponential Integral function
$I_n(\cdot)$	nth-order modified Bessel function of the first kind
$J_0(\cdot)$	Zeroth-order Bessel function of the first kind.
$\Gamma_o(\cdot, \cdot)$	Incomplete Gamma function.
$\Gamma(\cdot, \cdot)$	Complementary Incomplete Gamma function.

Acronyms

3G	Third Generation.
AM	Adaptive Modulation.
AMC	Adaptive Modulation and Coding.
ARQ	Automatic Repeat-reQuest.
AS	Antenna Selection.
AWGN	Additive White Gaussian Noise.
BC	Broadcast Channel.
BER	Bit Error Rate.
BLAST	Bell-Labs Layered Space-Time.
BPSK	Binary Phase Shift Keying.
BS	Base Station.
CL	Cross-Layer.
cdf	cumulative density function.
CSI	Channel State Information.
DLC	Data Link Control.
DPC	Dirty Paper Coding.
DSP	Digital Signal Processor.
D-STTD	Double Space-Time Transmit Diversity.
FDD	Frequency Division Duplexing.
FS	Full Selection.
H-ARQ	Hybrid Automatic Repeat-reQuest.
HSDPA	High Speed Downlink Packet Access.
IEEE	Institute of Electrical and Electronics Engineers.
i.i.d.	independent and identically distributed.
MAC	Medium Access Control.
MIMO	Multiple-Input Multiple-Output.
MISO	Multiple-Input Single-Output.
ML	Maximum Likelihood.

MMSE Minimum Mean Square Error.
MRC Maximal Ratio Combining.
MRT Maximal Ratio Transmission.
MS Mobile Station.
MSE Mean Square Error.
MUD Multi-User Diversity.
N-SAW N Stop-and-Wait.
ORB Orthogonal Random Beamforming
OSI Open System Interconnection.
OSTBC Orthogonal Space-Time Block Code.
OSTBC-AS Orthogonal Space-Time Block Code with Antenna Selection.
pdf probability density function.
PER Packet Error Rate.
PHY Physical layer.
P/S Parallel to Serial.
PSK Phase Shift Keying.
QAM Quadrature Amplitude Modulation.
QoS Quality of Service.
Q-OSTBC Quasi-Orthogonal Space-Time Block Code.
QPSK Quadrature Phase Shift Keying.
RF Radio Frequency.
RS Restricted Selection.
RTT Round-Trip Time.
RX Receiver.
SAW Stop-and-Wait.
SDMA Space Division Multiple Access.
SIMO Single-Input Multiple-Output.
SIR Signal to Interference Ratio.
SINR Signal to Interference plus Noise Ratio.
SISO Single-Input Single-Output.
SISO-AS Single-Input Single-Output with Antenna Selection.
SM Spatial Multiplexing.
SMUD Selective Multi-User Diversity.
SNR Signal to Noise Ratio.
S/P Serial to Parallel.
STBC Space-Time Block Code.

STTC	Space-Time Trellis Code.
SVD	Singular Value Decomposition.
TCP	Transmission Control Protocol.
TDD	Time Division Duplexing.
TX	Transmitter.
UMTS	Universal Mobile Telecommunications System.
UTRA	UMTS Terrestrial Radio Access.
vs.	versus.
ZF	Zero Forcing.

Chapter 1

Introduction

1.1 Motivation

Future wireless communication systems aim to satisfy the ever increasing demand for high data rate communication services. In addition, these new communications schemes are expected to provide mobility and ubiquitous global coverage. In order to cope with these requirements, the traditional communication concept based on wired networks has had to evolve to a heterogeneous approach. In a heterogeneous network, global and mobile Internet is provided by introducing different access technologies (such as cellular networks, wireless local area networks or satellite constellations) to the terrestrial core network. In order to facilitate interoperability, standard Open Systems Interconnection (OSI) protocol stack is planned to be adopted. However, some drawbacks are related with the implementation of the layered OSI architecture in a wireless system due to the fragile nature of wireless links. In such a context, all constituent layers of the OSI protocol stack should work together in order to attain the highest possible system performance. For that reason, the interest in adopting cross-layer designs combining functionalities and information in different layers has been growing during the last years.

Communication schemes with multiple antennas at the transmit and/or receive edges are known to provide remarkable capacity improvements with respect to single-antenna configurations. Due to limitations in the radio spectrum available for wireless systems, multi-antenna approaches have been considered as promising techniques to increase capacity of future wireless systems. In a multiple-antenna context, channel capacity can be approached by conducting some sort of pre-processing on the transmit side. Unless reciprocity between the forward and reverse links can be assumed, a feedback channel is required to convey Channel State Information (CSI) to the transmitter. However, the amount of information allowed over feedback channels is limited. As a result, perfect and instantaneous CSI is seldom available at the transmitter, specially in those scenarios with fast fading and/or a high number of antennas. An effective solution with low feedback requirements is transmit antenna selection. By selecting the *best* sub-set of

transmit antennas, most of the gain provided by multi-antenna schemes can be obtained, but only a few bits must be fed back. As for the selection criteria, it is common practice to select the subset of transmit antennas maximizing metrics at the physical layer. Such approaches, though, do not exploit all the information available in specific system scenarios concerning the schemes and algorithms being used at the upper layers.

Within this framework, this PhD dissertation provides a contribution to the study of antenna selection algorithms from a cross-layer perspective. More precisely, by focusing our attention in the interaction between the lowest two layers of the OSI protocol stack (physical layer and data link layer), we design and study antenna selection algorithms aimed at maximizing performance at the data link layer. Our study begins with a single user scenario and is extended to the downlink of a multi-user environment, where scheduling algorithms based on cross-layer strategies are encompassed. The interests of this thesis are devoted to study the broadcast channel because the multiple-access channel has been widely discussed in the literature. Indeed, most of current research in wireless communications is oriented to the broadcast channel due to the massive demand for high-rate downlink applications.

1.2 Outline of Dissertation

The focus of this dissertation is on the study of antenna selection mechanisms in single- and multi-user scenarios. In contrast with other works based on restricting the optimization procedure to the physical layer, we consider a cross-layer approach. More specifically, we adopt methodologies aimed at maximizing data link layer performance. Special emphasis is given to those strategies requiring low amount of information in the feedback channel. For that reason, when the study is extended to the multi-user case, we consider the use of opportunistic scheduling algorithms as only partial CSI must be reported by the different users. Next, we describe the topics addressed in the remaining chapters of this dissertation.

Chapter 2 presents an overview of cross-layer designs and multi-antenna techniques. Besides, a complete description of algorithms and classical results related to antenna selection mechanisms is given.

Chapter 3 considers a single-user Multiple-Input Multiple-Output (MIMO) system with a Hybrid Automatic Repeat-reQuest (Hybrid-ARQ) strategy at the data link layer. In that scenario, we investigate a transmit antenna selection algorithm capable of adapting the number of active antennas to varying channel conditions. We address a cross-layer methodology in the sense that the *criterion* for the selection of antenna subsets is the maximization of *data link layer throughput*, which takes into account characteristics both at the physical and data link layers. We also analyze performance trade-offs with respect to other existing antenna selection algorithms where the number of streams remains constant but dynamically mapped into different antennas

and, also, with respect to conventional selection criteria based on channel capacity measures. In order to enhance system performance, adaptive modulation schemes are included to jointly perform antenna selection *and* rate adaptation. Finally, a problem based on switching between spatial diversity and spatial multiplexing is also addressed from a cross-layer point of view. More specifically, the transmitter is provided with three transmission schemes exploiting different degrees of spatial diversity and multiplexing gains. Then, according to scenario conditions, the transmission scheme maximizing link layer throughput is selected.

Chapter 4 considers a multi-user system where only one user is scheduled for transmission in each time-slot by means of an opportunistic scheduling algorithm. In that situation, antenna selection algorithms are adopted in order to further exploit spatial and multi-user diversity gains with low feedback requirements. In that chapter, different scenarios are considered in terms of the information available at the scheduler: ideal scenario, selective-feedback scheduling and degraded information due to feedback delays or noisy channel estimates. Then, in order to provide the system with robustness against wrong scheduling decisions, antenna selection mechanisms are combined with Orthogonal Space-Time Block Coding (OSTBC) approaches.

Chapter 5 adopts Orthogonal Random Beamforming (ORB) with the aim of serving several users simultaneously in the same time-slot. The purpose of this chapter is to improve ORB performance in scenarios with a reduced (practical) number of users. To do so, we adopt the idea of antenna selection to develop a beam selection mechanism. Instead of activating all beamforming vectors, only those beams maximizing link layer throughput are selected. Due to the combinational nature of the optimization process (all user-beam combinations must be tested), the optimum solution requires an exhaustive search. Several sub-optimum algorithms are proposed and performance vs. complexity trade-offs are assessed. However, all the proposed beam selection procedures require all the channel gains. In order to circumvent that, a beam selection approach requiring only the Signal to Interference plus Noise Ratio (SINR) information of the available beam sub-sets is proposed. Finally, the effects of feedback quantization on the proposed beam selection strategy are studied as well.

Chapter 6 concludes this PhD dissertation with a summary and discussion of the obtained results. Some suggestions for future work in the field are also outlined.

1.3 Research Contributions

The main contribution of this thesis is the study of antenna selection algorithms from a cross-layer perspective in single- and multi-user systems. Next, the details of research contributions in each chapter are presented.

Chapter 3

The main results of this chapter addressing the study of antenna selection algorithms in a MIMO single-user scenario have been published in one journal paper and nine conference papers:

- J. L. Vicario, M. A. Lagunas, C. Antón-Haro, "A Cross-Layer Approach to Transmit Antenna Selection", accepted for publication in IEEE Trans. on Wireless Comm., 2006.
- J. López-Vicario, C. Antón-Haro, "Full vs. Restricted Transmit Antenna Selection Schemes for MIMO Cross-layer Designs", in Proc. 12th Wireless World Research Forum Meeting (WWRF12), Nov. 2004.
- J. López-Vicario, C. Antón-Haro, "Adaptación de la Velocidad de Transmisión en Sistemas Multi-antena HSDPA Mediante Criterios Cross-layer", in Proc. URSI, Sept. 2004.
- J. López-Vicario, C. Antón-Haro, "Joint Transmit Antenna Selection and Adaptive Modulation in Cross-layer Oriented Designs for HSDPA Systems", in Proc. IEEE Statistical Array and Multi-channel Signal Processing Workshop (SAM), July 2004.
- J. López-Vicario, C.F. Mecklenbrauker, C. Antón-Haro, "Reduced-complexity Methods for Throughput Maximization in MIMO Channels", in Proc. IEEE International Communications Conference (ICC), June 2004.
- J. López-Vicario, C. Antón-Haro, "Transmit antenna Selection for Rate Adaptation in HSDPA Systems", in Proc. Wireless World Conference (WWC), May 2004.
- J. López-Vicario, C. Antón-Haro, "Performance and Complexity Issues in Cross-Layer MIMO Designs for HSDPA", in Proc. 8th Wireless World Research Forum Meeting (WWRF8bis), Feb. 2004.
- J. López-Vicario, C. Antón-Haro, "Optimización del Throughput en Sistemas V-BLAST Usando Configuración de Antenas de Transmisión Adaptativa", in Proc. Jornadas Telecom I+D, Nov. 2003.
- J. López-Vicario, C. Antón-Haro, "Throughput Optimization for MIMO Systems via Cross-layer Designs", in Proc. 9th Wireless World Research Forum Meeting (WWRF9), July 2003.

The work related with the adaptive switching between diversity and multiplexing transmission schemes has been presented in one conference paper:

- J. L. Vicario, C. Antón-Haro, "Adaptive Switching between Spatial Diversity and Multiplexing: a Cross-layer Approach", in Proc. IST Mobile & Wireless Communications Summit, June 2005.

Chapter 4

The main results of this chapter are related to the cross-layer interaction between spatial and multi-user diversity in scenarios with different assumptions for the CSI available at the scheduler. The results have been presented in two journal papers and six conference papers.

- J. L. Vicario, C. Antón-Haro, "Spatial vs. Multi-user Diversity Trade-offs for Cross-layer Scheduling in Limited Feedback Systems", accepted for publication in Eurasip Signal Processing Journal - Special issue on Signal Processing-assisted Cross-layer designs, 2006.
- J. L. Vicario, C. Antón-Haro, "Analytical Assessment of Multi-user vs. Spatial Diversity Trade-offs with Delayed Channel State Information", accepted for publication in IEEE Communications Letters, 2006.
- J. L. Vicario, C. Antón-Haro, "Cross-layer Interaction Between Spatial and Multi-user Diversity In Selective Feedback Systems: Outage Capacity Analysis", in Proc. IEEE International Wireless Communications and Mobile Computing Conference (IWCMC), July 2006.
- J. L. Vicario, C. Antón-Haro, "A Unified Approach to the Analytical Assessment of Multi-user Diversity with Imperfect Channel State Information", in Proc. European Wireless Conference, April 2006.
- J. L. Vicario, C. Antón-Haro, "Analytical Assessment of Capacity vs. Robustness trade-offs in Systems with Selective Multi-user Diversity", in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May 2006.
- J. L. Vicario, C. Antón-Haro, "Joint Exploitation of Spatial and Multi-user Diversity via Space-time Block-coding and Antenna Selection", in Proc. IEEE International Communications Conference (ICC), May 2005.
- J. L. Vicario, C. Antón-Haro, "Robust Exploitation of Spatial and Multi-user Diversity in Limited-Feedback Systems", in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2005.
- J. L. Vicario, C. Antón-Haro, "Spatial vs. Multi-user Diversity Trade-offs in SMUD Systems", in Proc. Winterschool on Coding and Information Theory, Feb. 2005.

Chapter 5

The results in this chapter extend the antenna selection concept to a multi-user scenario adopting an orthogonal random beamforming scheme by means of beam selection algorithms. This work has been presented in two conference papers and one journal paper is still in preparation:

- J. L. Vicario, R. Bosisio, U. Spagnolini, C. Antón-Haro, "A Throughput Analysis for Opportunistic Beamforming with Quantized Feedback", in Proc. IEEE Personal, Indoor and Mobile Radio Communications (PIMRC), Sept. 2006.
- J. L. Vicario, R. Bosisio, U. Spagnolini, C. Antón-Haro, "Adaptive Beam Selection Techniques for Opportunistic Beamforming", in Proc. IEEE Personal, Indoor and Mobile Radio Communications (PIMRC), Sept. 2006.

Other contributions not presented in this dissertation

During the PhD period, several collaborations has been carried out with the Politecnico di Milano and the KTH resulting in several conference papers and one journal paper in preparation.

Results concerning the comparison of OSTBC and ORB schemes in SMUD scenarios has been presented in:

- R. Bosisio, J. L. Vicario, C. Antón-Haro, U. Spagnolini, "Diversity-Multiplexing Tradeoff in Multi-user Scenario with Selective Feedback", in Proc. IST Mobile & Wireless Communications Summit, June 2006.

The impact of imperfect feedback information on a system based on a cluster-eigenbeamforming approach has been discussed in:

- R. Bosisio, J. L. Vicario, C. Antón-Haro, U. Spagnolini, "The Effect of Imperfect Feedback on Opportunistic Beamforming Schemes", in Proc. European Wireless Conference, April 2006.

Finally, a comparative study between optimal and Zero Forcing beamforming in combination with practical bit-allocation algorithms has been published in:

- M. Bengtsson, D. Bartolome, J. L. Vicario, C. Antón-Haro, "Beamforming and Bit-loading Strategies for Multi-User SDMA with Admission Control", in Proc. IEEE Personal, Indoor and Mobile Radio Communications (PIMRC), Sept. 2005.

Chapter 2

An Overview of Cross-layer Designs and MIMO Systems

This chapter is devoted to provide the reader with some background on cross-layer mechanisms and the multi-antenna techniques encompassed in this work. First, cross-layer principles are presented and some examples justifying the utilization of cross-layer designs in wireless communication schemes are given. After that, the state of the art on Multiple-Input Multiple-Output (MIMO) techniques along with a complete overview on antenna selection mechanisms is described. Finally, some fundamental results and practical schemes concerning MIMO systems in the downlink of a multi-user scenario are discussed.

2.1 Cross-layer Principles

A first attempt to standardize the new wireless communication systems (3G and beyond) has been the direct application of the standard OSI protocol stack utilized in the Internet. The OSI layered architecture is divided into seven clearly defined functional layers [Ber92], as it can be observed in Fig. 2.1, and is considered one of the main foundations of the Internet. This is because, thanks to its modular architecture, research areas have been parallelized and, thus, technological development has been accelerated. As a result, services offered to the subscribers have improved and proliferated. On the other hand, by following a layered philosophy, networks of different nature can be easily interconnected.

However, the current layered design was originally designed for wireline environments and, then, several difficulties appear when it is adopted in a wireless scenario. In particular, wireless scenarios introduce special problems that do not exist in wireline networks. These include high error rate, bursty errors, location-dependent and time varying wireless link capacity, user mobility, and power constraint of the mobile hosts. Therefore, in such an environment, information exchange between layers seems to be mandatory in order to improve the Quality of

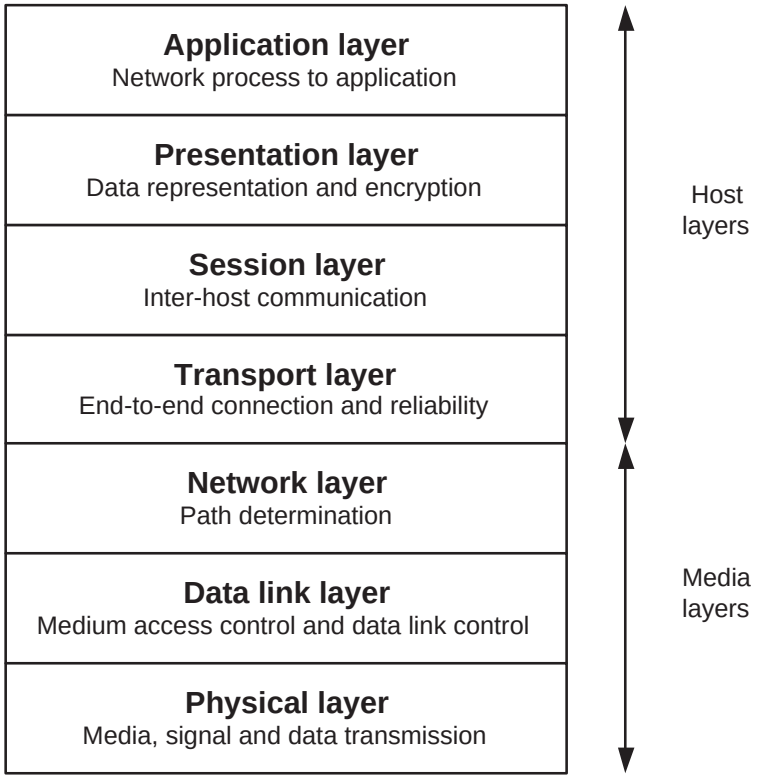


Figure 2.1: OSI protocol stack.

Service (QoS). For that reason, the interest in adopting cross-layer designs devoted to combine functionalities and information between layers has been growing recently.

One of the pioneering examples that reflected the need for exchanging information between layers in the wireless community was the paper by Knopp and Humblet [Kno95]. It was shown that, in the uplink of a wireless multi-user system, the average cell throughput can be increased when in each slot the user with the best channel conditions is scheduled. That is, by simply reporting the channel state (physical layer) to the scheduler (data link layer) the system performance can be substantially improved.

Another situation that clearly justifies the utilization of cross-layer design is the adoption of the Transmission Control Protocol (TCP) in a wireless scenario. The TCP protocol is not aware of packet loss nature, i.e., all the packet losses are treated as congestion losses. Therefore, in a wireless environment, where packet losses are also produced by transmissions errors, the system throughput can be considerably reduced by the congestion avoidance and retransmission procedures [Xyl99]. A number of solutions aimed at avoiding such behavior exist, however, substantial modifications of the TCP protocol are required. Such modifications are not welcome due to the economical consequences that are implied. As a possible alternative, an Automatic Repeat-reQuest (ARQ) strategy [Lin84] at the data link layer can be incorporated to reduce packet losses. Nevertheless, inter-layer coordination must be considered to counteract conflicts

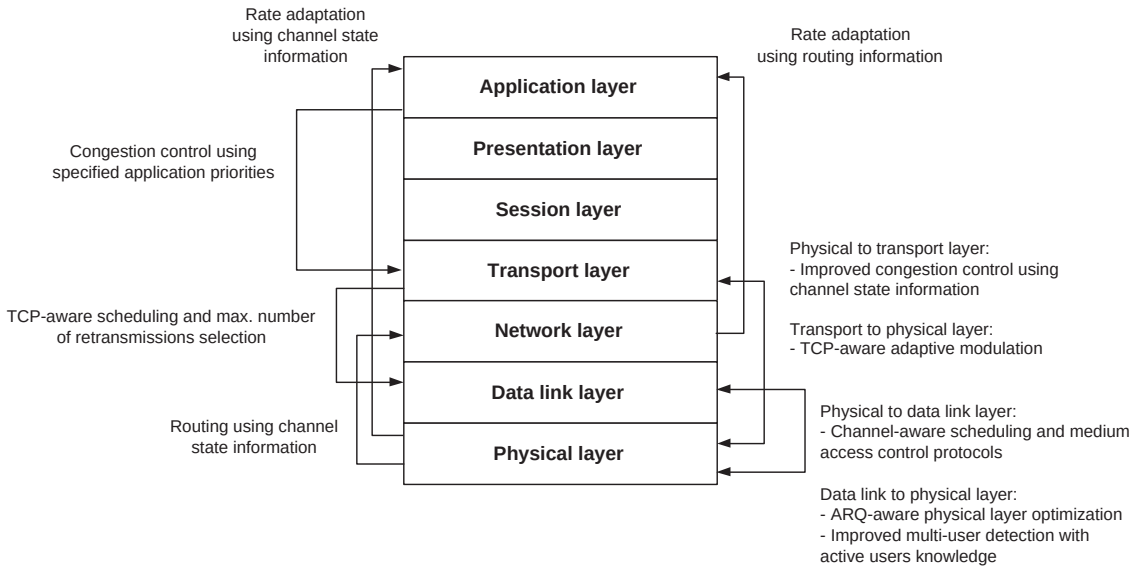


Figure 2.2: Cross-layer interaction examples.

between TCP and ARQ protocols [Car04]. For instance, delays associated with ARQ retransmissions can deteriorate the Round-Trip Time (RTT) estimation. This estimation is used by TCP to scale the growth of the congestion window, and, then, end-to-end throughput can be harmed. To alleviate that, information about the link layer ARQ strategy can be sent to TCP in order to improve the RTT estimation.

The above-mentioned examples are not the only ones in which a cross-layer philosophy is reflected. For instance, in [Sha03, Rai04, DS05, Sri05] and references therein more examples can be found (see Fig. 2.2). A complementary point of view in identifying and managing the cross-layer interactions is proposed in [Car04]. In that paper, the problems associated with the layered architecture in a wireless system are identified and classified into four coordinated planes: security, QoS, mobility and wireless link adaptation. After that, a cross-layer manager is assigned to each control plane. This control manager is responsible for collecting the information of the involved layers and managing cross-layer algorithms aimed at optimizing the objective of the corresponding control plane.

It is worth noting that cross-layer algorithms must be carefully designed. As noticed in [Kaw05], consequences difficult to foreseen appear when functionalities in different layers are interacted. In some cases, system stability becomes an issue as it is difficult to appreciate how a new modification will affect system performance due to the multiplicity of existing interactions. Besides, cross-layer optimization implies an increased system complexity and, therefore, complexity vs. performance trade-offs must be assessed. In summary, cross-layer design is expected to provide substantial gains in wireless networks, but one must carefully analyze all possible consequences and justify the increased complexity.

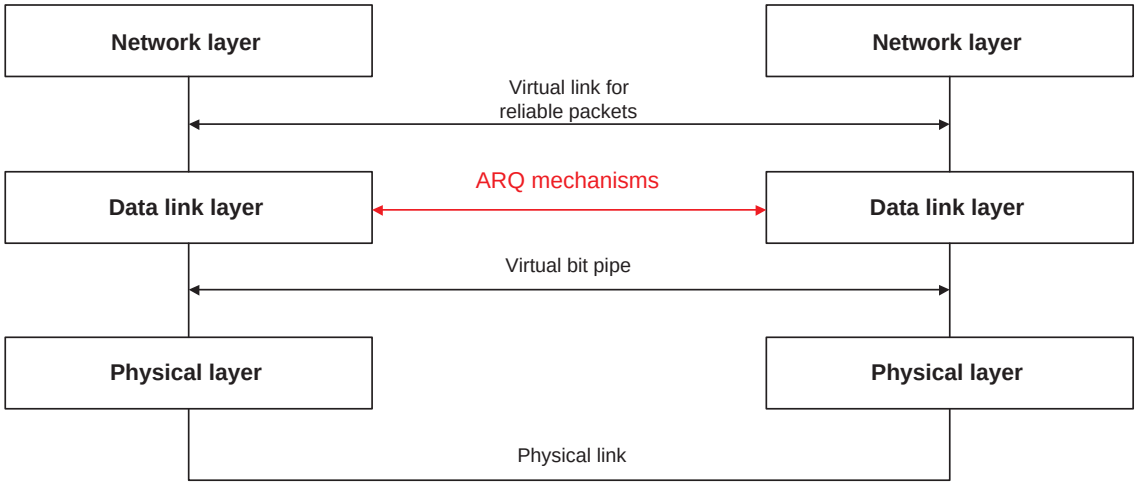


Figure 2.3: DLC functionality of the data link layer.

2.2 Cross-layer Interaction between Physical and Data Link Layers

As mentioned in the previous chapter, the interests of this thesis lie in the cross-layer interaction between the physical (PHY) and data link layers. In order to provide the reader with an overview in this topic, we next present some cross-layer techniques grouped according to the functionalities exploited at the data link layer: Data Link Control (DLC), which performs error and flow control of the transmission link, and Medium Access Control (MAC), responsible for controlling access and managing the available resources of the shared medium.

2.2.1 PHY-DLC Interaction: ARQ-aware Physical Layer Optimization

The purpose of the DLC functionality is to provide the network layer with an error-free virtual link (see Fig. 2.3). To do so, ARQ mechanisms can be employed in the link between the transmit and receive nodes. In this context, there exists a recent trend aimed at directly maximizing data link layer metrics rather than PHY-layer parameters like the Signal to Noise Ratio (SNR), Bit Error Rate (BER) or mutual information. In other words, system design is devoted to maximize the effective number of transmitted bits per second while taking into consideration ARQ at the data link layer. For instance, in a specific communication system (with a given receiver structure, packet size, modulation scheme, ARQ strategy...), selecting the transmit antenna subset that maximizes the data link layer throughput instead of the mutual information seems to be more appropriate. This concept was considered in [Mil02a] and, as expected, considerable throughput gains were achieved. In [Mil02b, Mil06], power allocation was also carried out but it was shown that minor performance gains can be obtained with respect to antenna selection only.

In [Yoo06c], parameters in the physical layer were also optimized with the aim of maximizing

data link layer throughput. In that paper, a generic expression of throughput was derived as a function of the data-rate, modulation scheme and packet size. Although joint optimization was performed, it was shown that not all three parameters need to be jointly adapted. For the high SNR region only packet size and the constellation size matter, since data rate cannot be further increased due to bandwidth limitations. In the low SNR region, the constellation size cannot be reduced below the lowest available modulation scheme. In this case, it was shown that the optimal strategy is to fixe an optimal packet size and only adapt the symbol rate to ensure a received SNR above a specific threshold.

In [Liu04], the SNR thresholds of an Adaptive Modulation and Coding (AMC) scheme were determined with information from the data link layer. In particular, the information was the maximum number of retransmissions, chosen to fulfill delay requirements, and the maximum probability of packet loss tolerated by the system. Apart from that, the existing trade-off between delay requirements and throughput was discussed. That is, with a higher number of retransmissions the packet loss is reduced, higher AMC levels can be used and, as a result, throughput is improved; but packet delay is increased too. In practice, a desirable delay-throughput trade-off can be achieved with a small number of retransmissions. A similar approach was adopted in [Gon04]. In that paper, however, SNR thresholds were adaptively modified in accordance with the number of packet retransmissions in the ARQ process. In order to obtain a high spectral efficiency, a low value of SNR threshold was chosen in the first transmission of the packet; whereas a more conservative election of the thresholds was adopted for an increasing number of retransmissions. Differently from the approaches mentioned above, the size of the queue at the data link layer was assumed finite in [Liu05]. In such a situation, packets are dropped when the queue is full. Then, aimed at maximizing the average throughput by keeping a reduced probability of packet losses, the cross-layer mechanism was redesigned to optimize PER thresholds in accordance with the queue size.

The work in [Liu04] was extended to exploit multiple antenna capabilities at the PHY layer in several papers. In [Maa04b], the authors considered an Orthogonal Space-Time Block Coding (OSTBC) aproach to improve link reliability. The generalization of that work to Nakagami fading channels [Nak60] was carried out in [Maa04a, Fem04] and the effect of channel estimation errors on system performance was analyzed in [Maa05]. In [Lu05], rate adaption was carried out by means of an adaptive switching between space-time codes with different transmission rates.

2.2.2 PHY-MAC Interaction: Opportunistic Scheduling

As commented in the previous section, conventional scheduling algorithms based on wireline communication systems cannot be directly applied to wireless communication systems. For that reason, wireless scheduling algorithms must be designed taking into account wireless channel

characteristics. In that context, the cross-layer design remains in the fact that the scheduling algorithm (data link layer) uses physical layer information in order to combat wireless channel adversities.

In the uplink of a Single-Input Single-Output (SISO) cellular wireless system with constraints on the long-term average power, the system throughput is maximized when water-filling is applied in the time-domain [Kno95]. When doing so, under unfavorable channel states no user transmits. Otherwise, only the user with the best channel conditions is scheduled, whereby more power is allocated to favorable channel states. Similar conclusions were obtained for the SISO broadcast channel[1] (downlink) in [Tse97]. This strategy is known as opportunistic scheduling and relies on the assumption that different users experience independent fading processes, i.e., the scheduler exploits the *Multi-User Diversity* (MUD) gains of the system. In particular, if channel responses associated with the different users are i.i.d. Rayleigh distributed, a growth-rate in the system capacity equal to $\log_2(\log_2(K))$ can be obtained (with K standing for the number of users in the system). If channel conditions are not statistically equal for all users, these algorithms may be unfair among users. That is, users with the worst channels are never scheduled. To circumvent that, a Proportional Fair Scheduling algorithm was proposed in [Hol00, Vis02]. By monitoring the throughput of the different users, the scheduler raises the priority to those users that do not transmit for too long.

In order to exploit MUD gains, only partial Channel State Information (CSI), i.e. SNR, has to be estimated by the terminals and reported back to the base station over a feedback channel. In [Flo03] the authors showed that most of the MUD can still be extracted when partial CSI is quantized with a few bits. Going one step further, authors in [San05] proved that indeed multiuser diversity can be exploited with only one bit in the feedback channel, provided that the threshold value of the quantization is appropriately optimized. It was shown that, besides from preserving the growth-rate $\log_2(\log_2(K))$, almost the same capacity as that with unquantized feedback can be obtained. In the mentioned works, all users must continually report their channel status to the scheduler. A Selective MUD (SMUD) approach was introduced by Gesbert *et al.* [Ges03] (and extended to a non-homogenous scenario in [Yan04]) in order to reduce the amount of feedback load generated by CSI signalling. By exclusively letting users report CSI when channel quality exceeds a pre-defined threshold, remarkable benefits can be obtained in terms of reduced load in the feedback channel at the expense of moderate performance losses. These schemes work well in systems with a large number of active users. System performance deteriorates when the number of users is low or the feedback channel load is considerably restricted. The same effect is observed when the CSI available at the scheduler is degraded, since wrong scheduling decisions may be done (further details about this topic can be found in Chapter 4).

[1] In the rest of this section, we focus our attention on the broadcast channel due to the interests of this PhD dissertation.

All the described algorithms are based on the exploitation of the MUD gain, which relies on the dynamic range of the channel fluctuations. Therefore, in scenarios with reduced dynamic range, or slow fading, MUD cannot be efficiently exploited. Still, channel fluctuations can be easily induced with multiple antennas at the base station as shown in [Vis02]. The authors in that paper proposed a method, referred to as opportunistic beamforming, where in each time-slot a random beamformer is generated. In scenarios with a high number of users with independent channels, the probability that the generated random beamformer matches the channel signature of one user is high. In the asymptotic case, the same performance as that obtained with coherent (optimal) beamforming can be achieved. However, this is not the best way to exploit multi-antenna capabilities in a multi-user system as we will show later.

2.3 Single-user MIMO Systems

Multiple-Input Multiple-Output (MIMO) wireless systems are known to provide remarkable benefits in scenarios with rich scattering as the multi-path fading is exploited in a beneficial way. More specifically, the key concept is that MIMO systems take advantage of multi-path fading to send several streams in parallel. Then, the link reliability of the system can be improved by sending replicas of the same symbol or the data-rate can be increased by sending independent symbols in parallel.

This section is devoted to provide the reader with some background on MIMO wireless schemes. We first present some of the fundamental limits obtained in the study of MIMO wireless systems from an information theory viewpoint. After that, we describe some practical schemes according to the purpose for which MIMO techniques are considered: spatial diversity or spatial multiplexing. Finally, since the main topic of this thesis is the study of antenna selection techniques from a cross-layer perspective, we conclude this section with a complete overview of algorithms and some classical results related with antenna selection mechanisms.

2.3.1 An Overview of MIMO Systems

A MIMO wireless system is a communication link where both the transmitter and the receiver are equipped with multiple antennas. In Fig. 2.4, we show a typical MIMO wireless system with M transmit antennas and N receive antennas. Usually, the MIMO signal model is represented in matrix form as follows:

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{n}$$

where $\mathbf{r} \in \mathbb{C}^{N\times 1}$ is the received signal vector, $\mathbf{H} \in \mathbb{C}^{N\times M}$ is the channel matrix whose elements are the channel responses between each pair of antennas, $\mathbf{s} \in \mathbb{C}^{M\times 1}$ denotes the transmitted symbols and $\mathbf{n} \in \mathbb{C}^{N\times 1}$ stands for an additive Gaussian noise vector of complex, random variables

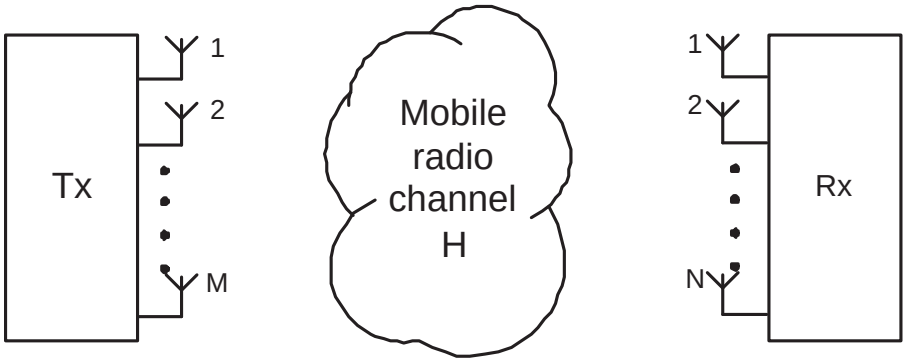


Figure 2.4: Block diagram of a MIMO wireless system.

with zero mean and unit variance, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$. We define $\bar{\gamma}$ as the received SNR at any receive antenna and, due to the noise normalization, it is equal to the total transmit power at the transmitter. Concerning CSI, it is commonly assumed perfectly known at the receiver (this can be easily arranged by training), whereas different considerations are adopted for the knowledge at the transmit side.

We start by considering the case where CSI is completely known at the transmitter. In this case, the transmitter configuration can be adapted to all the channel realizations. The mutual information between vectors $\mathbf{s}$ and $\mathbf{r}$ conditioned on $\mathbf{H}$ and expressed in terms of bits/s/Hz is given by:

$$I(\mathbf{s};\mathbf{r}) = \log_2 \det\left(\mathbf{I}_N + \mathbf{H}\mathbf{Q}\mathbf{H}^H\right)$$

where $\mathbf{Q}$ is the covariance matrix of the input signal, $\mathbf{Q} = \mathbb{E}[\mathbf{s}\mathbf{s}^H]$. The instantaneous capacity of the MIMO channel is obtained by maximizing the above expression [Tel95, Fos96]:

$$C_{inst} = \max_{\mathbf{Q}:Tr(\mathbf{Q})\leq\bar{\gamma}} I(\mathbf{s};\mathbf{r}) = \max_{\mathbf{Q}:Tr(\mathbf{Q})\leq\bar{\gamma}} \log_2 \det\left(\mathbf{I}_N + \mathbf{H}\mathbf{Q}\mathbf{H}^H\right)$$

In [Tel95], it was shown that the above equation is maximized by transmitting through the eigenmodes of the channel matrix. In other words, $\mathbf{H}$ must be decomposed by means of the Singular Value Decomposition (SVD), i.e., $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. Then, the transmitter uses the columns of $\mathbf{V}$ as beamfoming vectors, whereas the columns of $\mathbf{U}$ are employed at the receiver to estimate the different transmitted signals. Finally, power allocation is carried out following a waterfilling strategy [Cov91, Tel95]. As a consequence, the instantaneous capacity turns out to be equal to:

$$C_{inst} = \sum_{i=1}^{r} \log_2 \left(\mu\lambda_i\right)^+$$

where r and λ_i stand for the rank and the non-zero eigenvalues of the matrix $\mathbf{H}\mathbf{H}^H$ respectively, and μ is the *water-level* chosen to satisfy

$$\bar{\gamma} = \sum_{i=1}^{r} \left(\mu - \lambda_i^{-1}\right)^+$$

If channel state information (CSI) is not known at the transmitter, rate adaption cannot be carried by the transmitter. Instead, a constant rate is used for all the channel realizations and two statistics are commonly used to characterize it: ergodic and outage capacity.

Ergodic capacity is appropriate for those cases where codewords can span all the channel realizations, accomplished in systems without delay constraints or in scenarios with fast fading. This is because ergodic capacity is defined as the maximum rate achievable by averaging over all the realizations of $\mathbf{H}$:

$$C_{erg} = \max_{\mathbf{Q}:Tr(\mathbf{Q})\leq\bar{\gamma}} \mathbb{E}_{\mathbf{H}} \left[\log_2 \det\left(\mathbf{I}_N + \mathbf{HQH}^H\right)\right]$$

For the i.i.d Rayleigh fading case, Telatar [Tel95] showed that the optimal strategy consists in uniformly distributing the available power among the transmit antennas, that is:

$$\mathbf{Q} = \frac{\bar{\gamma}}{M}\mathbf{I}_M$$

By doing so, the resulting ergodic capacity is

$$C_{erg} = \mathbb{E}_{\mathbf{H}} \left[\log_2 \det\left(\mathbf{I}_N + \frac{\bar{\gamma}}{M}\mathbf{HH}^H\right)\right]$$

At high SNR, the ergodic capacity linearly grows as $\min(M, N)$ [Fos98, Tel95]. Such an effect is one of the main motivations for the use of MIMO approaches and is known as spatial multiplexing gain. In particular, the improvement obtained with full CSIT with respect to the blind case is the array gain resulting from adapting the transmission configuration to the channel *direction*, which is negligible in the asymptotic SNR regime. It is worth noting that the capacity can be significantly degraded when the correlation between antennas increases [Shi00, Mes03, Kie98] or in Ricean channels with a strong line-of-sight component [Dri99].

In delay-limited applications where channel coding is conducted over a (potentially) low number of frames, the outage capacity becomes more relevant [Oza94]. Outage capacity is defined as the level of performance assured with a certain level of reliability. In other words, outage capacity $C_{out,q}$ is defined as the rate supported the $(100-q)\%$ of the channel realizations, that is:

$$\text{Prob}\left(\log_2 \det\left(\mathbf{I}_N + \mathbf{HQH}^H\right) < C_{out,q}\right) = q\%$$

The expression above is usually not available in closed-form and it is common practice to resort to numerical methods.

For the case where CSI at the transmitter is partial different strategies hold. If only the covariance matrix or the channel matrix is known at the transmitter, the transmission is given through the eigenvectors of that matrix [Vis01, Jaf01a]. For the case that only the norm of the channel coefficients are known, independent signalling must be carried out over the different transmit antennas [Pay06]. In particular, the optimal power allocation for the 2×2 case can be found in closed-form in [Pay04].

Thus far we have discussed the information theoretical limits of MIMO wireless systems, where it has been assumed that the system is implemented with ideal Gaussian codebooks and receivers with infinite complexity. In real systems, however, suboptimum signal constellations with practical coding schemes are utilized along with complexity-limited receivers. The fundamental limits are then useful to indicate how far is the designed communication system from the fundamental upper bound. Traditionally, practical MIMO systems are designed with the aim of exploiting either spatial diversity or spatial multiplexing gains. In accordance with that classification, we next present some practical schemes. Some emphasis is given to the Vertical Bell-Labs LAyered Space-Time (V-BLAST) and OSTBC architectures as these schemes will be adopted in Chapters 3 and 4, respectively.

Spatial Diversity

In a MIMO environment with rich scattering, the link reliability can be considerably improved since the probability of losing the signal decreases exponentially as a function of the number of uncorrelated paths. Practical schemes based on that concept are usually designed to maximize the received SNR of the link. According to the degree of CSI knowledge at the transmitter[2], different strategies can be taken into consideration. When the *CSI is perfectly known* at the transmitter, the optimal strategy consists in matching the transmit beamformer to the channel by means of Maximal-Ratio Transmission (MRT) (or optimal transmit beamforming), i.e., transmitting through the right eigenvector associated with the largest eigenvalue of $\mathbf{H}$ [Pau03, Lar03] (recall that matrix $\mathbf{H}$ can be expressed as $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$). At the receiver, Maximal-Ratio Combining (MRC) is performed in the sense that the left eigenvector associated to the maximum eigenvalue is selected as receive beamformer. In this case, the diversity order (i.e., the number of uncorrelated paths or the slope as the BER decays at high SNR on a log-log scale) is equal to MN.

The most popular transmission schemes designed to exploit spatial diversity *without CSI* at the transmitter are Space-Time Trellis Codes [Tar98] and Orthogonal Space-Time Block Coding [Ala98][Tar99]. Space-Time Trellis Codes (STTC) are the extension of conventional trellis codes [Big91] to the multi-antenna case. In addition to the diversity gain, these schemes provide benefits in terms of coding gain. However, a multidimensional Viterbi decoder is required at the receiver, which results in a prohibitive complexity inappropriate for mobile terminals. For that reason, more attention has been paid to the use of Orthogonal Space-Time Block Coding (OSTBC) since optimal reception can be carried out with low complexity receivers. Consider, for instance, a system where an OSTBC system with two transmit antennas is adopted at the transmitter and the receiver is equipped with a single receive antenna. By stacking two consecutive data

[2]Recall that CSI is assumed perfectly known at the receiver.

samples, the received vector $\mathbf{r} = [r_1, r_2]^T$ can be written as:

$$\mathbf{r} = \mathbf{Sh} + \mathbf{n}$$

where $\mathbf{h} = [h_1, h_2]^T$ is the channel vector, $\mathbf{n}$ stands for an additive Gaussian noise vector of complex, random variables with zero mean and unit variance, and $\mathbf{S}$ is the symbol block that describes the space-time block code. For the Alamouti case, the matrix $\mathbf{S}$ is given by the following expression:

$$\mathbf{S} = \begin{bmatrix} s_1 & s_2 \\ s_2^* & -s_1^* \end{bmatrix}$$

The received signal can be reformulated as:

$$\mathbf{y} = \begin{bmatrix} r_1 \\ r_2^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ -h_2^* & h_1^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix} = \mathbf{H}_{eq}\mathbf{s} + \mathbf{v}$$

From the above expression, one can extract the main particularity of OSTBC: the resulting equivalent channel matrix $\mathbf{H}_{eq}$ is orthogonal, that is

$$\mathbf{H}_{eq}^H \mathbf{H}_{eq} = h^2 \mathbf{I_2}$$

where $h^2 = |h_1|^2 + |h_2|^2$ is the channel gain.

As a consequence, transmitted symbols can be easily decoupled at the receiver by using simple matched filtering. That is, we can estimate the symbols as:

$$\hat{\mathbf{s}} = \mathbf{H}_{eq}^H \mathbf{y} = h^2 \mathbf{s} + \mathbf{H}_{eq}^H \mathbf{v}$$

Due to the orthogonality of $\mathbf{H}_{eq}$, the components of the noise vector are i.i.d. and, hence, independent optimum ML detection can be carried out for each transmitted symbol. By doing so, the number of decoding metrics for ML detection is reduced from 2^{2^b} to 2×2^b. In addition, the received SNR is the same for the two transmitted symbols and can be written as:

$$\gamma = \frac{\bar{\gamma}}{2}\left(|h_1|^2 + |h_2|^2\right)$$

where $\bar{\gamma}$ stands for the total transmitted power at the transmitter. In the above expression, one can observe that the diversity order is equal to two if the channel fades are uncorrelated.

For the general case with M transmit and N receive antennas, a diversity order equal to MN can be achieved by constructing the symbol matrices in such a way that orthogonality is preserved in the resulting equivalent channel matrix:

$$\mathbf{H}_{eq}^H \mathbf{H}_{eq} = \sum_{i=1}^{N} \sum_{j=1}^{M} |h_{i,j}|^2 \mathbf{I_M}$$

where $h_{i,j}$ stands for the channel response between the $i-th$ receiver and the $j-th$ transmitter.

When complex constellations are used, the only known orthogonal-based scheme allowing full rate is the Alamouti scheme, i.e. $M = 2$. With three and four transmit antennas, for instance, a rate equal to 3/4 can be achieved. Some schemes have been proposed in order to achieve higher rates at the expense of some diversity loss. Jafarkhani proposed a Quasi-Orthogonal STBC scheme able to achieve full rate with four transmit antennas [Jaf01b]. Full diversity cannot be achieved, but it was shown that better performance than with orthogonal designs can be obtained in the low-SNR regime. The extension of Q-OSTBC schemes to a higher number of transmit antennas was derived in [Rup02, Mec04].

Although the same diversity order can be achieved with OSTBC and optimal beamforming, better performance can be obtained with the latter one due to the array gain resulting from focusing all the power in the *channel direction*. However, perfect CSI is seldom available at the transmitter. Some works in the literature proposed improving OSTBC performance with partial CSI. In order to reduce the gap between OSTBC and optimal beamforming, the outputs of the OSTBC scheme can be linearly combined by means of a linear precoder which is optimized with the partial CSI available at the transmitter. In [Gan02], the linear precoder was restricted to be a diagonal matrix resulting in a scheme where the only function was distributing the power among the different transmit antennas. It was shown that if a non-delayed and error-free feedback channel is available, the BER can be minimized by selecting the antenna with the largest gain. For the case that the feedback channel is subject to errors, the power was distributed among the different transmit antennas in accordance with the probability of feedback errors. That is, only a fraction of power is given to the *best* antenna and this fraction depends on the feedback error probability. Other authors considered that an estimated version of the matrix $\mathbf{H}$ is available at the receiver. According to the degree of accuracy on the channel estimate, the precoding matrix was optimized with the purpose of minimizing the BER. In [Jön02], for instance, the optimization was carried out following a Bayesian approach, i.e., the system was designed to become *statistically* robust to channel uncertainties. In order to optimize the worst case, a maximin approach [Boy04] was encompassed in [PI06]. In particular, the eigenvectors of the estimated channel were used as transmit beamformers and, according to the CSI uncertainty, the transmit power was distributed among the transmission modes. In the extreme cases all the power is focused on the mode associated to the highest eigenmode if perfect CSI is available; whereas uniform power allocation (i.e, OSTBC) is used in the situation where the CSI estimation is completely inaccurate.

Spatial Multiplexing

MIMO schemes can also be used to improve the transmission data rate by sending several streams in parallel. When CSI is perfectly known at the transmitter, the system can be adapted to each channel realization and, in order to reduce the complexity, it is common practice to

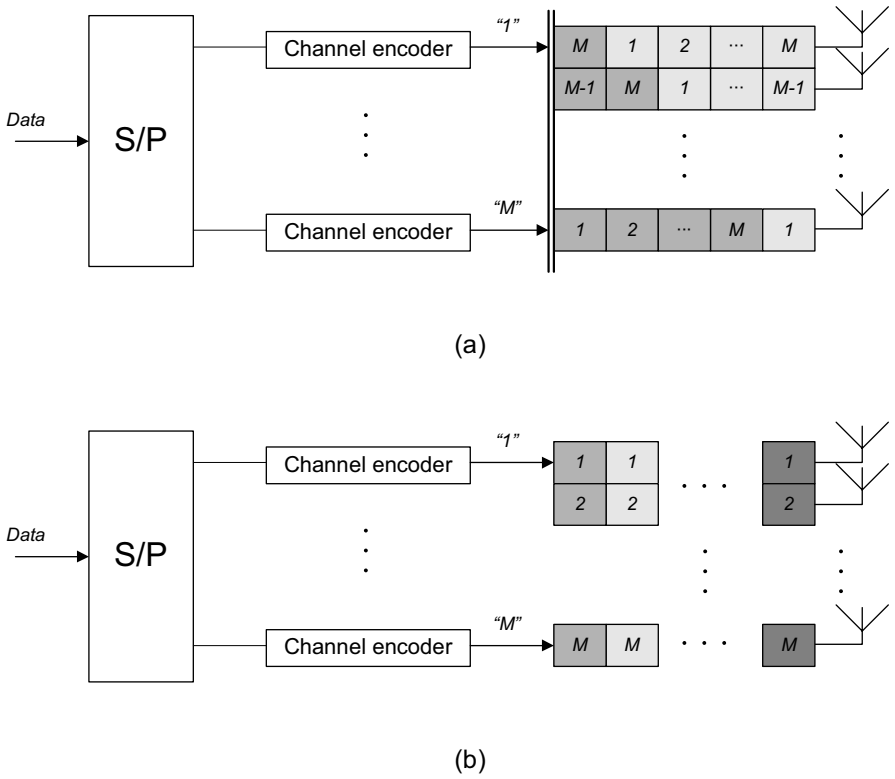


Figure 2.5: Block diagram of transmission scheme for BLAST architectures: (a) D-BLAST (b) V-BLAST.

adopt linear transceivers. Concerning the optimization procedure, some criteria can be found in the literature. As a case in point, in [Yan94] the optimization criterion is based on the minimization of the sum of the Mean Square Error (MSE) of all the streams and in [Ong03b] the average BER was considered as cost function. In order to obtain a generalized view of the design of MIMO linear transceivers, a unified framework was proposed in [Pal03]. It was shown that most of the objective functions used in communications can be grouped in two families: Schur-concave and Schur-convex functions. For the case of Schur-concave functions, the optimal solution consists in diagonalizing the channel, whereas in the Schur-convex case the channel must be rotated before diagonalization. Once the structure of the transceivers was imposed, the optimization problem was shown to be easily solved with the help of convex optimization tools. Some objective functions belonging to the Schur-concave family are those involved with the minimization of the sum of the MSEs, the maximization of the mutual information or the minimization of the product of the BERs. Concerning the Schur-convex, some examples are the minimization of the maximum of the MSE, the minimization of the average BER or the minimization of the maximum of the BER.

As noted previously, scenarios with perfect CSI at the transmitter are rarely found in wireless communications systems. A widely known set of architectures to exploit spatial multiplexing

Table 2.1: Detection algorithm proposed in [Gol99] for the V-BLAST scheme.

$$
\begin{aligned}
&\text{Initialization} \quad : \\
&\mathbf{r}_1 = \mathbf{r} \\
&\mathbf{G}_1 = \begin{cases} \mathbf{H}^{\#} & \text{for ZF receiver} \\ \sqrt{\frac{M}{\bar{\gamma}}}\left(\mathbf{H}^H\mathbf{H} + \frac{M}{\bar{\gamma}}\right)^{-1}\mathbf{H}^H & \text{for MMSE receiver} \end{cases} \\
&\text{Recursion} \quad : \quad for\ k = 1, ..., M\ do \\
&l_k = \arg\max_j \mathsf{SINR}_j = \arg\min_j \left\| (\mathbf{G}_k)_j \right\|^2 \qquad (2.1) \\
&\mathbf{w}_k = \left(\mathbf{G}_k^T\right)_{l_k} \qquad (2.2) \\
&y_{l_k} = \mathbf{w}_k^T \mathbf{r}_k \qquad (2.3) \\
&\hat{s}_{l_k} = Q\left(y_{l_k}\right) \qquad (2.4) \\
&\mathbf{r}_{k+1} = \mathbf{r}_k - \hat{s}_{l_k}(\mathbf{H})_{l_k} \qquad (2.5) \\
&\mathbf{G}_{k+1} = \begin{cases} \mathbf{H}_{l_k}^{\#} & \text{for ZF receiver} \\ \sqrt{\frac{M}{\bar{\gamma}}}\left(\mathbf{H}_{l_k}^H\mathbf{H}_{\overline{l_k}} + \frac{M}{\bar{\gamma}}\right)^{-1}\mathbf{H}_{l_k}^H & \text{for MMSE receiver} \end{cases}
\end{aligned}
$$

gains in scenarios without CSI are the so-called Bell-Labs LAyered Space-Time (BLAST) architectures [Fos96, Gol99]. These architectures are based on splitting the original data stream into several data streams, each of them being independently encoded. At the receiver, interference cancelation is performed with the aim of converting the MIMO channel into a set of parallel channels. According to how the encoding streams are distributed among the transmit antennas, different architectures result. The Diagonal-BLAST (D-BLAST) architecture consists in sending the symbols of each encoding stream over all the transmit antennas in a sequential order (see Fig. 2.5.(a)). A simplified architecture called Vertical-BLAST (V-BLAST) was proposed in [Gol99], where the encoding streams are sent in parallel as can be observed in Fig. 2.5.(b). In terms of system performance, D-BLAST obtains better results because the symbols are transmitted over all the transmit antennas and, as a result, spatial diversity is higher. However, since this scheme is not easy implementable, the V-BLAST strategy is preferable for practical applications [Fos99]. Furthermore, it was pointed out in [Chu04] that if the transmitter is able to adapt the rates of the symbols conveyed in the different transmit antennas on the basis of the information provided with a feedback channel, the capacity of the MIMO channel in an i.i.d Rayleigh fading scenario can be achieved with the V-BLAST scheme.

Finally, it is noteworthy that performance of V-BLAST can be quite good due to the detection algorithm proposed by Golden et. al in [Gol99]. Basically, this algorithm is based on the classical successive cancelation receiver [Pau03] but, in each iteration, the strongest symbol is detected. As a result, the detection procedure can be substantivally improved since a means of selection diversity is exploited. The complete description of the algorithm can be found in Table 2.1, where $(\mathbf{G}^T)_{l_k}$ stands for the l_k-th column of matrix $\mathbf{G}^T$, $\mathbf{H}_{\overline{l_k}}$ is the matrix constructed by dropping the

columns $l_1, l_2, ..., l_k$ of matrix $\mathbf{H}$ and it is assumed that transmit power is uniformly distributed (i.e., $\mathbb{E}[|s_k|^2] = \frac{\bar{\gamma}}{M}$). At the k-th iteration of the algorithm, the stream with the highest Signal to Interference plus Noise Ratio (SINR) out of those that have not been detected yet is selected (Table 2.1, Eq. (2.1)), symbols are filtered out with the corresponding Zero Forcing (ZF) or Minimum Mean Square Error (MMSE) spatial filter[3], $\mathbf{w}_k$, (Eq. (2.2)) and, then, a sliced version of the *soft* output is produced (Eq. (2.4)). Contribution from symbols in the current stream to the overall received signal is removed by re-generating transmitted signals with the symbol estimates (Eq. (2.5)). By doing so, constraints for the spatial filters in subsequent streams can be relaxed since the number of actual interfering signals to be canceled out becomes lower.

2.3.2 Antenna Selection

In a MIMO system, adding complete Radio Frequency (RF) chains may result in increased complexity, size and cost. These negative effects can be drastically reduced by using antenna selection. This is because antenna elements and digital signal processing is considerably cheaper than introducing complete RF chains. In addition, many of the benefits of MIMO schemes can still be obtained [Mol04, San04a]. Besides, perfect CSI is not required at the transmitter as the antenna selection command can be computed at the receiver and reported to the transmitter by means of a low-rate feedback channel.

In Fig. 2.6, we show a typical MIMO wireless system with antenna selection capabilities at both the transmit and the receive sides. The system is equipped with M transmit and N receive antennas, whereas a lower number of RF chains has been considered ($L_M < M$ and $L_N < N$ at the transmitter and receiver, respectively). In accordance with the selection criterion, the best sub-set of L_M transmit and L_N receive antennas are selected. In order to convey the antenna selection command to the transmitter, a feedback channel is needed but this can be done with a low-rate feedback as only $\binom{M}{L_M}$ bits are required.

Originally, antenna selection algorithms were born with the purpose of improving link *reliability* by exploiting spatial diversity. More precisely, a reduced complexity system with antenna selection can achieve the same diversity order as the system with all antennas in use. However, as MIMO schemes gained popularity, antenna selection algorithms began to be adopted in *spatial multiplexing* schemes aimed at increasing the system capacity.

A brief review of the state of the art is presented below, where different methodologies are classified according to the context: spatial diversity or spatial multiplexing.

[3]The optimality of the V-BLAST scheme discussed in [Chu04] is only valid for the MMSE receiver.

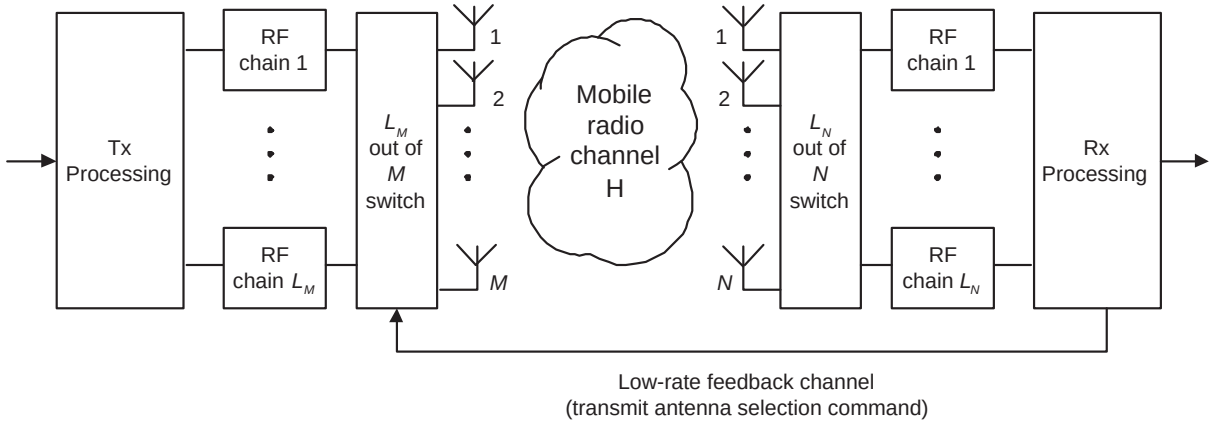


Figure 2.6: Block diagram of a MIMO scheme with transmit and receive antenna selection.

Antenna Selection for Spatial Diversity

Antenna selection was introduced by Jakes as a simple and low-cost solution capable of exploiting receive diversity in a Single-Input Multiple-Output (SIMO) scheme [Jak74]. In a wireless environment, by separating the receive antennas far enough[4] the correlation between the channel fades is low. Then, by selecting the best receive antenna in terms of channel gains, a diversity order equal to the number of receive antennas is obtained. Winters considered a similar procedure in a Multiple-Input Single-Output (MISO) system to exploit diversity at the transmit side with the help of a feedback channel [Win83]. In that work, the antenna selection algorithm was very simple: when the received SNR was below an specific threshold a command is sent to the transmitter to indicate that the transmit antenna must be switched.

For the SIMO case, more sophisticated receive antenna selection algorithms based on Hybrid Selection/maximal-ratio combining techniques were derived in [Win99a, Win99b, Win01]. The basic idea of those algorithms was to select the best (in terms of SNR) L_N out of N receive antennas and combine the received signals by means of a MRC procedure. By doing so, apart from exploiting the diversity gain, array gain can also be extracted. The extension to MIMO systems were presented by Molisch *et al.* [Mol01b, Mol03b] in a scenario where antenna selection was only performed at the transmitter in combination with a MRT strategy. It was shown that by selecting the best sub-set of transmit antennas, the degradation in system performance is slight in comparison with the saving in terms of hardware cost. The obtained results can be easily generalized to those cases performing antenna selection at the receive side of the MIMO link due to the reciprocity of the SNR maximization problem. An interesting result was obtained in [Che04] for those systems performing MRC at the receiver side and an antenna selection mechanism (with a single active antenna) at the transmitter. It was shown that the achieved

[4]For mobile terminals surrounded by other objects, quarter-wavelength spacing is sufficient, whereas for high base station a separation of 10-20 wavelengths is required [Win98].

diversity order is equal to NB, with B standing for the position taken by the channel gain of the selected antenna when arranging the channel gains of the different transmitters in an increasing order.

The combination of antenna selection with OSTBC was studied by Gore and Paulraj in [Gor02a]. It was proven that the diversity order obtained through antenna selection is identical to that of a situation with all the antennas in use. Regarding the degradation in terms of SNR when antenna selection is carried out at the receiver, it was shown in [Zen04] that it can be upper bounded by $10\log_{10}(N/L_N)$ dB. In a similar context, both transmit and receive antenna selection mechanisms in combination with Q-OSTBC schemes were analyzed in [Bad04]. For the case that antenna selection is combined with STTC, different results were found: by increasing the total number of receive antennas N, the coding gain can be improved but the diversity order remains fixed [San04b].

Antenna Selection for Spatial Multiplexing

In spatially correlated MIMO fading channels, capacity gains can be lower than expected since spatial multiplexing gains mainly come from resolving parallel paths in rich scattering MIMO environments. With this problem in mind, Gore *et al.* proposed one of the first papers where antenna selection was adopted in a MIMO context [Gor00]. In that paper, the authors showed that system capacity cannot be improved by using a number of transmit antennas greater than the rank of the channel matrix. By considering that, an algorithm was proposed where only antennas satisfying the full rank condition were selected. As a result, system capacity gains were obtained with respect to the full antenna system, since transmit power was efficiently distributed. In order to reduce the complexity of the proposed algorithm (exhaustive search), various sub-optimal algorithms based on the waterfilling principle [Cov91] were proposed in [San00].

Upper bounds of the achievable capacity with antenna selection were derived in [Mol01a]. In particular, it was shown that capacity results close to those of the full antenna system can be achieved by selecting the best $L_N \geq M$ out of N receive antennas. In [Gor02c], a sub-optimal approach was proposed for both transmit and receive antenna selection. By starting with the full channel matrix, those rows (columns) corresponding to the receivers (transmitters) minimizing the capacity loss are iteratively dropped. As shown in [Gor03, GA04], almost the same capacity as with an optimal selection scheme can be achieved with an incremental version of the mentioned selection algorithm, i.e., by using a bottom-up selection procedure. In [Gor03] it was also proven that the diversity order achieved with receive antenna selection is the same as that with the full antenna scheme, where the diversity order was defined as the slope of the outage rate. Although a sub-optimal approach with decoupled transmit and receive selection was adopted in [Gor04],

similar conclusions in terms of the diversity-multiplexing trade-off curve [Zhe03] were drawn. That is, the same trade-off curve as with all antennas in use can be obtained with transmit and receive antenna selection.

Heath *et al.*, on the other hand, pointed out that antenna selection approaches based on maximizing the mutual information do not necessarily minimize the error rate when practical receivers are in use [Hea01]. As an alternative, minimum error rate algorithms were derived and analyzed in systems with ZF and MMSE linear receivers. As for the ZF approach, selection algorithms were also derived in [Gor02b] for the case that only channel statistics (covariance matrix) are known at the transmitter. A geometrical approach was presented in [Ber04] in order to reduce the computational complexity.

2.4 Multi-user MIMO Systems

So far we have discussed the benefits obtained with MIMO systems in a single-user scenario. In multi-user systems, however, multi-antenna capabilities play a different role. Apart from exploiting spatial diversity and multiplexing gains, multi-antenna capabilities can also be exploited to send (receive) information to (from) several users simultaneously.

As previously explained, the interests of this thesis are oriented towards the study of broadcast channels. For that reason, in this section we give some insight on the model considered in the downlink of a wireless multi-user system, i.e., the MIMO Gaussian Broadcast Channel. We start the discussion by exploring the fundamental limits and, after that, we present some practical schemes. In particular, we present a scheme known as orthogonal random beamforming which requires a low amount of information in the feedback channel and, due to its applicability in wireless networks, is studied in detail in Chapter 5.

2.4.1 MIMO Gaussian Broadcast Channel

Consider the downlink of a wireless system with one Base Station (BS) equipped with M antennas and K Mobile Stations (MS) with N_1,...,N_K receive antennas, respectively (see Fig. 2.7). The received signal for user k can be modeled as:

$$\mathbf{r}_k = \mathbf{H}_k\mathbf{s} + \mathbf{n}_k$$

where $\mathbf{H}_k \in \mathbb{C}^{N_k \times M}$ is the channel matrix gain between the BS and the k-th MS, $\mathbf{s} \in \mathbb{C}^{M\times 1}$ is the symbol vector broadcasted from the BS and $\mathbf{n}_k \in \mathbb{C}^{N_k\times 1}$ stands for an additive Gaussian noise vector of complex, random variables with zero mean and unit variance, $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_k})$.

For the case that $N_k = M = 1$, this model falls into the category of degraded broadcast channel, which capacity region is widely known for the case with perfect CSI at the transmitter

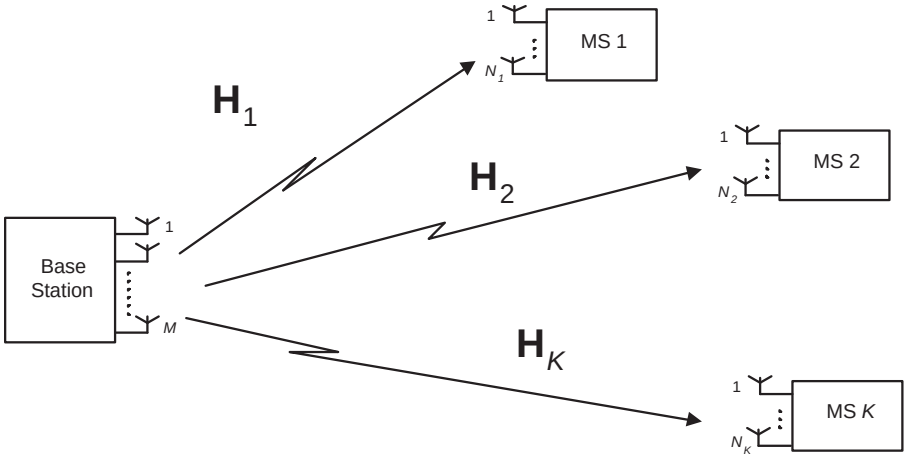


Figure 2.7: Block diagram of a MIMO broadcast wireless system.

[Li01]. As explained in Subsection 2.2.2, the average cell throughput can be increased when in each time-slot the user with the best channel conditions is scheduled [Tse97, Li01]. However, in the multiple antenna case (i.e., in the MIMO Gaussian Broadcast Channel setup) the problem is more complicated since the channel is in general non-degraded. Roughly speaking, the ordering among users is lost. Indeed, for the general non-degraded broadcast channel the capacity region is still unknown. Fortunately, the capacity region of the MIMO Gaussian Broadcast Channel has been recently solved in [Wei04], where the authors proved that Dirty Paper Coding (DPC) is the capacity-achieving strategy. In order to obtain that result, some previous steps were done, which are briefly overviewed in the next paragraphs.

Costa [Cos83] showed that in a Gaussian scalar channel with Gaussian interference non-causally known at the transmitter, the same capacity can be obtained as that achieved without interference. To do so, Costa proposed a technique called *writing on dirty paper* (or *dirty paper coding*) consisting in the exploitation of the interference knowledge at the transmitter when the transmitted signal is constructed. In particular, an interference pre-substraction technique was used to completely remove the interference component. This strategy was considered by Caire and Shamai [Cai03] in order to obtain an achievable region for the MIMO Gaussian Broadcast Channel. In particular, the authors obtained a closed-form expression for a scenario with two transmit antennas and two single-antenna receivers. Besides, they proved that the maximum sum-rate achievable with the DPC strategy is in fact the sum-rate capacity of the MIMO Gaussian Broadcast Channel. The extension of that work to an arbitrary number of transmit antennas was derived in [Vis03c] and the generalization to a scenario with multiple antennas at both the transmit and the receivers was carried out in [Vis03a]. In those papers, the problem was considerably simplified due to the use of the duality principle between the uplink and downlink channels. This is because the concept of duality, first proposed by Jindal *et al.* [Jin02, Jin04] for the scalar Gaussian Broadcast Channel, facilitates the computation of

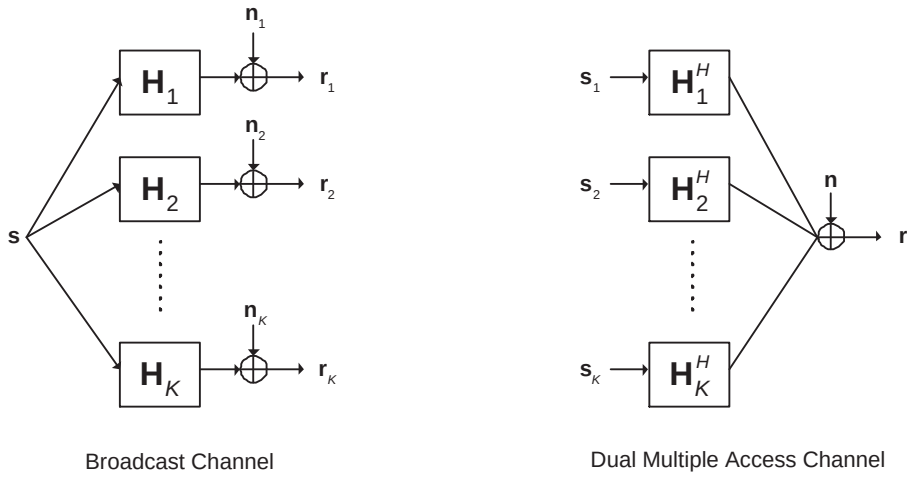


Figure 2.8: Block diagram of the Broadcast (left) and Multiple Access Channels (right).

the DPC achievable region thanks to the following Theorem: "*the achievable region of DPC with power constraint P is equal to the capacity region of the dual MIMO Multiple Access Channel with sum power constraint P*". That is,

$$C_{DPC}(P, \mathbf{H}_1, ..., \mathbf{H}_K) = \bigcup_{\mathbf{P}_k : Tr\left(\sum_{k=1}^{K} \mathbf{P}_k\right) \leq P} C_{MAC}(\mathbf{P}_1, ..., \mathbf{P}_K, \mathbf{H}_1^H, ..., \mathbf{H}_K^H)$$

where $\mathbf{P}_1, ..., \mathbf{P}_K$ stand for the covariance matrices of the transmitted signals in the dual Multiple Access Channel, $\mathbf{P}_k = \mathbb{E}\left[\mathbf{s}_k \mathbf{s}_k^H\right]$ (see Fig. 2.8). Essentially, the computation of the achievable region of DPC is simplified because it is expressed in terms of the widely known capacity region of the MIMO Multiple Access Channel [Cov91, Pau03]. By exploiting the duality concept, the expression of the sum-rate capacity can also be easily written as follows:

$$R_{DPC}(P, \mathbf{H}_1, ..., \mathbf{H}_K) = \max_{\mathbf{P}_k : Tr\left(\sum_{k=1}^{K} \mathbf{P}_k\right) \leq P} \log_2 \det \left(\mathbf{I}_M + \sum_{k=1}^{K} \mathbf{H}_k^H \mathbf{P}_k \mathbf{H}_k \right)$$

In the case that the channels of the different users are independent, the above expression at high SNR linearly grows as a function of $\min(M, K)$ [Vis02]. For a high number of users and i.i.d Rayleigh fading, the DPC strategy is capable of exploiting both the spatial multiplexing and MUD gains. More precisely, the ergodic sum capacity grows linearly with the number of transmit antennas and double logarithmically with the number of users [Sha04], that is:

$$\lim_{K \to \infty} \mathbb{E}_{\mathbf{H}_1, ..., \mathbf{H}_K} \left[R_{DPC}(P, \mathbf{H}_1, ..., \mathbf{H}_K) \right] = M \log_2 \log_2 (KN)$$

where it is assumed that all the users have the same number of receive antennas N.

At that point, information theory community knew how to compute the achievable region of the DPC, and that the sum-capacity of the MIMO Gaussian broadcast channel could be achieved

with DPC. Nonetheless, it was not proved yet if DPC was the capacity achieving strategy. The final step was done thanks to a conjecture of Viswanath *et al.* [Vis03b] and Tse and Viswanath [Tse03]. By introducing the concept of enhanced broadcast channels, the conjecture was finally confirmed by Weingarten *et al.* [Wei04]. Enhanced broadcast channels are defined as improved versions of the original broadcast channel, whose capacity region contains that of the original one. Then, if one point lies outside the capacity region of the enhanced channel, this point must lie outside the capacity region of the original channel as well. Since it was shown that all the point outside the DPC achievable region were also outside the capacity region of the enhanced channel, it was concluded that DPC is the capacity achieving strategy for the MIMO Gaussian Broadcast Channel.

2.4.2 Practical Approaches

DPC has two main drawbacks: its computational and implementation complexity (due to successive encoding and decoding processes) and the need for perfect CSI at the transmitter. In order to alleviate that, several works concerning sub-optimal strategies have appeared in the literature during the last years. Due to their practical applicability, we restrict ourselves to the most popular schemes based on linear processing: optimal beamforming, ZF beamforming and orthogonal random beamforming. For an extensive overview of transmission strategies including linear and non-linear precoding, the interested reader is referred to [Cai06].

For the sake of simplicity, in this section we consider a scenario with single-antenna receivers. In order to serve multiple users in the same time-slot, a linear precoding matrix $\mathbf{W}$ is used at the base station. In this case, the received signal for user k can be modeled as:

$$r_k = \mathbf{h}_k^T \mathbf{W}\mathbf{s} + n_k \tag{2.6}$$

where $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ is the channel vector gain between the BS and the k-th MS and $n_k \in \mathbb{C}$ denotes Additive White Gaussian Noise (AWGN) with zero mean and unit variance. The above expression can be re-written as:

$$r_k = \mathbf{h}_k^T \mathbf{w}_k s_k + \sum_{j\neq k} \mathbf{h}_k^T \mathbf{w}_j s_j + n_k \tag{2.7}$$

where s_j stands for the symbol transmitted with beam j. Notice that the last two terms in the above expression correspond to the interference-plus-noise contribution and, hence, the corresponding long-term averaged SINR amounts to:

$$\mathsf{SINR}_k = \frac{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k}{\sum_{j\neq k} \mathbf{w}_j^H \mathbf{R}_k \mathbf{w}_j + 1}$$

where $\mathbf{R}_k$ is the channel covariance matrix, $\mathbf{R}_k = \mathbb{E}[\mathbf{h}_k \mathbf{h}_k^H]$, and it is assumed that $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_M$.

The *optimal beamforming* approach [Ben01] is based on the search of the optimal set of beamvectors $\mathbf{w}_k$ minimizing the transmit power and satisfying the QoS requirements of the system (expressed in terms of a minimum required SINR, γ_{th}), i.e.[5]:

$$\begin{aligned} &\min_{\mathbf{w}_k} \mathbf{w}_k^H \mathbf{w}_k \\ &\text{s.t. } \mathsf{SINR}_k \geq \gamma_{th,k} \qquad\qquad k = 1, 2, \ldots, K, \end{aligned}$$

The power component can be separated from the beamforming vectors themselves by defining $\mathbf{w}_k$ as the product of a unitary vector $\mathbf{u}_k$ and a scalar containing the power information ρ_k, i.e., $\mathbf{w}_k = \sqrt{\rho_k}\mathbf{u}_k$. Then, the problem can be reformulated as:

$$\begin{aligned} &\min_{\rho_k, \mathbf{u}_k : ||\mathbf{u}_k|| = 1} \rho_k \\ &\text{s.t. } \frac{\rho_k \mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k}{\sum_{j \neq k} \rho_j \mathbf{u}_j^H \mathbf{R}_k \mathbf{u}_j + 1} \geq \gamma_{th,k} \qquad\qquad k = 1, 2, \ldots, K, \end{aligned}$$

The above problem is non-linear and non-convex and, then, it is not a trivial task to find the direct solution. An iterative algorithm capable of finding the global optimum was proposed in [RF98]. In particular, it was shown that the problem is equivalent to a power minimization problem in a *virtual uplink channel* with the same channel response and the same SINR constraints. In each iteration of the algorithm, the optimal receive beamvector obtained in the uplink problem is used as beamvector for the downlink case. Then, powers in both the uplink and downlink problems are adapted in order to satisfy the SINR constraints. Indeed, this iterative algorithm was designed by exploiting the duality concept between the uplink and downlink channels discussed in the previous section, which was formalized for the case of linear beamforming in [Vis03c] and [Sch04]. The same optimization problem was solved from another perspective in [Ben99]. In that paper, the optimization problem was reformulated as a semidefinite programming problem [Boy04]. By doing so, standard tools for semidefinite optimization can be used to solve the problem, which provide a faster convergence. However, the computational complexity is rather high.

In order to reduce the computational cost, a simpler approach based on a transmit ZF (or channel inversion) strategy can be considered. Basically, the transmitter is designed to null the contributions of the interfering users, i.e., the transmit beamvectors are designed in such a way that:

$$\mathbf{h}_k^T \mathbf{w}_j = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases}$$

By computing the pseudo-inverse of the global channel response $\mathbf{H} = [\mathbf{h}_1, ..., \mathbf{h}_K]^T$:

$$\mathbf{W} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}$$

[5] It is worth noting that the optimal beamforming approach can also be formulated as a SINR maximization problem with power constraints [Sch04].

one can easily obtain the ZF beamvectors. When the channel *signatures* of two users are highly correlated, the required amount of power to orthogonalize them may be extremely high and here resides one of the main drawbacks of ZF: its power inefficiency. However, both the design and analysis of power and bit allocation techniques can be substantially simplified as the resulting channels are decoupled. For instance, different bit allocation algorithms were proposed in terms of system requirements [Bar05]. More precisely, the algorithms were designed in order to balance the existing trade-offs in terms of global system performance vs. fairness among users. Furthermore, ZF may be a reasonable alternative to optimal beamforming in some cases as pointed out in [Ben05]. In that paper, it was proved that the same system performance can be obtained in a single-cell and interference-limited scenario when performing practical bit allocation algorithms.

For the case when the number of users is higher than the number of transmit antennas, user selection algorithms must be carried out [Cai06]. In other words, the sub-set of users maximizing the instantaneous performance must be selected in each time-slot. Due to the combinatorial nature of the optimization process, an exhaustive search is mandatory to obtain the optimum solution. In [Bar05, Bay05, Yoo06a] sub-optimum procedures based on iteratively adding users nearly orthogonal to the previously selected ones were proposed. In addition, multi-user diversity gains can still be exploited in the sense that an identical growth-rate of the sum-rate capacity for an increasing number of active users can be obtained [Yoo06a]. In fact, for a large number of users the sum capacity expressions of DPC, optimal beamforming and ZF beamforming have the same growth-rate [Yoo06a, Sha04]:

$$\lim_{K\to\infty} \mathbb{E}\left[R_{ZF}\right] = \lim_{K\to\infty} \mathbb{E}\left[R_{BF}\right] = \lim_{K\to\infty} \mathbb{E}\left[R_{DPC}\right] = M \log_2\left(\log_2(K)\right)$$

Nonetheless, the ZF transmission technique requires perfect and complete CSI at the transmitter which in Frequency Division Duplexing (FDD) systems is difficult to obtain. In order to circumvent that, a modified version of the opportunistic beamforming scheme was proposed in [Vis02] and further analyzed in [Sha05]. The main idea of this scheme, herein referred as Orthogonal Random Beamforming (ORB), is to generate a set of random orthogonal beamformers in order to exploit the multiplexing capabilities of the system. Following the philosophy of opportunistic beamforming [Vis02], users only have to report SINR relative to the selected precoder and, then, the BS selects the best user for each beam. The rationale behind this algorithm is that in scenarios with a large enough number of users, there is a high probability of generating beams matching the channel signatures of a sub-set of users while maintaining low levels of interference.

By using ORB, the amount of information to be sent over the feedback channel is considerably reduced. More precisely, the information to be reported in the feedback channel are only the index and the SINR associated to the *best* beam. This is because the probability that one user is selected for transmission with two different beams is negligible when the number of users is

large in comparison with the number of active beams ($K >> M$)[Sha05]. Orthogonal random beamforming can be very effective in systems with a large number of users, since this scheme has the same capacity growth-rate as with perfect CSI at the BS (i.e., $M \log_2 \log_2(K)$) [Sha05]. For that reason, special attention has recently been paid to this scheme. However, the system performance of ORB is very dependent on the number of users with uncorrelated channels, being dramatically penalized in scenarios with a low number of users.

Chapter 3

A Cross-layer Approach to Transmit Antenna Selection

In this chapter, we begin our study by considering a single-user MIMO wireless system. In this scenario, we adopt a transmit antenna selection approach in order to exploit MIMO spatial multiplexing gains with a low-rate feedback channel. The selection criterion is based on a cross-layer approach in the sense that the transmit antenna sub-set maximizing data link layer performance is selected. The problem of switching between spatial diversity and multiplexing is addressed from a cross-layer point of view as well.

3.1 Introduction

In order to ensure high peak data rates, low latency and increased link throughput, new wireless communication systems often resort to advanced techniques. In the High-Speed Downlink Packet Access (HSDPA) technical specification of UMTS Terrestrial Radio Access (UTRA) [UTR01], for instance, sophisticated transmission schemes such as Adaptive Modulation and Coding (AMC), fast Hybrid Automatic Repeat-reQuest (H-ARQ) and Multiple-Input Multiple-Output (MIMO) are adopted.

In this context, where different functionalities at both the data link and physical layers are introduced in order to attain the highest possible throughput, some attention should be paid to Cross-Layer (CL) designs. In this chapter, we investigate several CL approaches to *transmit* antenna selection. We address a CL methodology in the sense that the *criterion* for the selection of antenna subsets is the maximization of data link layer throughput which, clearly, takes into account characteristics both at the physical *and* data link layers, such as receiver structure, packet size, modulation scheme and ARQ strategy.

As commented in Subsection 2.3.2 of the previous chapter, transmit Antenna Selection (AS)

approaches emerged as an effective means of exploiting MIMO capabilities in systems with low-rate feedback channels. As for the selection criteria, it is common practice to select the subset of transmit antennas maximizing mutual information. However, such PHY-based approaches, do not exploit all the information available in *specific* system scenarios concerning the *actual* schemes and algorithms in use at the physical and data link layers. Therefore, maximizing mutual information does not necessarily lead to an improved performance in terms of link throughput. For that reason, we adopt a *throughput*-based approach to AS that, as shown later, brings remarkable benefits in comparison with conventional PHY-based criteria.

Some pioneering examples of cross-layer antenna selection can be found in [Mil02a, Mil02b, Mil06]. Our contribution extends those initial works in order to encompass H-ARQ and AMC mechanisms. In other words, we obtain a version of the scheme tailored to H-ARQ needs that, in addition, is capable of jointly encompassing both transmit antenna selection *and* adaptive modulation. In contrast with other antenna selection algorithms existing in the literature, the number of parallel data streams may actually change from burst to burst. Then, we also analyze performance trade-offs with respect to those antenna selection algorithms where the number of streams remains constant but are dynamically mapped on different antennas (out of a larger pool set). In addition, reduced-complexity versions are derived for the proposed schemes and, also, a detailed complexity analysis is conducted. Finally, we also contemplate the problem of switching between spatial diversity and multiplexing from a cross-layer perspective. As we will show later, the switching problem is analogous to the AS one in the sense that transmission mode selection is carried out at the transmitter and, then, this problem can be solved in a similar way.

In summary, the contributions of this chapter are the following:

- Work carried out in [Mil02a, Mil02b, Mil06] is extended to encompass H-ARQ and AMC.
- The proposed CL antenna selection algorithm is adapted to H-ARQ needs.
- Existing trade-offs in terms of allowing a variable number of active antennas vs. selecting a pre-determined number of active antennas (out of a larger set) are assessed.
- A reduced-complexity (albeit optimal) version of the AS mechanism is derived.
- The problem of switching between spatial diversity and multiplexing is also solved from a cross-layer perspective.

The chapter is organized as follows: the system model under consideration is presented in Section 3.2. In Section 3.3, the throughput-based criterion is introduced, and the corresponding cross-layer AS scheme is further refined to obtain an adapted version to H-ARQ needs. The proposed AS mechanisms are compared with other strategies that can be found in the literature in Section 3.4 and further enhanced to incorporate adaptive modulation in Section 3.5. Reduced

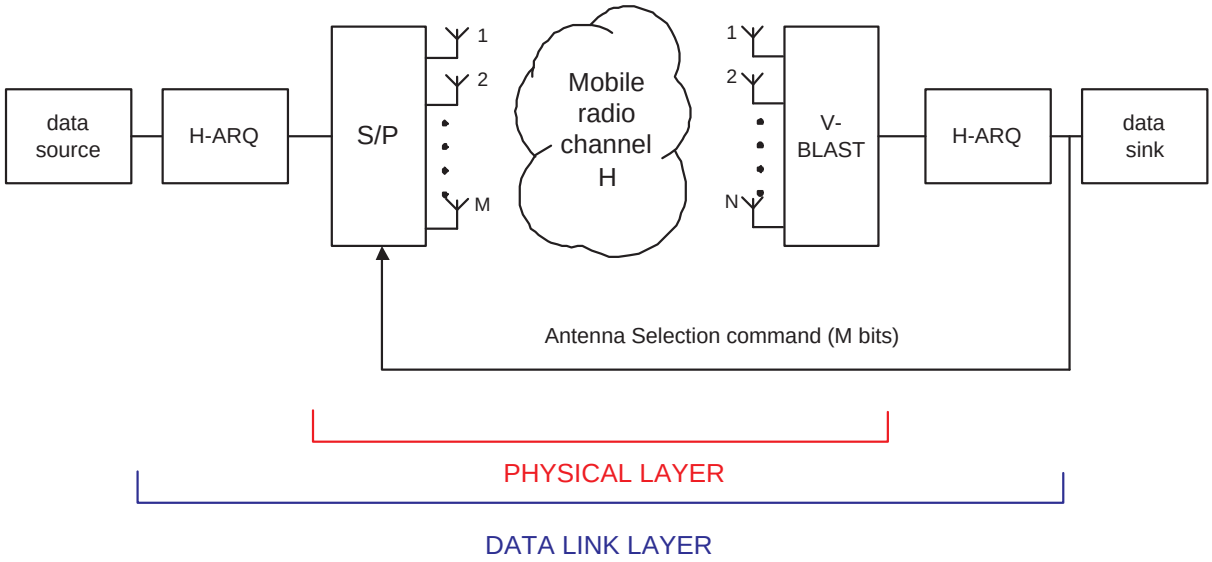


Figure 3.1: Block diagram of the communication system.

complexity versions of the algorithm are presented in Section 3.6. After that, the problem of switching between spatial and multiplexing is formulated from a cross-layer viewpoint in Section 3.7. Finally, the conclusions of this chapter are provided in Section 3.8.

The results obtained in this chapter have been published in [Vic03b, Vic03a, Vic04e, Vic04d, Vic04b, Vic04a, Vic04c, Vic04f, Vic05b, Vic06c].

3.2 System Description

Consider a transmission link between an M-antenna Base Station (BS) and an N-antenna Mobile Station (MS). In such a scenario, assume that the channel, which can be represented by an $N \times M$ matrix $\mathbf{H}$ (see Fig. 3.1), is flat-fading and Rayleigh-distributed. In order to exploit MIMO gains with a low-rate feedback channel, antenna selection is used at the transmitter. More precisely, the receiver estimates the *best* subset of transmit antennas[1], of dimension $m \leq M$, and feeds back this information to the transmitter with the help of the vector, $\mathbf{m}_m$ $(m_i \in \{0,1\}\,;\, i = 1, ..., M)$, where '1' and '0' values are used to indicate if the i-th transmit antenna must be activated or not, respectively. As a result, the received signal associated to the proposed scenario with AS capabilities can be modelled as:

$$\mathbf{r} = \mathbf{H}_m(\mathbf{m}_m)\mathbf{s} + \mathbf{n} \tag{3.1}$$

where $\mathbf{H}_m(\mathbf{m}_m) \in \mathbb{C}^{N \times m}$ is the matrix defined by the m columns of $\mathbf{H}$ corresponding to the subset of active antennas given by vector $\mathbf{m}_m$, $\mathbf{s} \in \mathbb{C}^{m \times 1}$ denotes the transmitted symbols vector drawn out of the selected constellation (QPSK or 16-QAM), and $\mathbf{n} \in \mathbb{C}^{N \times 1}$ stands for an

[1] Antenna selection criteria will be discussed in the next Section.

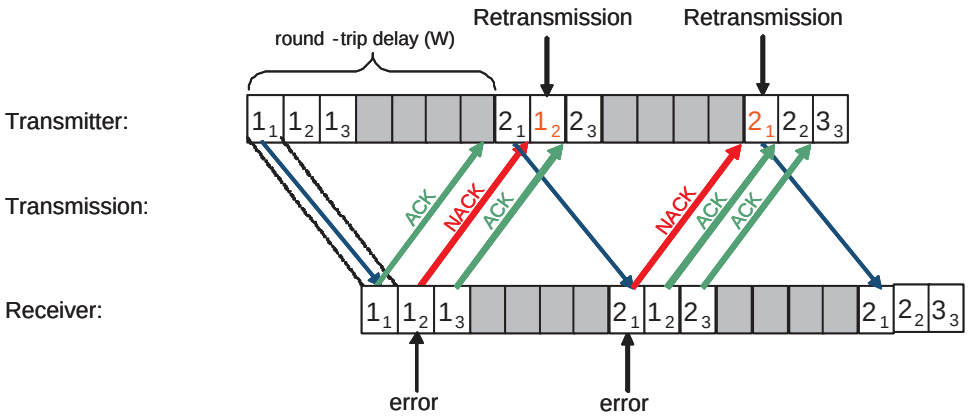


Figure 3.2: N-SAW strategy with three SAW processes running in parallel.

additive Gaussian noise vector of complex, random variables with zero mean and variance σ^2. For the ease of notation, we will denote $\mathbf{H}_m(\mathbf{m}_m)$ as $\mathbf{H}_m$, that is, we will not explicitly state the dependence of the channel matrix on vector $\mathbf{m}_m$. Since only scenarios with low-mobility users are considered, we will assume the channel responses to be stationary within each burst. Last, it is also assumed that perfect CSI is available at the receive side. At the transmit side, power is evenly distributed among active antennas and m streams are independently transmitted by means of a V-BLAST strategy. At the receiver, the detection procedure proposed by [Gol99] for the V-BLAST scheme is used for detecting the transmitted symbols, where spatial interference is removed on the basis of a ZF criterion (see Section 2.3 in the previous chapter for further details).

In alignment with the HSDPA technical specifications, we consider a Type III H-ARQ scheme at the data link layer. More precisely, we adopt an N Stop-and-Wait (N-SAW) re-transmission protocol along with a chase combining strategy [UTR01] for re-transmitted packets. The Stop-and-Wait ARQ protocol has been considered in the HSDPA standardization due to it minimizes signalling and buffering requirements at the MS [Lin84], which is very appropriate for wireless networks. Actually, a number of instances of the protocol are run in parallel so as to compensate for the idle time associated to round-trip delays (see Fig. 3.2) and, accordingly, the resulting scheme is referred to as the N-SAW protocol [UTR01]. As for the chase combining strategy, packets are combined by averaging soft symbols (i.e. equal gain combining) at the output of the V-BLAST scheme [Ong03a]. Therefore, the resulting symbol estimates after $p-1$ consecutive retransmissions can be expressed as:

$$\mathbf{y}_p = \frac{1}{p}\sum_{i=1}^{p}\mathbf{y}_{vb,i}$$

where $\mathbf{y}_{vb,i}$ denotes the *soft-symbol* vector at the output of the V-BLAST scheme at the i-th transmission. Clearly, a more efficient transmission can be achieved by exploiting Incremental Redundancy (IR) approaches for packet combination (where only additional coding information

is provided in consecutive re-transmissions) but this poses additional constraints in terms of memory, signalling requirements and computational complexity that were deemed inappropriate for practical implementation [Fre01].

3.3 Cross-layer Transmit Antenna Selection

In this section, we introduce a cross-layer method for transmit antenna selection on the basis of a throughput-based criterion. Before comparing such criterion to PHY-based ones, the AS method is refined in order to adapt the proposed methodology to the H-ARQ protocol presented above.

3.3.1 Capacity- vs. Throughput-based Criteria

A widely-used [San04a, Mol04] antenna selection criterion in MIMO systems exploiting spatial multiplexing gains consists in selecting the transmit antenna subset maximizing mutual information:

$$C_m = \log_2 \det\left(\mathbf{I}_N + \frac{\bar{\gamma}}{m}\mathbf{H}_m\mathbf{H}_m^H\right)$$

where $\bar{\gamma}$ is the *average* SNR per receive antenna. In this criterion, only physical parameters (i.e. channel response and SNR) are actually considered and, thus, the resulting scheme will be referred to as *PHY-layer only* approach.

In order to maximize the link throughput and, ultimately, enhance performance at the data link layer, additional information such as receiver structure, packet size, modulation scheme or H-ARQ strategy, etc. could be exploited as well, this leading to a *cross-layer* approach. As a previous step, an expression for data link layer throughput will be obtained in subsequent paragraphs.

For symbols conveyed in the *first* packet transmission (recall the H-ARQ strategy) which are detected in the k-th iteration of the recursive detection algorithm, we can define the post-filtering SNR as:

$$\gamma_{k,1} = \frac{\mathbb{E}\left[|s_k|^2\right]}{\sigma^2 \left\|\mathbf{w}_{k,1}\right\|^2} \qquad k = 1, 2, \ldots, m,$$

being s_k the transmitted symbol, and $\mathbf{w}_{k,1}$ the zero-forcing spatial filter used to obtain an estimate of the transmitted symbol. When considering a chase combining strategy the effective SNR after p transmissions (i.e. $p-1$ re-transmissions) can be expressed as:

$$\gamma_{k,p} = \alpha^{p-1} p \gamma_{k,1}$$

where α is the combining efficiency factor, that models the combining gain loss with respect to the theoretical model [Fre02]. Notice that expression $\alpha^{p-1}p$ is concave on the number of

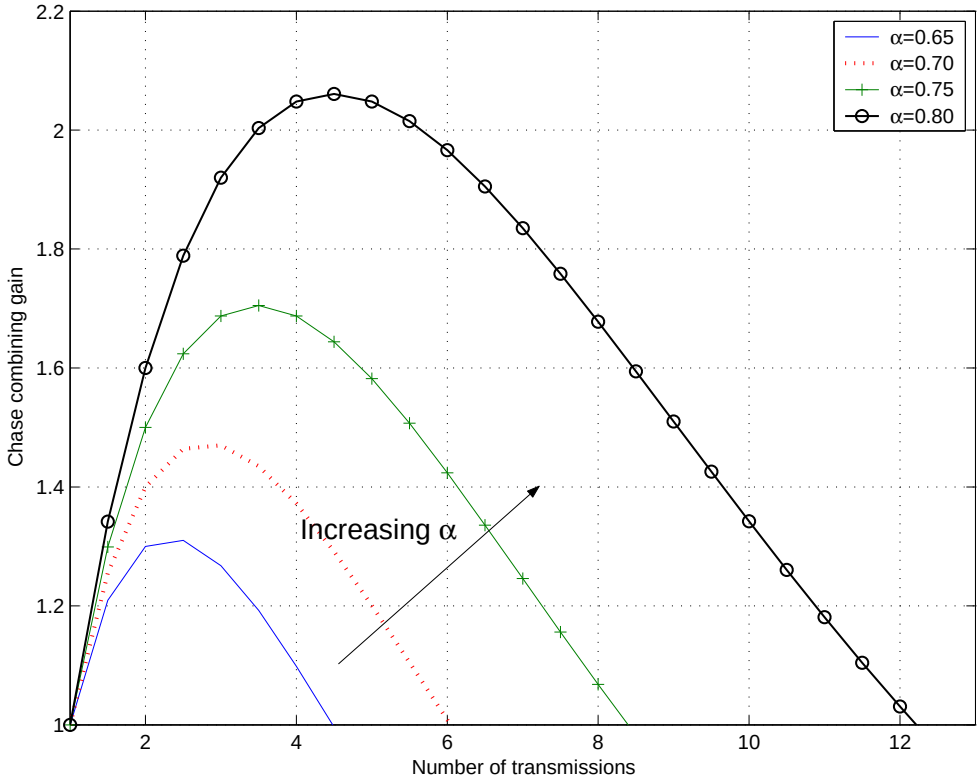


Figure 3.3: Chase combining gain as a function of the number of retransmissions.

packet transmissions, when relaxing p to be a continuous variable (see Fig.3.3). Therefore, there is no point in increasing the number of retransmissions when the optimum number of packet transmissions, P, is reached.

By disregarding error propagation in previous layers of the detection procedure and taking into account the specific modulation scheme in use, R, a closed-form expression for the Symbol Error Rate (SER) at each layer can be obtained, $\mathrm{SER}_{k,p} = f(\gamma_{k,p}, R)$ (e.g., for a QPSK case, $\mathrm{SER}_{k,p} = 2Q\left(\sqrt{\gamma_{k,p}}\right) - Q^2\left(\sqrt{\gamma_{k,p}}\right)$[Pro01]). From that, we can derive an expression of the uncoded Packet Error Rate (PER)[2]. Since as many L-symbol packets as actual antennas are sent in parallel (see Fig. 3.4), the resulting PER expression for the $L{\cdot}m$-symbol packet being transmitted from the subset of antennas defined by vector $\mathbf{m}$, is given by:

$$\mathrm{PER}_p(\mathbf{m}) = 1 - \left[\prod_{k=1}^{m}\left(1 - \mathrm{SER}_{k,p}\right)\right]^L \tag{3.2}$$

where we have assumed independent errors to occur.

The link throughput is measured as the effective number of correctly received bits at the data

[2]For a coded system, a different PER expression would result. In the case of convolutional codes, for instance, one could resort to the accurate and simple approximations in [Liu04]. Since, this chapter is aimed at illustrating a method to encompass both physical and data link layers parameters in the cost function and for the sake of simplicity, we will restrict ourselves to uncoded systems.

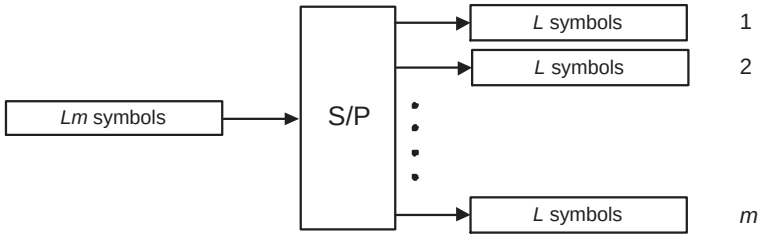


Figure 3.4: Structure of the transmitted packets.

link layer per channel use. Therefore, packet retransmissions due to the H-ARQ strategy must be taken into consideration. For the SAW protocol, the throughput expression can be written as [Lin84]:

$$\eta_{SAW} = \frac{l \cdot b}{W \cdot \mathbb{E}[p]}$$

where W stands for the round-trip delay expressed in slots, $\mathbb{E}[p]$ is the average number of packet transmissions, l is the ratio of information symbols per packet and b the number of bits per symbol, according to the modulation scheme in use. Then, by obtaining an estimate of the mean number of packet retransmissions, considering that m packets are sent in parallel and the fact that N_{SAW} Stop-and-Wait processes are run, an expression for the link layer throughput can be found:

$$\eta\left(\mathbf{m}\right) = \frac{N_{SAW} \cdot l \cdot m \cdot b}{W \cdot \mathbb{E}[p]} = \frac{l \cdot m \cdot b \cdot N_{SAW}/W}{\left(1 - PER_1\left(\mathbf{m}\right)\right) + \sum_{p=2}^{P}\left[p(1 - PER_p\left(\mathbf{m}\right)\right)\prod_{t=1}^{p-1} PER_t\left(\mathbf{m}\right)\Bigg]} \tag{3.3}$$

The above expression is used by the receiver in order to determine the optimal subset of transmit antennas, that is, the one maximizing link throughput. After that, the M-dimensional vector with the active antenna subset is signalled to the transmitter over the feedback channel. Given the highly non-linear nature of both the throughput expression and the maximization problem itself, a straightforward approach consists in checking every single combination of transmit antennas. In order to keep computational complexity affordable, though, it is possible to derive a reduced-complexity version of the algorithm (see Section 3.6 ahead).

To conclude, notice that the instantaneous throughput in Eq. (3.3) has been derived by considering a time invariant channel. However, this approximation is in general accurate. For a low number of retransmissions, the channel response will not change very much with respect to the original one. Besides, the number of packet retransmissions is limited by the system, $p \leq P$, thus, the number of summation terms in the denominator expression is potentially low. In order to validate this expression, performance results (corresponding to a system *without* antenna selection) are presented in Fig. 3.5. It can be observed a close matching between theoretical and empirical curves for the ARQ scheme (without chase combining), whereas a slight misadjustment is experienced in the H-ARQ scenario due to the non-ideal packet combining effect.

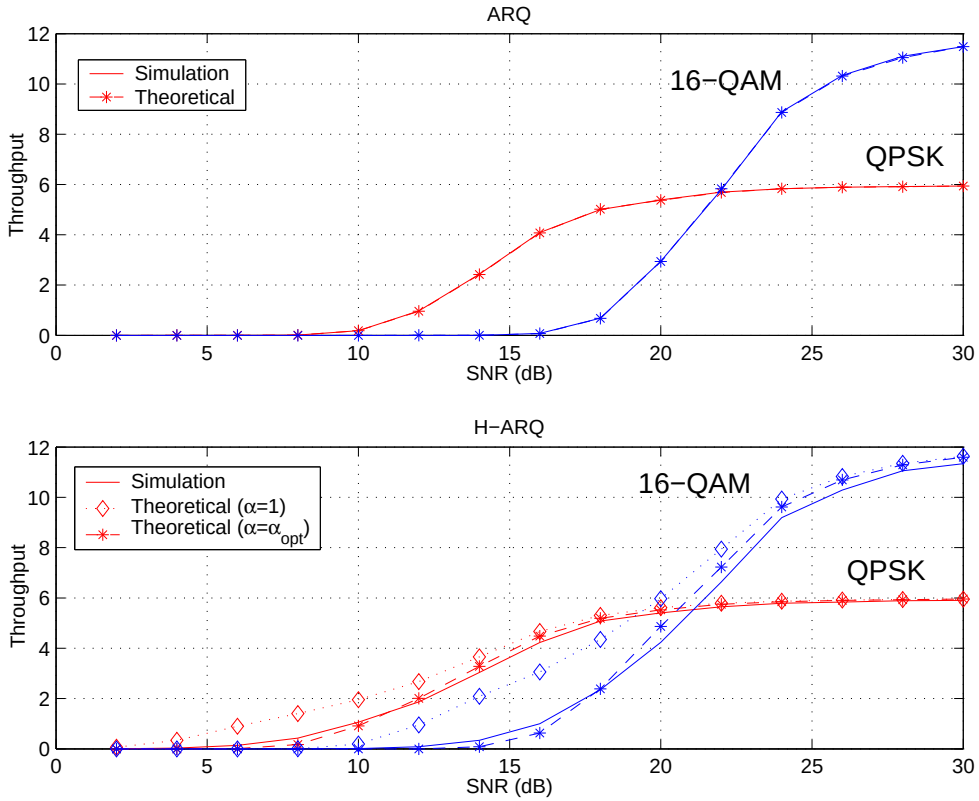


Figure 3.5: Validation of the theoretical throughput expression ($N = M = 3$, $W = 3$, $N_{SAW} = 3$, no antenna selection).

It is noteworthy that the optimum value of α has been empirically determined (by means of extensive Montecarlo simulations and least-squares fitting) and found to be equal to 0.74 and 0.72 for QPSK and 16-QAM, respectively.

3.3.2 Adaptation of Antenna Selection to H-ARQ

The proposed AS criterion is based on the adaptation of the transmit data rate by modifying the number of transmit antennas, i.e., different information symbols are transmitted over different transmit antennas. For that reason, some considerations must be taken into account to adapt the proposed methodology to the H-ARQ protocol in use and the existence of low complexity MS terminals.

First, the actual subset of transmit antennas must remain unchanged during packet retransmissions. This is needed in order to appropriately add-up packets with the chase combining approach. This extent, though, has little impact on performance in a context of slow moving terminals where the same antenna subset is valid over several consecutive frames. Moreover, for scenarios with high SNR the *all-active* antenna combination dominates and, thus, coherence periods can be even longer. This effect is shown in Fig. 3.6, where the validity period for the

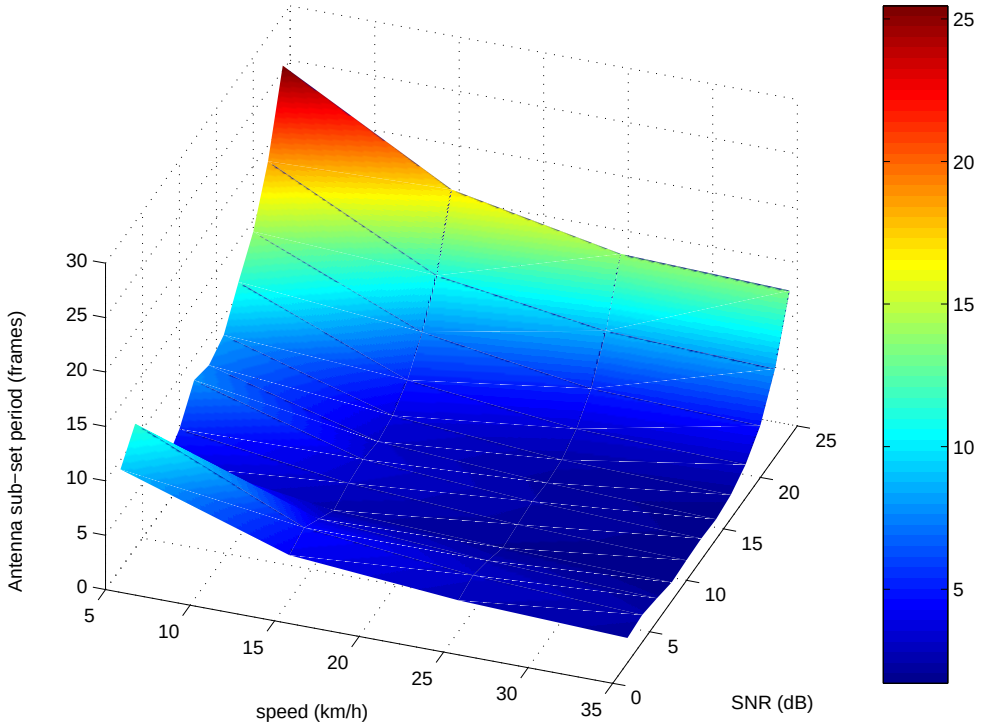


Figure 3.6: Validity period for the selected antenna sub-set (H-ARQ, $W = 3$, $N_{SAW} = 3$).

selected antenna sub-set, expressed in number of HSDPA frames, is presented as a function of SNR and MS's speed.

Second, packets being transmitted in parallel from different antennas must be accepted/discarded as a block. Since no numbering strategy is adopted by the N-SAW protocol, it is not possible to retransmit part of the m-packet block. Block discarding caused by, possibly, a single corrupted packet, has no critical impact on system performance provided that a throughput maximization strategy is adopted. In other words, the throughput criterion takes care of adaptively selecting the most appropriate packet size and, for that reason, the impact of discarding all the packets as a block is reduced.

And, finally, since multiple SAW processes are run in parallel, the same number of antenna selection processes can be conducted independently.

3.3.3 Computer Simulation Results

As far as computer simulations are concerned, we will consider a scenario of low mobility terminals ($v_{MS} = 3$ km/h) where 2 ms frames are divided into three slots, each featuring a $L = 160$-symbol packet (adopted from HSDPA). We will assume a round-trip delay of $W = 3$ slots and, accordingly, we will run $N_{SAW} = W = 3$ concurrent SAW processes. The FeedBack channel Information (FBI) field of HSDPA will be used to convey AS commands to the transmit side

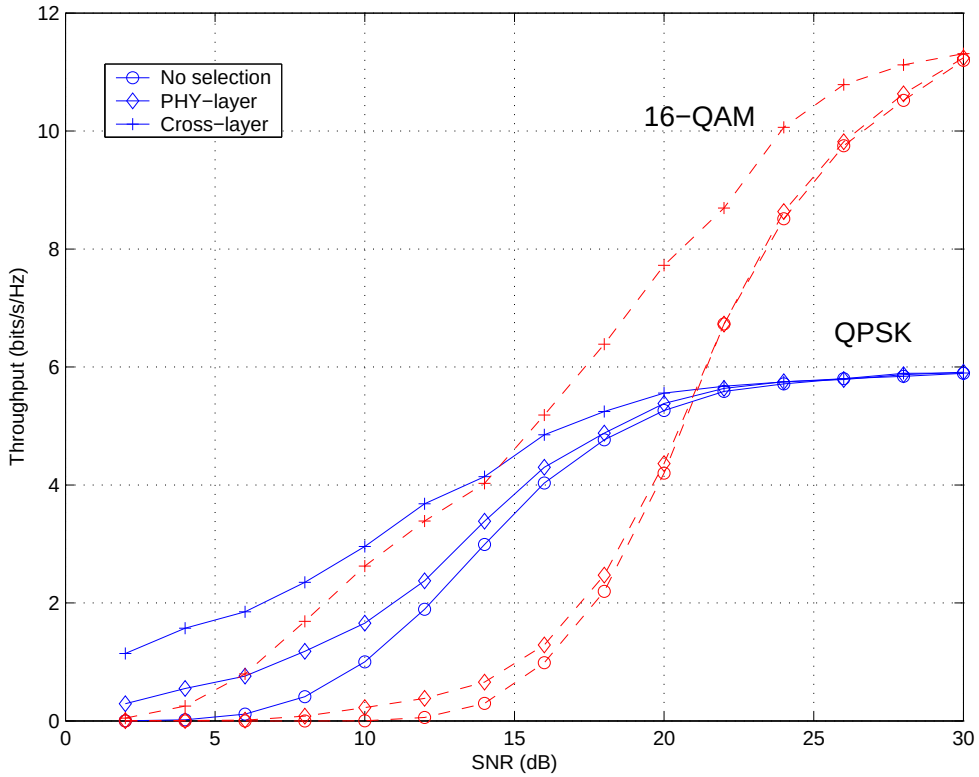


Figure 3.7: Throughput vs. SNR for the throughput-based (cross-layer) and PHY-based antenna selection mechanism in a 3 x 3 MIMO configuration ($W = 3$, $N_{SAW} = 3$).

over an error- and delay-free feedback channel. Since only one feedback bit per uplink control *slot* is available, AS commands are limited to three bits per uplink *frame*. Therefore, in order to issue commands on a frame-by-frame basis, the maximum number of transmit antennas is limited to $M = 3$. As for the number of receive antennas, it is set to $N = 3$ as well. Concerning the maximum number of packet transmissions is accordingly set equal to $P = 4$ so as to maximize the chase combining gain[3].

In Figure 3.7, the performance in terms of data link layer throughput vs. SNR for the AS methodologies based on maximizing mutual information (PHY) and throughput (CL) measures is compared. For benchmarking, an additional curve for a system without antenna selection is included, as well. It becomes apparent that the throughput-based criterion significantly outperforms its PHY-based counterpart for the whole range of SNR and for both modulation schemes. Conversely, mutual information for the antenna subsets selected with the PHY-based criterion exceeds in 2 bits/s/Hz (in average) that of subsets maximizing link throughput (see Fig. 3.8). Unfortunately, this does not revert in an increased spectral efficiency since channel bandwidth is simply wasted in packet retransmissions. Going back to Fig. 3.7, the larger the constellation size (16-QAM) the wider the gap between curves for solutions maximizing mutual information and

[3]The chase combining reliability parameter α has been previously set to 0.74 and 0.72 for QPSK and 16-QAM, respectively. Then, one can easily determine from Fig. 3.3 the most appropriate value for P.

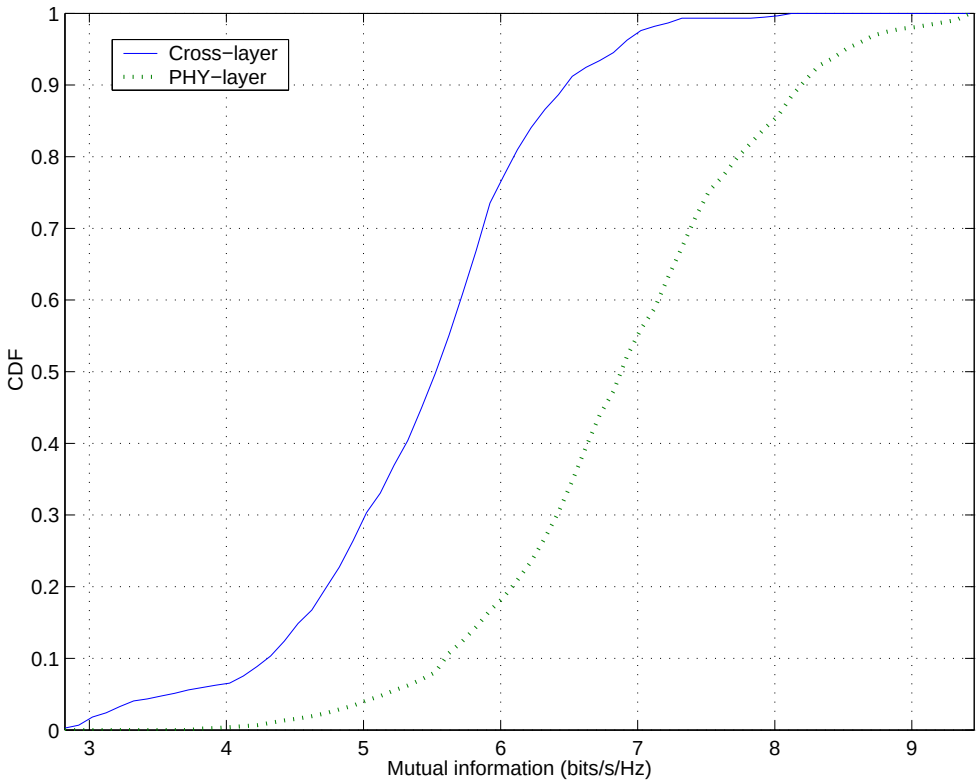


Figure 3.8: CDF for mutual information for the different antenna selection criteria (SNR=8dB).

throughput. Clearly, the former one does not *explicitly* consider the actual modulation scheme whereas the latter takes that into account (along with packet size and other transmission parameters). As a result, the throughput-based criterion is capable of further concentrating transmit power on the *best* channels when using 16-QAM. By doing so, the transmit power is efficiently exploited and the detection procedure can be improved because error propagation between layers is reduced. Indeed, little or no improvement is achieved by the PHY-based solution with respect to the *no antenna selection* case. Such behavior can also be observed in Fig.3.9 where a histogram of antenna usage ratio is shown for both optimization criteria. Clearly, the PHY-based method tends to use a larger number of antennas whereas throughput criterion is more conservative, in particular when no packet combining method is used.

3.4 Full vs. Restricted Antenna Selection Mechanisms

The antenna selection algorithms presented above allow the number of active transmit antennas vary from burst to burst, up to a maximum number of transmit antennas. Throughout this section, such an approach will be referred to as *Full Selection* since, any combination of *up to* M_{FS} (referred to as M, in previous sections) transmit antennas can be used. Clearly, this becomes an effective means of conducting rate adaptation. However, complexity for this approach

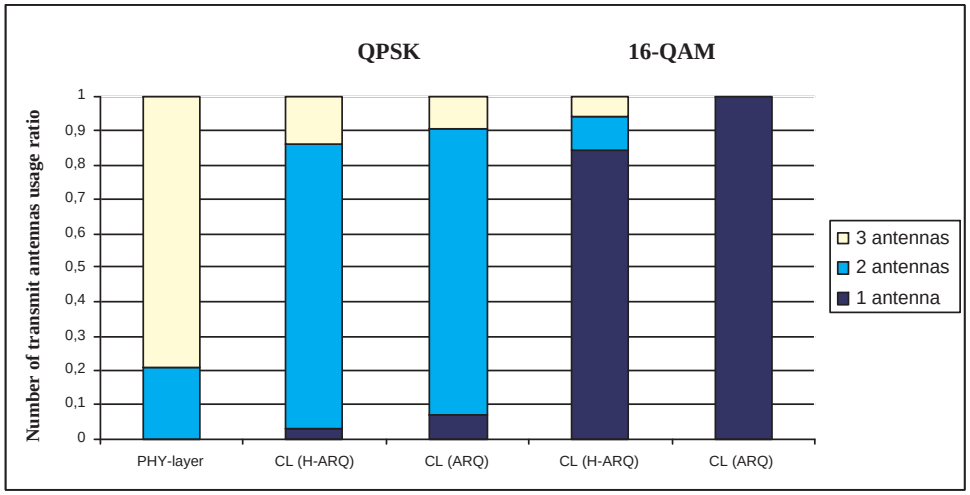
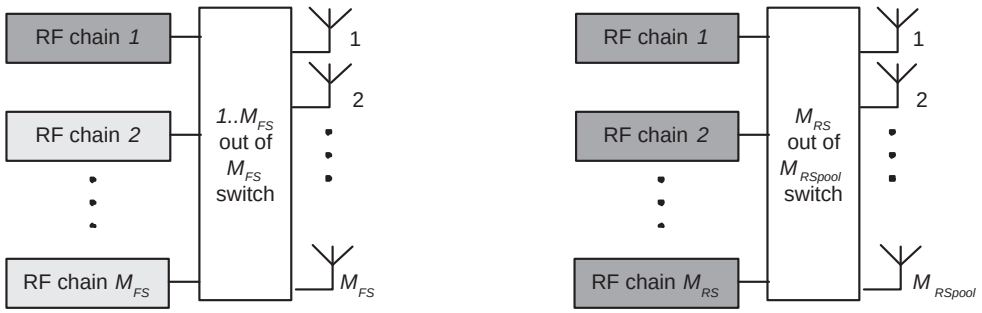


Figure 3.9: Relative frequency of 1, 2 and 3-antenna combinations for different antenna selection strategies and modulation schemes. No chase combining of retransmitted packets was conducted for bars labelled with 'ARQ' (SNR=12 dB).



Variable number of active RF chains

Pre-determined number of active RF chains

Figure 3.10: Full (left) and Restricted (right) Selection mechanisms.

increases exponentially in M_{FS}. Alternatively, one could think of selecting a pre-determined number of transmit antennas, M_{RS}, out of a set of M_{RSpool} so as to maximize a specific target function like mutual information [San04a, Mol04]. Ultimately, this approach is aimed at reducing hardware costs by mapping the outputs of a reduced number (M_{RS}) of costly RF chains into a larger number of inexpensive antenna elements ($M_{RSpool} > M_{RS}$). As shown in [Mol03a], the behavior of a reduced system with AS approaches that of a full system for a moderate number of antennas in the pool. In the sequel, this approach will be referred to as *Restricted Selection* (RS) in the sense that the actual number of active transmit antennas is defined beforehand (see Fig. 3.10), this acting as an additional constraint in the optimization process. On the one hand, the existence of a pool of antenna elements may provide additional degrees of spatial diversity. On the other hand, setting the number of antennas in advance could favor error propagation

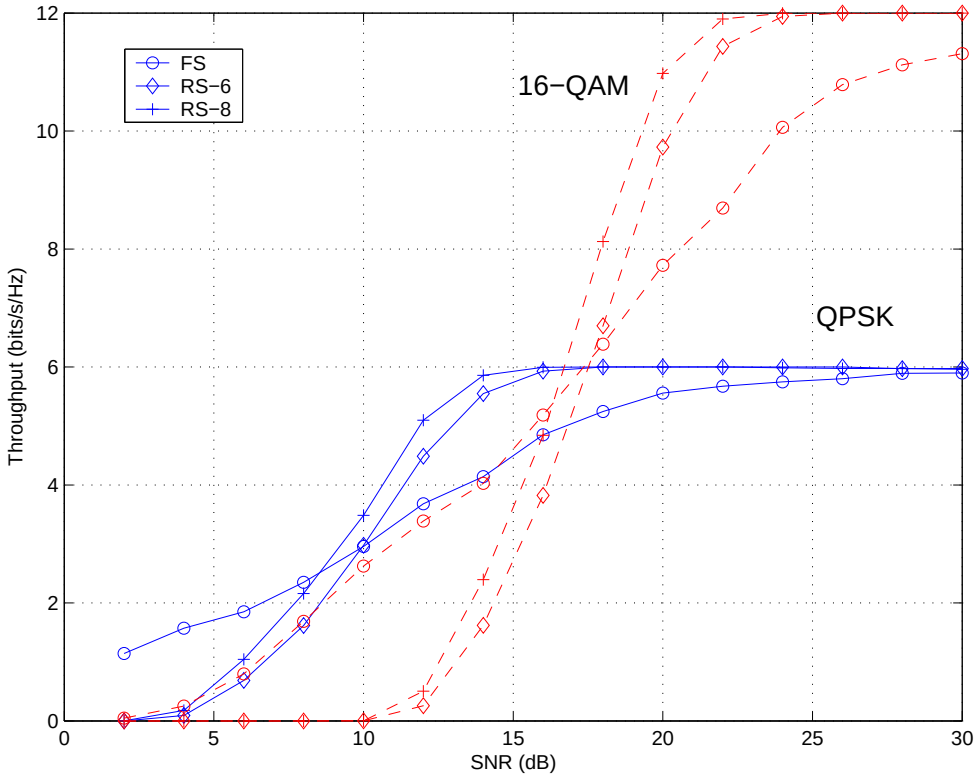


Figure 3.11: Throughput vs. SNR for the full-selection (FS) and restricted selection (RS) mechanisms operating under a throughput-based (cross-layer) antenna selection criterion ($N = 3, W = 3, N_{SAW} = 3$).

in the detection procedure. Results presented in the following paragraphs provide some insight into those issues.

For the sake of fairness, algorithm performance will be assessed for a fixed number of RF chains on the transmit side, that is, $M_{FS} = M_{RS}$ (i.e. same hardware cost). As for the RS scheme, two different cases with pools of $M_{RSpool} = 6$ and $M_{RSpool} = 8$ antenna elements will be analyzed (referred to as RS-6 and RS-8, in the sequel). As commented previously antenna selection commands are limited to three bits per frame and, as a consequence, RS antenna selection commands need a higher number of frames to be transmitted in comparison with the FS scheme (where $M_{FS} = 3$). This means a more limited degree of adaptability to channel variations as commands must be sent every two or three frames in RS-6 and RS-8 schemes, respectively.

In Fig. 3.11, some performance results for the FS and RS antenna selection schemes are shown. In the low-SNR region, the FS scheme clearly outperforms its RS counterpart; it is much better to reduce rate (by lowering the number of active transmit antennas) and, thus, focus transmit power, rather than keep the maximum number of actual transmit antennas and rely on the enhanced diversity coming from a larger pool of antenna elements. However, algorithm behavior is totally different in the high-SNR parts of the plot since, as soon as the FS approach

attains more frequently the maximum number of transmit antennas (i.e., three), diversity gains further favor RS schemes. In the mid-SNRregion (8-12 dB and 14-18 dB for QPSK and 16-QAM respectively), it can be observed that an additional benefit can be obtained by selecting transmit antennas out of a set of $M_{RSpool} = 8$, instead of $M_{RSpool} = 6$.

3.5 Inclusion of Adaptive Modulation Schemes

In this section, we will further enhance the CL antenna selection algorithm presented in Section 3.3 by including adaptive modulation (AM) schemes in the overall framework. As shown later in this section, the resulting scheme makes an efficient use of this additional feature, in terms of delay and throughput.

3.5.1 AM-enhanced Cross-layer Approach (AM-CL)

It is widely acknowledged that adaptive modulation and coding techniques can further improve spectral efficiency by adjusting transmission parameters to time-varying channel conditions [Gol97], in particular when applied to MIMO configurations [Cau02]. Adaptive modulation schemes in combination with antenna selection strategies were introduced in e.g. [Seb00, Zhu03]; by switching off antennas elements unable to meet specific QoS requirements the overall interference level can be reduced and, thus, higher constellation sizes can be supported by the remaining ones.

When adopting a CL criterion, though, the integration of an adaptive modulation scheme with the AS strategy comes in a very natural way, this rendering unnecessary separate optimizations (often resorting to heuristic criteria). More precisely, the optimization consists in jointly determining (by exhaustive search) both the *best* subset of transmit antennas along with the corresponding modulation scheme which maximize the link throughput given by Eq. (3.3). In order not to further increase algorithmic complexity and keep signalling in the feedback channel low, the same modulation scheme is shared by every single active antenna. In a CL antenna selection context, though, this additional restriction does not have a major impact in performance as illustrated in Fig. 3.12. In the low and high-SNR regions, selecting a common modulation scheme for all active antennas (the lowest and the highest, respectively) does definitely make sense. In the mid-SNR region, additional degrees of freedom are provided by the antenna selection mechanisms, this reverting in an improved granularity in terms of effective data rates. In other words, the fact that both the constellation size *and* the actual transmit antennas are jointly selected, partially compensates for the losses resulting from failing to adapt constellation size on a per-antenna basis. Clearly, this would no longer be possible in a system where the number of transmit antennas is kept constant.

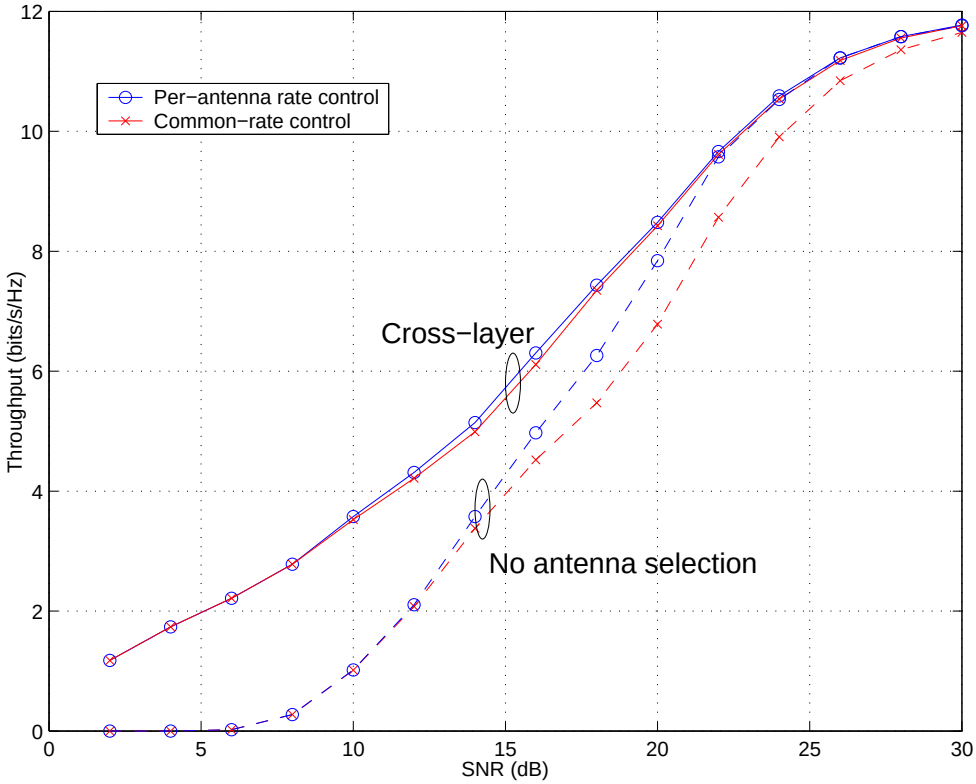


Figure 3.12: Performance penalty resulting from the use of identical constellation size in all streams (*common rate control*) instead of specific constellation sizes (*per-antenna rate control*).

As an additional remark, the (linear) increase in computational burden resulting from throughput scoring over different modulation schemes can be considered a minor issue since the largest contribution to computational complexity comes from the computation of vector norms $\|\mathbf{w}_k\|$ (see Section 3.6 ahead for details).

3.5.2 Computer Simulation Results

Figure 3.13 shows how the AM extension of the cross-layer antenna selection algorithm makes the most of the different modulation schemes (QPSK and 16 QAM). Results are compared with an extension of the PHY-layer antenna selection methodology presented in Subsection 3.3.1. More precisely, this extension (referred as AM-PHY) consists in first selecting the transmit antennas sub-set according to the mentioned PHY-based criterion and, after that, the largest constellation size for which the resulting PER is below a specific threshold, PER_{th}, is chosen. In this case, the same modulation scheme is also shared by all active antennas. As for the threshold PER, it has been empirically set to $\text{PER}_{th} = 0.2$ in order to maximize the resulting data link layer throughput[4].

[4]In a practical system, PER_{th} would be given by the QoS requirements of the associated data service.

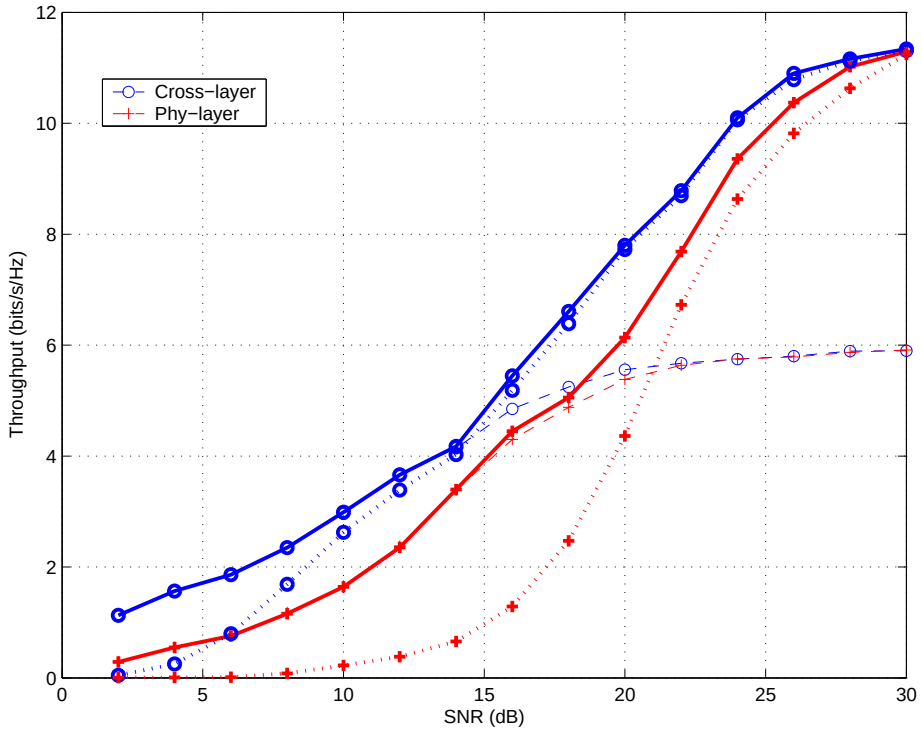


Figure 3.13: Performance in terms of throughput vs. SNR for the throughput-based (cross-layer) and PHY-based antenna selection schemes *with* adaptive modulation. (Solid lines: adaptive modulation, dashed lines: QPSK, dotted lines: 16-QAM).

.

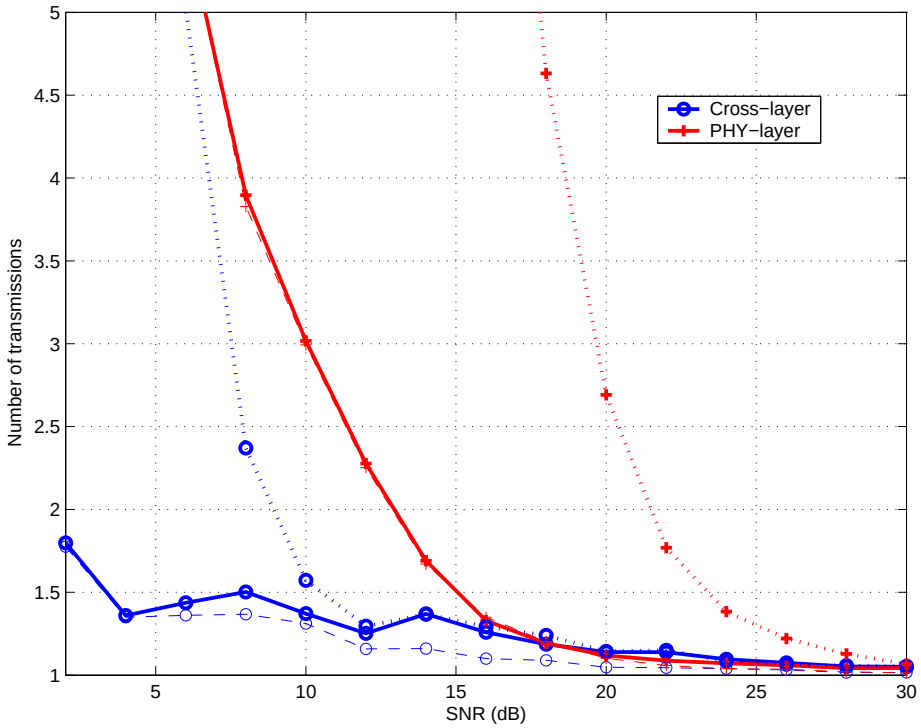


Figure 3.14: Average number of transmissions with and without adaptive modulation schemes for both antenna selection criteria. (Solid lines: adaptive modulation, dashed lines: QPSK, dotted lines: 16-QAM.)

To start with, one can observe that the resulting AM-PHY throughput curve runs well below its AM-CL counterpart for the whole range of SNR ratios. Actually, there is a 3-4 dB shift in average SNR. Besides, in terms of transmission delay (Fig.3.14), the AM-CL approach, in combination with H-ARQ mechanisms, provides a means to effectively keep the number of *individual* packet transmissions low for the whole range of signal to noise ratios (less than 2). Conversely, the number of packet retransmissions grows exponentially in the low-SNR region for the AM-PHY approach. This fact indicates that error propagation cannot be circumvented by including an AM scheme since the initial selection of transmit antennas (on the basis of PHY-based criterion) is not appropriate.

3.6 Reduced-complexity Schemes for CL Antenna Selection

As mentioned in Subsection 3.3.1 above, there is a need for an exhaustive search in the proposed full antenna selection process. More precisely, a total of $N_{subs} = 2^M - 1$ different subsets of $m = 1, ..., M$ antennas must be taken into account for throughput scoring. This may be considerably restrictive in practical systems due to the iterative nature of the detection algorithm as can be observed in Table 3.1, where $(\mathbf{G}^T)_{l_k}$ stands for the l_k-th column of matrix $\mathbf{G}^T$ and $\mathbf{H}_{\overline{l_k}}$ is the matrix constructed by dropping the columns $l_1, l_2, ..., l_k$ of matrix $\mathbf{H}$. For each antenna subset of dimension m, a total of $m-1$ pseudo-inverse matrices must be computed for the corresponding $N \times m$ channel matrices in subsequent iterations (layers) of the detection procedure loop, each with a complexity of $O\left(m^3\right)$. Additionally, the Euclidean norm for every single row in those matrices has to be computed so as to determine the right ordering for data detection and, also, for throughput scoring. However, we can do better by observing that:

1. *No data detection/cancellation steps are needed.* Ultimately, throughput expressions only depend on $\|\mathbf{w}_k\|^2$ and $\mathbb{E}\left[\left|s_{l_k}\right|^2\right]$ [5] which, in turn, can be directly obtained from $\mathbf{H}_m$.

2. *The computation of matrix* $\mathbf{G}$ *is not necessary.* As a matter of fact, we are not interested in the rows of $\mathbf{G}$, $\mathbf{w}_k^T$, but in their norms.

3. *A finite set of* N_{subs} *channel matrices* $\mathbf{H}_m$ *exist.* As a consequence, some throughput expressions for $(m-1)$-antenna subsets can use vector norms already computed for specific m-antenna combinations.

In the following, the ideas listed above will be exploited in order to reduce the inherent complexity of the exhaustive search. It should be noted that, in contrast with the recursive *sub-optimal* approaches derived in [GA04] and [Gor02c], we derive a recursive approach capable

[5] Notice that the energy of the transmitted symbols do depend on the number of antennas (m) in the subset under consideration (total transmitted power is kept constant).

Table 3.1: Detection algorithm for the V-BLAST scheme with ZF spatial filtering as proposed in [Gol99].

Initialization	:		
$\mathbf{r}_1$	=	$\mathbf{r}$	
$\mathbf{G}_1$	=	$\mathbf{H}^{\#}$	
Recursion	:	$for\ k = 1, ..., M\ do$	
l_k	=	$\arg\max_j \mathsf{SINR}_j = \arg\min_j \left\|(\mathbf{G}_k)_j\right\|^2$	(3.4)
$\mathbf{w}_k$	=	$\left(\mathbf{G}_k^T\right)_{l_k}$	(3.5)
y_{l_k}	=	$\mathbf{w}_k^T \mathbf{r}_k$	(3.6)
$\hat{s}_{l_k}$	=	$Q\left(y_{l_k}\right)$	(3.7)
$\mathbf{r}_{k+1}$	=	$\mathbf{r}_k - \hat{s}_{l_k}(\mathbf{H})_{l_k}$	(3.8)
$\mathbf{G}_{k+1}$	=	$\mathbf{H}_{\overline{l_k}}^{\#}$	

of identifying the optimal antenna subset (in a CL sense) since it incurs no inaccuracy in the recursive computation of vector norms or in the exploration of different antenna subsets.

3.6.1 Recursive Computation of Vector Norms

In order to recursively compute the row vectors of an arbitrary matrix $\mathbf{G}$ (Remark 2), we will start by re-writing their expression as:

$$\|\mathbf{w}_k\|^2 = \mathbf{w}_k^H \mathbf{w}_k = \left[\mathbf{G}\mathbf{G}^H\right]_{k,k}$$

By simply expressing $\mathbf{G}\mathbf{G}^H$ in terms of the SVD of $\mathbf{H}$ and taking into consideration that $\mathbf{G} = \mathbf{H}^{\#} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H$, one can realize that:

$$\mathbf{G}\mathbf{G}^H = \mathbf{V}\left(\mathbf{\Sigma}^H\mathbf{\Sigma}\right)^{-1}\mathbf{V}^H = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}$$

The right-handside of the equation above lends itself more easily to a recursive implementation by invoking the block matrix inversion lemma [Mag02]. This approach was also adopted in [Ben03] but, in a context where no antenna selection mechanisms (i.e., only a detection procedure for V-BLAST) were considered and for the MMSE interference cancellation criterion. Now, the recursive computation problem can be stated as follows: "Given $\mathbf{H}_m$, $\mathbf{H}_{m-1}$ and $\mathbf{G}_m\mathbf{G}_m^H = \left(\mathbf{H}_m^H\mathbf{H}_m\right)^{-1}$ compute $\mathbf{G}_{m-1}\mathbf{G}_{m-1}^H = \left(\mathbf{H}_{m-1}^H\mathbf{H}_{m-1}\right)^{-1}$ without resorting to any matrix inversion". Channel matrix $\mathbf{H}_{m-1}$ is obtained by removing the column vector in $\mathbf{H}_m$ associated to the row vector in $\mathbf{G}_m$ with the lowest norm. In order to exploit the block matrix inversion

lemma, assume that such a row is the first one, that is, $\mathbf{H}_m = \left[\begin{array}{cc} \mathbf{h} & \mathbf{H}_{m-1} \end{array} \right]$. Consequently,

$$\begin{aligned} & \left(\mathbf{H}_m^H \mathbf{H}_m\right)^{-1} \\ &= \begin{pmatrix} \mathbf{h}_1^H \mathbf{h}_1 & \mathbf{h}_1^H \mathbf{H}_{m-1} \\ \mathbf{H}_{m-1}^H \mathbf{h}_1 & \mathbf{H}_{m-1}^H \mathbf{H}_{m-1} \end{pmatrix}^{-1} \doteq \begin{pmatrix} a_{11} & \mathbf{a}_{12} \\ \mathbf{a}_{21} & \mathbf{A}_{22} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} e^{-1} & -e^{-1}\mathbf{a}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{a}_{21}e^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{a}_{21}e^{-1}\mathbf{a}_{12}\mathbf{A}_{22}^{-1} \end{pmatrix} \\ &\doteq \begin{pmatrix} d_{11} & \mathbf{d}_{12} \\ \mathbf{d}_{21} & \mathbf{D}_{22} \end{pmatrix} \end{aligned} \tag{3.9}$$

It is worth noting that d_{11}, $\mathbf{d}_{12}$, $\mathbf{d}_{21}$, $\mathbf{D}_{22}$ are known since we initially assumed $\left(\mathbf{H}_m^H \mathbf{H}_m\right)^{-1}$ to be known. Besides, one should also realize that $\left(\mathbf{H}_{m-1}^H \mathbf{H}_{m-1}\right)^{-1} = \mathbf{A}_{22}^{-1}$. Hence, by simply identifying elements in the above expression

$$\begin{aligned} \mathbf{D}_{22} &= \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{a}_{21}e^{-1}\mathbf{a}_{12}\mathbf{A}_{22}^{-1} \\ &= \mathbf{A}_{22}^{-1} + \left(\mathbf{A}_{22}^{-1}\mathbf{a}_{21}e^{-1}\right) e \left(e^{-1}\mathbf{a}_{12}\mathbf{A}_{22}^{-1}\right) \\ &= \mathbf{A}_{22}^{-1} + \mathbf{d}_{21}d_{11}^{-1}\mathbf{d}_{12} \end{aligned}$$

a recursion for $\left(\mathbf{H}_{m-1}^H \mathbf{H}_{m-1}\right)^{-1}$ in terms of the elements in $\left(\mathbf{H}_m^H \mathbf{H}_m\right)^{-1}$can be finally obtained:

$$\left(\mathbf{H}_{m-1}^H \mathbf{H}_{m-1}\right)^{-1} = \mathbf{A}_{22}^{-1} = \mathbf{D}_{22} - \mathbf{d}_{21}d_{11}^{-1}\mathbf{d}_{12} \tag{3.10}$$

Indeed only one addition of matrices, one vector product and a single scalar inversion (but no matrix inversion) are needed. Reciprocally, Eq. (3.9) can also be used to obtain $\left(\mathbf{H}_m^H \mathbf{H}_m\right)^{-1}$ from the known matrices/vectors $\left(\mathbf{H}_{m-1}^H \mathbf{H}_{m-1}\right)^{-1}$, $\mathbf{H}_{m-1}$ and $\mathbf{h_1}$ since, clearly, all the elements in that expression can be computed from them, that is:

$$\begin{aligned} e &= \mathbf{h}_1^H \mathbf{h}_1 - \mathbf{a}_{12}\mathbf{A}_{22}^{-1}\mathbf{a}_{21} \\ \mathbf{a}_{12} &= \mathbf{h}_1^H \mathbf{H}_{m-1} \\ \mathbf{a}_{21} &= \mathbf{H}_{m-1}^H \mathbf{h}_1 \\ \mathbf{A}_{22}^{-1} &= \left(\mathbf{H}_{m-1}^H \mathbf{H}_{m-1}\right)^{-1} \end{aligned}$$

Finally, we will go back to the problem of column ordering. In our derivation, we initially assumed the column in $\mathbf{H}_m$ to be removed was the first one. Should this not be the case, the appropriate column in $\mathbf{H}_m$ can be moved to the first position by simply using a permutation matrix $\mathbf{P}_H$,that is, $\mathbf{H}_{m,p} = \mathbf{H}_m\mathbf{P}_H$ where the subscript p stands for *permutation.* Accordingly, the following equation

$$\begin{aligned} \left(\mathbf{H}_{m,p}^H \mathbf{H}_{m,p}\right)^{-1} &= \left(\mathbf{P}_H^H \mathbf{H}_m^H \mathbf{H}_m \mathbf{P}_H\right)^{-1} \\ &= \mathbf{P}_H^H \left(\mathbf{H}_m^H \mathbf{H}_m\right)^{-1} \mathbf{P}_H \end{aligned}$$

relates the know data, $\left(\mathbf{H}_m^H \mathbf{H}_m\right)^{-1}$, with the elements on the right handside of Eq. (3.10) ($\mathbf{D}_{22}$, $\mathbf{d}_{21}$, d_{11}^{-1} and $\mathbf{d}_{12}$) that, now, can be directly obtained from matrix $\left(\mathbf{H}_{m,p}^H \mathbf{H}_{m,p}\right)^{-1}$. For simplicity, we will make no distinction between $\mathbf{H}_m$ and $\mathbf{H}_{m,p}$ in subsequent sections.

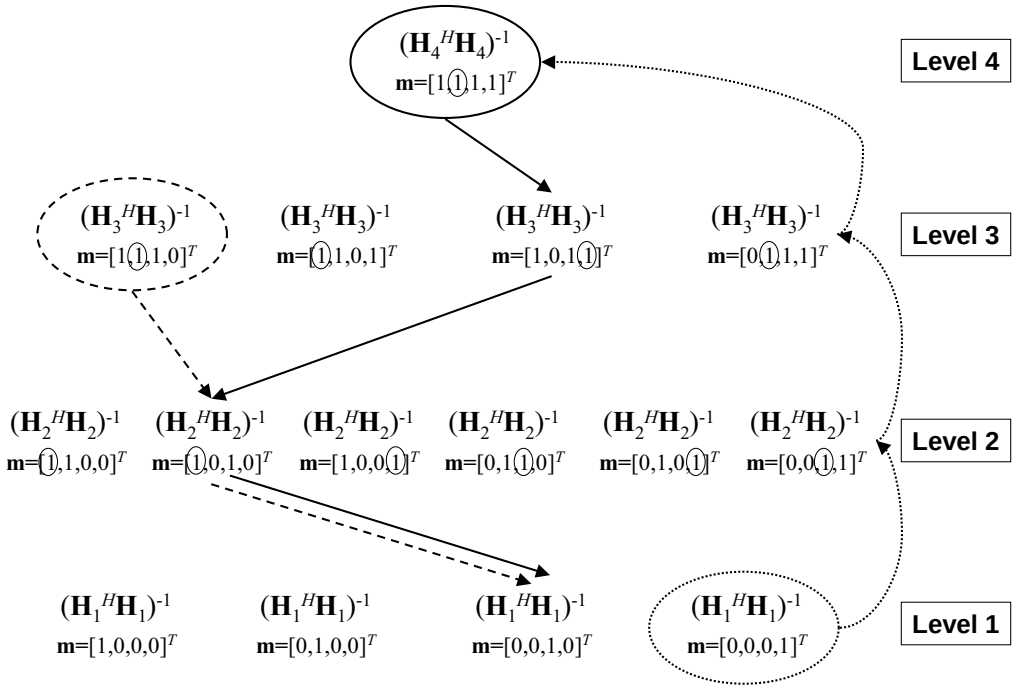


Figure 3.15: Trellis arrangement for the overall set of norm matrices of up to $M = 4$ transmit antennas. Circles on some elements in the corresponding $\mathbf{m}$-vectors denote vector $\mathbf{w}_k$ with the lowest norm. Solid and dashed arrows reflect *descending* paths for two antennas subsets of 3 and 4 transmit antennas, respectively. Dotted curves show how to obtain an initial estimate for $\left(\mathbf{H}_4^H\mathbf{H}_4\right)^{-1}$ by means of the *bottom-up* approach.

3.6.2 Simplified Computation of the Overall Set of Vector Norms

On the basis of the recursive procedure presented in the previous subsection, we are actually interested in determining the norms of the elements in the diagonal of the $\left(\mathbf{H}_m^H\mathbf{H}_m\right)^{-1}$ matrices for *every single* combination of *up to* M transmit antennas. As an example, Fig. 3.15 displays the whole set of $\left(\mathbf{H}_m^H\mathbf{H}_m\right)^{-1}$ matrices of up to $M = 4$ transmit antennas in a trellis arrangement.

- *Top-down approach*:

An exhaustive search for a specific channel realization would start by considering as a starting point the only combination of $M = 4$ antennas. A descending path to the last row of channel matrices (of dimension $N \times 1$) would result. A sequence of vector norms associated to each transition in that descending path would thus be obtained. By plugging those norms along with $\mathbb{E}\left[|s_k|^2\right]$ (depending on the inverse of the number of antennas in the starting point matrix) into

Eq. (3.2) and Eq. (3.3) one would obtain the corresponding score in terms of throughput. Next, combinations with matrices of $M = 3$ antennas as a starting point would follow.

From the rationale given in the preceding paragraph, we conclude that every single channel matrix of up to M transmit antennas and their corresponding norm matrices, $\left(\mathbf{H}_m^H\mathbf{H}_m\right)^{-1}$, have to be computed in order to find out the best subset of transmit antennas. However, it is not necessary to recompute those norm matrices for channel matrices included in the paths defined by starting points in upper levels since paths will merge from that point on (see Fig. 3.15). Instead, computations of those matrices and the norms associated to transitions can be done in advance (recall Remark 3) by applying the recursion in Eq. (3.10).

Such a *top-down* recursive approach achieves substantial computational savings but, unfortunately, knowledge of $\left(\mathbf{H}_M^H\mathbf{H}_M\right)^{-1}$ is still needed. One could think about exploiting the recursion in the reverse order, that is, starting with an arbitrary vector channel $\mathbf{H}_1$ (of dimension $N \times 1$) and proceed with the addition of columns (i.e. transmit antennas) all the way up through the only $N \times M$ matrix on level M. By considering the recursive formula in Eq. (3.9), the number of operations is remarkably lower with respect to those of direct inversion via SVD decomposition.

- *Bottom-up approach*:

Going one step further, a bottom-up approach could also be used to obtain the whole set of norm matrices on levels $1, ..., M$. Recursion would start on Level 1 in the trellis. Next, by appending additional channel vectors as the left-most column, matrices in the upper level and the associated vector norms for every row in the ZF solution would be computed. This procedure would be iterated until all the elements/norms in the trellis are computed. After that, *descending* paths are identified and, finally, throughput for each combination of transmit antennas can be readily computed so as to find the optimal transmit subset out of all the possible ones.

3.6.3 Computational Complexity Analysis

In this section, the so-called *top-down* and *bottom-up* recursive approaches will be compared with the *full-complexity* (i.e. non-recursive) one in terms of computational complexity. In all the cases, contributions associated to throughput scoring computations will be neglected since computational burden is much lower than that of the antenna selection and data detection steps.

Computational complexity, expressed in the number of floating-point operations (FLOPS), for the *full-complexity*, *bottom-up* and *top-down* procedures amount to[6] (see Appendix 3.A for

[6]Notice that six and two FLOPS are needed for each complex product and addition, respectively.

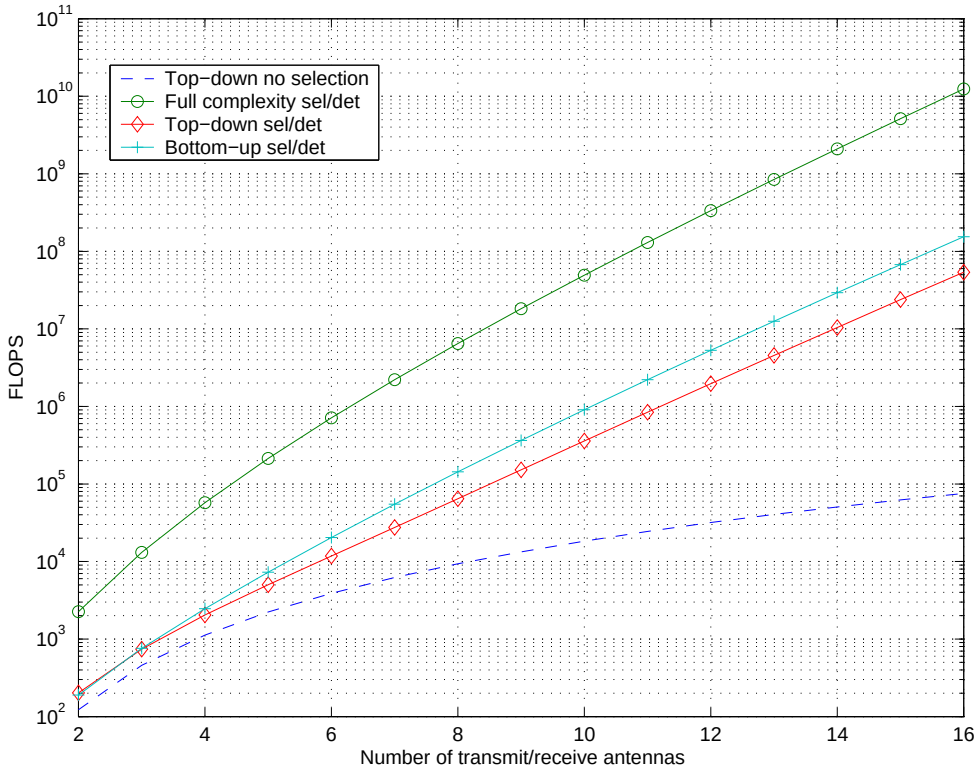


Figure 3.16: Number of operations (FLOPS) vs. maximum number of transmit antennas (M).

details):

$$F_{full}(M,N) = 16NM - 2M - 8N + \sum_{m=1}^{M} \binom{M}{m} \left[72m^3 + 72Nm^2 + 32N^2m + 12Nm + 4m\right] \quad (3.11)$$

$$F_{bu}(M,N) = 4NM^2 + 26NM - 8N + 2 + \sum_{m=1}^{M-1} \binom{M}{m+1} \left[24m^2 + 8Nm + 8N + 8m + 4\right] \quad (3.12)$$

$$F_{td}(M,N) = 8M^3 + 8NM^2 - 8M^2 + 22NM - 8N + 2 + \sum_{m=1}^{M-1} \binom{M}{m} \left[14m^2 + 6\right] \quad (3.13)$$

Curves showing results in terms of FLOPS vs. the maximum number of transmit antennas (M) can be found in Figure 3.16, where, additionally, it has been considered the same number of receive and transmit antennas ($N = M$). The curve labelled with *top-down-no selection* is used for benchmarking since it constitutes a lower bound for complexity. Clearly, both recursive approaches attain substantial savings in comparison with the *full complexity* version. For a moderate number of transmit antennas, say $M = 3$, computational burden is reduced by a factor of 15, approximately. The difference becomes even larger as the number of transmit antennas grows (almost two orders of magnitude for a case with $M = 16$ antennas). Out of the two recursive versions, the *top-down* approach features lower number of FLOPS (half the number of operations, in an $M = 6$ case).

Table 3.2: Computational complexity analysis for a 3x3 MIMO system. The first column reflects the operation count for the antenna selection part (if applicable) and the computation of spatial filters. The second column accounts for the spatial filtering of $L = 160$-symbol vectors packets. The forth column stands for the number of operations per second considering packets of 0.667 ms time duration.

	Antenna selection + computation of spatial filters [FLOPS]	Spatial filtering [FLOPS]	Total [FLOPS]	Complexity [MFLOPS/s]
Top-down, no sel.	344	18,240	18,584	27.9
Top-down, sel.	626	18,240	18,866	28.3
Conventional, no sel.	7,584	18,240	25,824	38.7
Conventional, sel.	13,008	18,240	31,248	46.9

In order to provide the reader with a complementary point of view, computational complexity requirements in terms of floating point operation per time unit will be discussed. Results for a system with $M = 3$ transmit antennas and 0.667 ms packets (like HSDPA) are shown in Table 3.2. With respect to a scheme with a fixed number of transmit antennas (*top-down, no sel.*), the additional cost of conducting CL transmit antenna selection on the basis of a *top-down* recursive approach turns out to be 1.52% (in MFLOPS/s). However, savings with respect to non-recursive versions of the AS procedure (*conventional, sel.*) are close to 40%.

3.7 Switching between Spatial Diversity and Multiplexing from a Cross-layer Perspective

In this section, we propose a cross-layer approach to solve the problem of switching between spatial diversity and multiplexing in a H-ARQ context. In particular, three transmission modes are considered (diversity, spatial multiplexing and a hybrid mode) and the main purpose of the proposed scheme is the adaptive selection of those modes maximizing data link layer throughput.

3.7.1 Problem Formulation

As explained in Chapter 2, MIMO techniques are aimed at either enhancing diversity or providing spatial multiplexing capabilities. In fact, most of the current research is focused on exploiting only one of the gains provided by MIMO but, recently, studies that combine both gains have appeared in the literature [Zhe03]. In that direction, a system based on switching between multiplexing and diversity is proposed in [Hea00, Hea05]. According to instantaneous channels conditions, the transmission mode is switched in order to minimize the resulting SER. For a *constant* data rate, it was shown that by choosing the best mode for a given channel realization,

better results can be obtained than with the original approaches separately. A third transmission mode combining advantages of both MIMO approaches was included in [Cha04] with the aim of improving granularity in terms of SER. More precisely, the hybrid mode consists in using a Double Space-Time Transmit Diversity (D-STTD) technique [DST01], i.e., transmitting an independent Alamouti scheme in each pair of transmit antennas.

Switching between transmission modes can also be used in order to increase the data rate. For instance, authors in [For04] proposed selecting those transmission modes that maximize the spectral efficiency for a pre-determined target SER. In other words, the selection algorithm chooses the transmission mode with the highest data rate for which the resulting SER is below a specific threshold.

Our contribution is directed towards achieving the maximum possible throughput. To do so, we propose a system capable of adaptively switching between multiplexing and diversity but, rather than selecting the maximum transmission data rate subject to SER constraints, we directly maximize the link layer throughput expression. It is worth noting that a cross-layer approach was independently derived in [Lu05]. In that paper, however, the data link layer throughput was not directly maximized. Instead, a sub-optimum approach based on derive *effective* PER expressions according to the ARQ strategy in use was taken into consideration. In addition, the granularity is reduced in comparison with our approach as adaptive modulation is not performed and the available transmission schemes are OSTBC, Q-OSTBC and V-BLAST (i.e., rate 3/4, rate 1 and rate 4, respectively).

In this section we consider a transmission link between an M-antenna BS and an N-antenna MS like Figure 3.17, where the same system parameters as in Sect. 3.2 are adopted except for the transmission schemes and the receiver structure. As shown in the figure, three transmission modes are used: diversity, spatial multiplexing and hybrid diversity/multiplexing mode. As for the receiver structure, ZF detection scheme is used for all the transmission modes in order to keep computational complexity moderate.

By stacking T consecutive data samples, the received vector at the i-th sensor ($\mathbf{r}_i = [r_i(1), ..., r_i(T)]^T$) can be written as:

$$\mathbf{r}_i = \mathbf{S}\mathbf{h}_i + \mathbf{n}_i$$

where $\mathbf{S} \in \mathbb{C}^{T \times M}$ is the symbol matrix describing the transmission block code, according to the modulation scheme R (QPSK or 16-QAM), $\mathbf{h}_i = [h_{i_1}, ..., h_{i_M}]^T \in \mathbb{C}^{T \times 1}$ is the channel vector corresponding to the i-th receiver, and $\mathbf{n}_i \in \mathbb{C}^{T \times 1}$ stands for an additive Gaussian noise vector of complex, random variables with zero mean and variance σ^2. Again, power is evenly distributed among transmit antennas. Channel knowledge is used at the receiver to jointly estimate both the optimal transmission mode and modulation scheme maximizing link layer throughput according to the H-ARQ strategy under consideration. Once the transmit configuration is selected, three

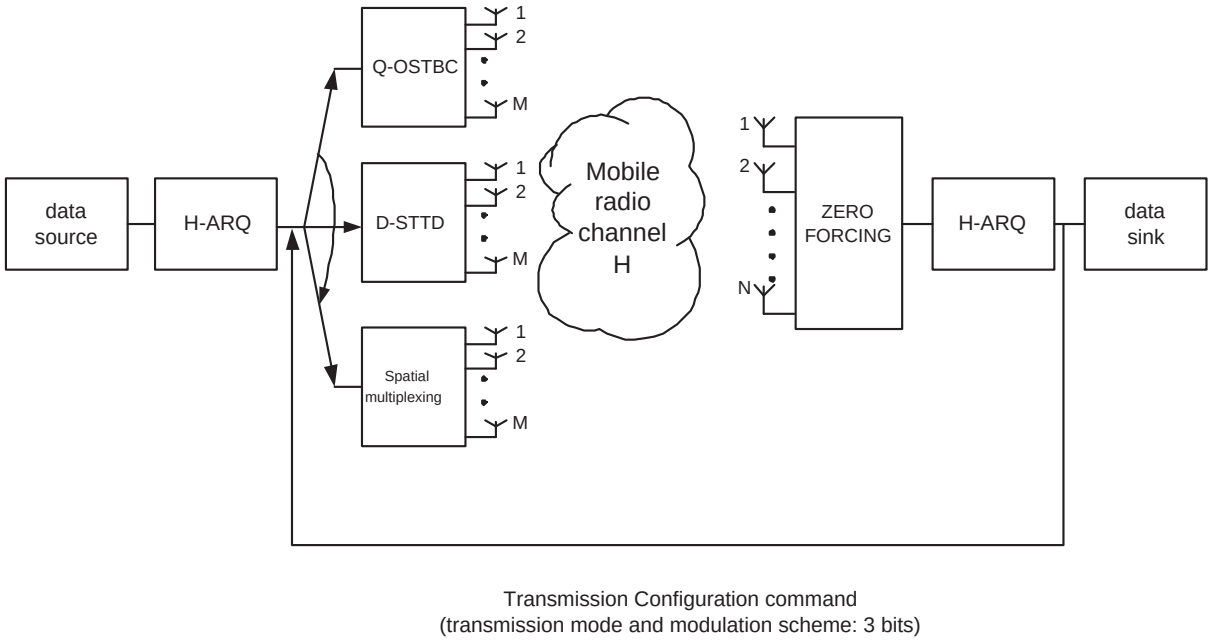


Figure 3.17: Block diagram of a MIMO communication system with an adaptive switching of the transmission mode.

bits of the feedback channel are utilized to convey this information to the transmitter[7].

In order to properly analyze the different transmission modes with a minimum number of receive antennas at the MS, the number of transmit and receive antennas will be set to $M = N = 4$. This is because an even number of antennas $M \geq 4$ is required for the D-STTD scheme at the transmit side, whereas $N \geq M$ antennas are needed for the spatial multiplexing mode at the receiver.

3.7.2 Transmission Modes

Next, we present the different transmission modes available at the BS:

- *Diversity mode*:

A Quasi-Orthogonal STBC (Q-OSTBC) code is considered for the diversity mode. Although full diversity is not obtained, we adopt this strategy because full rate ($r = 1$) is achieved. Moreover, better performance than with orthogonal designs is obtained in the low-SNR regime[8]

[7]This is due to three transmission modes and two modulation schemes (QPSK and 16-QAM) are considered, i.e., six transmission configurations are available at the transmitter.

[8]Notice that, the diversity mode will be usually selected in the low-SNR region.

[Jaf01b]. By using this strategy, the symbol block matrix $\mathbf{S}$ is given by:

$$\mathbf{S} = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 \\ s_2^* & -s_1^* & s_4^* & -s_3^* \\ s_3^* & s_4^* & -s_1^* & -s_2^* \\ s_4 & -s_3 & -s_2 & s_1 \end{bmatrix}$$

For the ease of notation, the received signal can be rewritten as:

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{s} + \mathbf{v}$$

where vectors $\mathbf{y}_i$ and $\mathbf{v}_i$ have been defined as:

$$\mathbf{y}_i = [r_{i_1}, r_{i_2}^*, r_{i_3}^*, r_{i_4}]$$
$$\mathbf{v}_i = [n_{i_1}, n_{i_2}^*, n_{i_3}^*, n_{i_4}]^T$$

respectively, and $\mathbf{H}_i$ stands for the equivalent space-time channel matrix:

$$\mathbf{H}_i = \begin{bmatrix} h_{i_1} & h_{i_2} & h_{i_3} & h_{i_4} \\ -h_{i_2}^* & h_{i_1}^* & -h_{i_4}^* & h_{i_3}^* \\ -h_{i_3}^* & -h_{i_4}^* & h_{i_1}^* & h_{i_2}^* \\ h_{i_4} & -h_{i_3} & -h_{i_2} & h_{i_1} \end{bmatrix} \tag{3.14}$$

Notice that matrix $\mathbf{H}_i$ is quasi-orthogonal in the sense that:

$$\mathbf{H}_i^H \mathbf{H}_i = h_i^2 \begin{bmatrix} \mathbf{I}_2 & X_i \mathbf{J} \\ -X_i \mathbf{J} & \mathbf{I}_2 \end{bmatrix} \tag{3.15}$$

where $\mathbf{J}$ is the 2×2 matrix $\mathbf{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, X_i is a real-valued random variable defined as $X_i = 2\mathrm{Re}\left(h_{i_1} h_{i_4}^* - h_{i_2} h_{i_3}^*\right)/h_i^2$ and $h_i^2 = |h_{i_1}|^2 + |h_{i_2}|^2 + |h_{i_3}|^2 + |h_{i_4}|^2$ is the channel gain (see Appendix 3.B for further details).

As commented in the previous section, a ZF detector is adopted in all the transmission modes. Prior to data detection, the received signal at the different branches are match-filtered and coherently combined:

$$\mathbf{z} = \sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{y}_i = \left[\sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{H}_i\right] \mathbf{s} + \sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{v}_i \tag{3.16}$$

After that, the transmitted symbols can be estimated as:

$$\hat{\mathbf{s}} = \left[\sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{H}_i\right]^{-1} \mathbf{z} = \mathbf{s} + \left[\sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{H}_i\right]^{-1} \sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{v}_i$$

Then, the signal-to-noise ratio corresponding to the k-th symbol can be estimated as:

$$\gamma_k = \frac{\mathbb{E}\left[|s_k|^2\right]}{\sigma^2 \|\mathbf{w}_k\|^2} \qquad k = 1, 2, \ldots, M, \tag{3.17}$$

where w_k, $k = 1, ..., M$, stand for the row vectors in matrix $\mathbf{W}$:

$$\mathbf{W} = \left[\sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{H}_i\right]^{-1} \sum_{i=1}^{N} \mathbf{H}_i^H = \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \mathbf{w}_3^T \\ \mathbf{w}_4^T \end{bmatrix} \tag{3.18}$$

By bearing in mind that an N-SAW hybrid ARQ strategy in combination with a chase combining scheme is considered, γ_k will ultimately depend on the actual number of recombined soft-symbol packets, p. That is, the effective SNR can be expressed as $\gamma_{k,p} = \alpha^{p-1} p \gamma_k$. Finally, given that all the row vector norms are identical, the signal-to-noise ratio corresponding to the diversity mode can be expressed as[9]:

$$\gamma_{\mathrm{DIV},p} = \alpha^{p-1} p \frac{\mathbb{E}\left[|s_k|^2\right]}{\sigma^2 \|\mathbf{w}_1\|^2} \tag{3.19}$$

- *Hybrid mode*:

The hybrid mode is based on the transmission of four different symbols during two consecutive time intervals. Then, at the expense of half diversity gain, data rate is doubled ($r = 2$). In particular, a D-STTD scheme is adopted, which results in the following symbol block:

$$\mathbf{S} = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 \\ s_2^* & -s_1^* & s_4^* & -s_3^* \end{bmatrix}$$

and the equivalent space-time channel matrix:

$$\mathbf{H}_i = \begin{bmatrix} h_{i_1} & h_{i_2} & h_{i_3} & h_{i_4} \\ -h_{i_2}^* & h_{i_1}^* & -h_{i_4}^* & h_{i_3}^* \end{bmatrix} \tag{3.20}$$

As in the diversity mode, the ZF detector is used. Hence, expressions (3.16) and (3.18) are still valid but one must take into account the new expression of matrix $\mathbf{H}$ given by Eq. (3.20).

Again, no matrix inversion is required to compute the vector norms (see Appendix 3.C) but, in this case, vector norms are related as follows:

$$\|\mathbf{w}_1\|^2 = \|\mathbf{w}_2\|^2$$
$$\|\mathbf{w}_3\|^2 = \|\mathbf{w}_4\|^2$$

As a consequence, the SNR expressions related to the two transmitted streams with the hybrid transmission mode can be written as:

$$\gamma_{\mathrm{HYB},1,p} = \alpha^{p-1} p \frac{\mathbb{E}\left[|s_k|^2\right]}{\sigma^2 \|\mathbf{w}_1\|^2} \qquad \gamma_{\mathrm{HYB},2,p} = \alpha^{p-1} p \frac{\mathbb{E}\left[|s_k|^2\right]}{\sigma^2 \|\mathbf{w}_3\|^2} \tag{3.21}$$

[9] In appendix 3.C, it is shown that all the row vector norms are identical and that they can be computed without resorting to any matrix inversion.

- *Spatial multiplexing mode*:

In this mode, four symbols are transmitted in parallel in each time-slot ($r = 4$):

$$\mathbf{S} = \left[\begin{array}{cccc} s_1 & s_2 & s_3 & s_4 \end{array} \right]$$

and, consequently, transmit spatial diversity cannot be exploited (for a $M = N = 4$ configuration).

The corresponding space-time channel matrix can be expressed as:

$$\mathbf{H}_i = \left[\begin{array}{cccc} h_{i_1} & h_{i_2} & h_{i_3} & h_{i_4} \end{array} \right]$$

and expressions (3.16) and (3.18) can also be used to obtain the SNR:

$$\gamma_{\mathrm{MULT},k,p} = \alpha^{p-1} p \frac{\mathbb{E}\left[|s_k|^2\right]}{\sigma^2 \left\|\mathbf{w}_k\right\|^2} \qquad k = 1,2,3,4 \tag{3.22}$$

3.7.3 Switching Criterion and Performance Evaluation

As commented above, the switching criterion is based on the maximization of the link layer throughput. In other words, the optimization process consists in jointly selecting the transmission mode *and* the modulation scheme that maximize the following expression:

$$\eta_{N-SAW} = \frac{N_{SAW}}{W} \frac{l \cdot r \cdot b}{(1 - \mathrm{PER}_1) + \sum_{p=2}^{P} \left[p(1 - \mathrm{PER}_p) \prod_{t=1}^{p-1} \mathrm{PER}_t \right]} \tag{3.23}$$

The above equation is identical to Eq. (3.3) except for the term r related to the rates of the different transmission modes and how the PER expressions are computed. In particular, the transmitted packet is always equal to LM symbols but the number of redundant symbols increases when spatial diversity is introduced. Then, the uncoded PER must be computed in accordance with the transmission modes as follows:

$$\begin{aligned} \mathrm{PER}_{\mathrm{DIV},p} &= 1 - (1 - \mathrm{SER}_{\mathrm{DIV},p})^L \\ \mathrm{PER}_{\mathrm{HYB},p} &= 1 - \left[\prod_{k=1}^{2} (1 - \mathrm{SER}_{\mathrm{HYB},k,p}) \right]^L \\ \mathrm{PER}_{\mathrm{MULT},p} &= 1 - \left[\prod_{k=1}^{4} (1 - \mathrm{SER}_{\mathrm{MULT},k,p}) \right]^L \end{aligned}$$

which can be easily computed with the accustomed expressions for the SER [Pro01] and the SNR expressions derived in the previous section:

$$\begin{aligned} &\mathrm{SER}_{\mathrm{DIV},p} = f(\gamma_{\mathrm{DIV},p}, R) && \\ &\mathrm{SER}_{\mathrm{HYB},k,p} = f(\gamma_{\mathrm{HYB},k,p}, R) && k = 1, 2 \\ &\mathrm{SER}_{\mathrm{MULT},k,p} = f(\gamma_{\mathrm{MULT},k,p}, R) && k = 1, ..., 4 \end{aligned}$$

To solve the optimization problem, an exhaustive search is also needed. Contributions to computational complexity mainly arise from the computation of the instantaneous signal-to-noise ratio expressions (equations (3.19), (3.21) and (3.22)) required for the estimation of the SER. For the spatial multiplexing case, the bottom-up approach developed in Section 3.6 can be adopted to obtain the vector norms. As for the Q-OSTBC and D-STTD schemes, it was previously shown that no matrix inversion is required.

Alternatively, when computational requirements at the receiver are further limited, a *sub-optimum* approach based on the scenario *statistics* can be adopted. That is, the average link throughput is pre-computed for the different transmission modes and modulation schemes. Then, a set of thresholds for the *average* SNR can be established. By doing so, the transmission mode and constellation size can be selected at the BS according to the average SNR of the system.

To conclude this section some simulation results are presented. In Figure 3.18, it can clearly be observed that considerably performance gains are obtained with an adaptive switching approach with respect to using the original schemes separately. It is also clear that, for the whole range of SNR, superior performance gain is obtained with the fast adaptive cross-layer technique (i.e., maximizing Eq. (3.23)) since this approach takes advantage of the instantaneous scenario conditions. Performance gains associated to CL are also reflected in the number of retransmissions behavior (Fig.3.19). It is curious enough to observe that the lowest number of transmissions are obtained with the sub-optimum approach. The reason for that being that the CL solution aims to obtain the maximum throughput and, to do so, it adaptively plays with the existing trade-offs in terms of the transmission rate of packets vs. the number of required retransmissions for successfully receiving these packets. This effect can be easily understood with the following example. In the 0-6 dB SNR region, the sub-optimum approach selects Q-OSTBC and QPSK in accordance with the SNR thresholds. In that case, the number of packets retransmission is minimized. The proposed CL selection, on the other hand, may increase the number of retransmissions by choosing Q-OSTBC and 16-QAM. By doing so, the packet delay is slightly increased but the obtained link throughput is higher when the transmitted packets are combined.

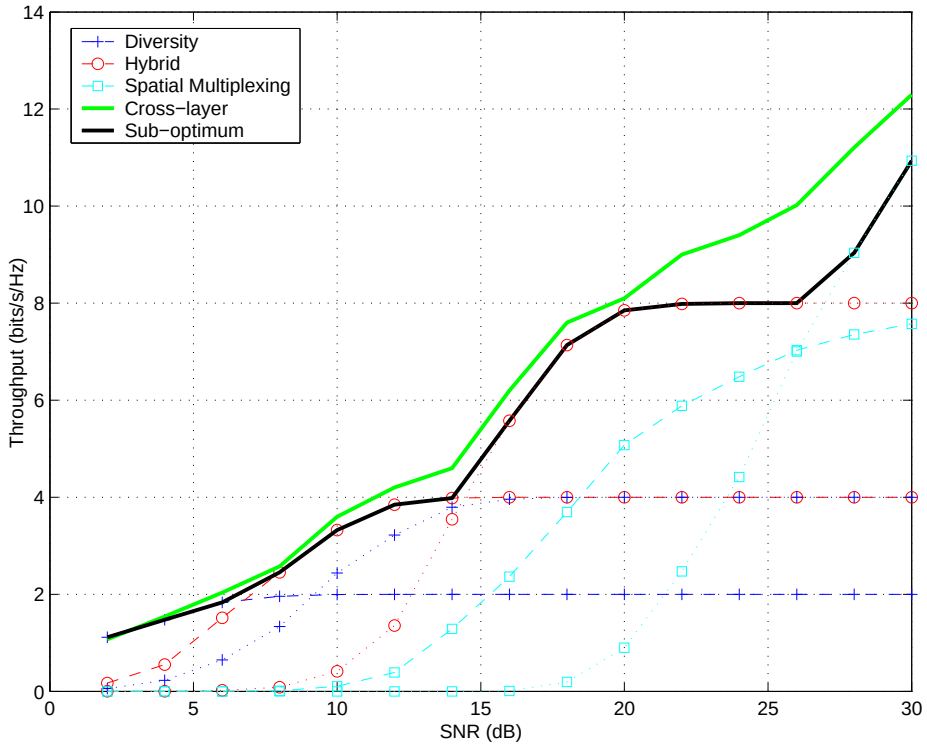


Figure 3.18: Link layer throughput vs. average SNR for the different transmission schemes. (Solid lines: switching approaches, dashed lines: QPSK, dotted lines: 16-QAM).

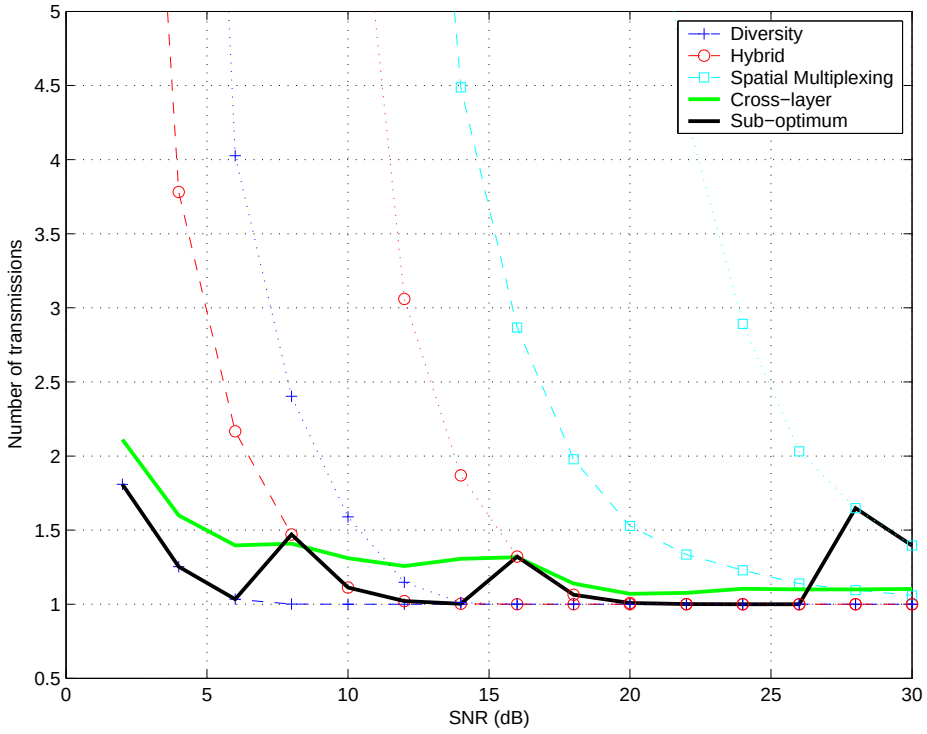


Figure 3.19: Average number of transmissions vs. average SNR for the different transmission schemes. (Solid lines: switching approaches, dashed lines: QPSK, dotted lines: 16-QAM).

3.8 Chapter Summary and Conclusions

In this chapter, a cross-layer approach to transmit antenna selection on the basis of a throughput-based criterion has been presented. For reference purposes, a conventional PHY-based antenna selection scheme based on maximizing mutual information has been considered. An enhanced version aimed at jointly performing antenna selection *and* rate adaptation has been derived for the CL approach in order to exploit better the available modulation schemes. In terms of throughput, it has been shown that the cross-layer approach outperforms the PHY-layer one for the whole range of signal-to-noise ratios with gains of up 3-4 dB in average SNR. Substantial benefits result in terms of transmission delay too. Moreover, it has been observed that, unless transmit antenna selection is conducted in a CL fashion, little or no improvement in terms of number of packet retransmissions can be achieved (due to error propagation in the detection procedure associated with V-BLAST).

Concerning the *full* and *restricted* antenna selection mechanisms, it has been found that in the low SNR region the FS scheme clearly outperforms RS. Conversely, the RS scheme performs better for high SNR ratios because of the additional diversity gain provided by a higher number of antenna elements. Further, a reduced-complexity version of the proposed CL antenna selection algorithms on the basis of the block-matrix inversion lemma has been derived. Computational complexity in terms of number of operation per time unit has been analyzed in detail.

Finally, the problem of switching between spatial diversity and multiplexing has been addressed from a cross-layer point of view, as well. Regarding system performance, it has been shown that both the fast adaptive and sub-optimum approaches exhibit superior performance in comparison with the original MIMO techniques separately. Particularly, a considerably gain is obtained with the fast adaptive cross-layer technique since instantaneous variations of the scenario conditions can be exploited.

3.A Appendix: Detailed Complexity Analysis of the AS Procedures

First, we will analyze the computational burden resulting from the *full-complexity* approach. For an arbitrary matrix on level m in the trellis, the following computations are required:

- Determine matrix $\mathbf{G}_m = \mathbf{H}_m^{\#}$. That can be accomplished via SVD due to its rich stability in a fixed-point implementation. Determining matrices $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{\Sigma}$ takes $9m^3 + 8Nm^2 + 4N^2m$ products and approximately the same number of additions [Gol96]. Also, $Nm^2 + Nm + m$ products and $Nm^2 - Nm$ additions are required for the matrix product $\mathbf{H}_m^{\#} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H$.
- Compute the norms for every row vector in $\mathbf{G}_m$, $\|\mathbf{w}_k\|^2$. For that purpose, mN products and $m(N-1)$ additions are needed.

Taking into account that the whole set of M levels in the trellis have to be explored, each featuring $\binom{M}{m}$ matrices $\mathbf{G}_m$, the overall complexity amounts to the number of FLOPS collected in Eq. (3.11). The terms outside the summation account for the computations involved in data detection steps (Eq. (3.6) and Eq. (3.8) in Table 3.1) where a *worst* case scenario consisting in all M transmit antennas being active is considered. Under those circumstances, a total of $2NM - N$ products and $2NM - M - N$ additions would be required. Therefore, the above expressions should be regarded as an upper bound for complexity.

Next, computational complexity for the *bottom-up* recursive approach will be analyzed:

- First step consists in computing $\left(\mathbf{h}_1^H\mathbf{h}_1\right)^{-1}$ for the M elements (matrices) on the first level. To do that, $NM + M$ products and $NM - M$ additions are required.
- Each transition from one matrix on level m to one matrix on level $m+1$, involves $3m^2 + Nm + 2m + N + 1$ products and $3m^2 + Nm - 2m + N - 1$ additions.

By considering all matrices in the trellis, the total number of required operations are collected in Eq. (3.12). Unlike in the *full complexity* approach, the vector set, $\mathbf{w}_k$ $k = 1, ..., m_{opt}$ is not available yet after throughput scoring. Actually, matrices $\left(\mathbf{H}_m^H\mathbf{H}_m\right)^{-1}$constitute the only output of the recursive procedure whereas $\mathbf{w}_m$ is given by one specific row in $\mathbf{G}_m = \left(\mathbf{H}_m^H\mathbf{H}_m\right)^{-1}\mathbf{H}_m^H$. In order to compute vector $\mathbf{w}_k$, Nm products and $Nm - N$ additions are needed. Taking a summation over $m = 1, ..., M$, that is, considering the worst case scenario, and adding up operations resulting from the data detection process itself, terms outside the summations in Eq. (3.12) can be found.

Similar considerations apply for the *top-down* approach, except for the following:

- First step consists now in computing $\left(\mathbf{H}_M^H\mathbf{H}_M\right)^{-1}$ on the top level. To do that, a single computationally-efficient *bottom-up* recursion will be considered (see Fig. 3.15) where the number of products and additions are $M^3-\frac{1}{2}M^2+\frac{1}{2}M+\frac{1}{2}NM^2+\frac{1}{2}NM$ and $M^3-\frac{5}{2}M^2+\frac{1}{2}M+\frac{1}{2}NM^2+\frac{1}{2}NM$, respectively.
- Each transition from one matrix on level $m+1$ to one matrix on level m involves $2m^2+1$ products and m^2 additions.

In summary, the resulting number of operations turns out to be that of Eq. (3.13).

3.B Appendix: Proof of Eq. (3.15)

By rewriting Eq. (3.14) as follows:

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{H}_{i,k_1} & \mathbf{H}_{i,k_2} \\ -\mathbf{H}_{i,k_2}^* & \mathbf{H}_{i,k_1}^* \end{bmatrix} \tag{3.24}$$

with matrices $\mathbf{H}_{i,k_1}$ and $\mathbf{H}_{i,k_2}$ defined as:

$$\mathbf{H}_{i,k_1} = \begin{bmatrix} h_{i_1} & h_{i_2} \\ -h_{i_2}^* & h_{i_1}^* \end{bmatrix} \quad \text{and} \quad \mathbf{H}_{i,k_2} = \begin{bmatrix} h_{i_3} & h_{i_4} \\ -h_{i_4}^* & h_{i_3}^* \end{bmatrix}$$

we can represent $\mathbf{H}_i^H\mathbf{H}_i$ as a block product of matrices:

$$\begin{aligned}\mathbf{H}_i^H\mathbf{H}_i &= \begin{bmatrix} \mathbf{H}_{i,k_1}^H & -\mathbf{H}_{i,k_2}^T \\ \mathbf{H}_{i,k_2}^H & \mathbf{H}_{i,k_1}^T \end{bmatrix}\begin{bmatrix} \mathbf{H}_{i,k_1} & \mathbf{H}_{i,k_2} \\ -\mathbf{H}_{i,k_2}^* & \mathbf{H}_{i,k_1}^* \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_1}+\mathbf{H}_{i,k_2}^T\mathbf{H}_{i,k_2}^* & \mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}-\mathbf{H}_{i,k_2}^T\mathbf{H}_{i,k_1}^* \\ \mathbf{H}_{i,k_2}^H\mathbf{H}_{i,k_1}-\mathbf{H}_{i,k_1}^T\mathbf{H}_{i,k_2}^* & \mathbf{H}_{i,k_1}^T\mathbf{H}_{i,k_1}^*+\mathbf{H}_{i,k_2}^H\mathbf{H}_{i,k_2} \end{bmatrix}\end{aligned}$$

The matrices in the diagonal of the above equation are identical and can be written as:

$$\begin{aligned}\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_1}+\mathbf{H}_{i,k_2}^T\mathbf{H}_{i,k_2}^* &= \mathbf{H}_{i,k_1}^T\mathbf{H}_{i,k_1}^*+\mathbf{H}_{i,k_2}^H\mathbf{H}_{i,k_2} \\ &= \left(|h_{i_1}|^2+|h_{i_2}|^2\right)\mathbf{I}_2+\left(|h_{i_3}|^2+|h_{i_4}|^2\right)\mathbf{I}_2 = h_i^2\mathbf{I}_2\end{aligned}$$

where $h_i^2 = |h_{i_1}|^2+|h_{i_2}|^2+|h_{i_3}|^2+|h_{i_4}|^2$ stands for the channel gain.

As for the matrices outside the diagonal, the first of them can be expressed as:

$$\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}-\mathbf{H}_{i,k_2}^T\mathbf{H}_{i,k_1}^* = \mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}-\left(\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}\right)^T \tag{3.25}$$

By identifying the elements of matrix $\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}$:

$$\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2} = \begin{bmatrix} h_{i_1}^* & -h_{i_2} \\ h_{i_2}^* & h_{i_1} \end{bmatrix}\begin{bmatrix} h_{i_3} & h_{i_4} \\ -h_{i_4}^* & h_{i_3}^* \end{bmatrix} = \begin{bmatrix} h_{i_1}^*h_{i_3}+h_{i_2}h_{i_4}^* & h_{i_1}^*h_{i_4}-h_{i_2}h_{i_3}^* \\ h_{i_2}^*h_{i_3}-h_{i_1}h_{i_4}^* & h_{i_2}^*h_{i_4}+h_{i_1}h_{i_3}^* \end{bmatrix} \tag{3.26}$$

one can compute the difference of matrices in Eq. (3.25) as follows:

$$\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}-\left(\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}\right)^T=\begin{bmatrix} h_{i_1}^*h_{i_3}+h_{i_2}h_{i_4}^* & h_{i_1}^*h_{i_4}-h_{i_2}h_{i_3}^* \\ h_{i_2}^*h_{i_3}-h_{i_1}h_{i_4}^* & h_{i_2}^*h_{i_4}+h_{i_1}h_{i_3}^* \end{bmatrix}-\begin{bmatrix} h_{i_1}^*h_{i_3}+h_{i_2}h_{i_4}^* & h_{i_2}^*h_{i_3}-h_{i_1}h_{i_4}^* \\ h_{i_1}^*h_{i_4}-h_{i_2}h_{i_3}^* & h_{i_2}^*h_{i_4}+h_{i_1}h_{i_3}^* \end{bmatrix}$$
$$=\begin{bmatrix} 0 & 2\mathrm{Re}\left\{h_{i_1}h_{i_4}^*-h_{i_2}h_{i_3}^*\right\} \\ -2\mathrm{Re}\left\{h_{i_1}h_{i_4}^*-h_{i_2}h_{i_3}^*\right\} & 0 \end{bmatrix} \tag{3.27}$$

The remaining matrix to be computed can be easily obtained by observing that all the elements of the matrix in Eq. (3.27) are real-valued and, as a consequence, the following equality holds:

$$\mathbf{H}_{i,k_2}^H\mathbf{H}_{i,k_1}-\mathbf{H}_{i,k_1}^T\mathbf{H}_{i,k_2}^*=\left(\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}-\mathbf{H}_{i,k_2}^T\mathbf{H}_{i,k_1}^*\right)^H=\left(\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}-\mathbf{H}_{i,k_2}^T\mathbf{H}_{i,k_1}^*\right)^T$$

Finally, we can collect all the previous results and verify the quasi-orthogonal behavior of $\mathbf{H}_i$, that is:

$$\mathbf{H}_i^H\mathbf{H}_i=h_i^2\begin{bmatrix} \mathbf{I}_2 & X_i\mathbf{J} \\ -X_i\mathbf{J} & \mathbf{I}_2 \end{bmatrix}$$

where $\mathbf{J}$ is the 2×2 matrix $\mathbf{J}=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and X_i is a real-valued random variable defined as $X_i=2\mathrm{Re}\left\{h_{i_1}h_{i_4}^*-h_{i_2}h_{i_3}^*\right\}/h_i^2$.

3.C Appendix: Computation of ZF Vector Norms

In this appendix, we are devoted to show the behavior of the norms associated to the ZF receiver vectors, that is the rows of matrix:

$$\mathbf{W}=\left[\sum_{i=1}^N\mathbf{H}_i^H\mathbf{H}_i\right]^{-1}\sum_{i=1}^N\mathbf{H}_i^H \tag{3.28}$$

In particular, the analysis is conducted for both the Q-OSTBC and D-STTD approaches:

3.C.1 Q-OSTBC

First, the inverse of matrix $\sum_{i=1}^N\mathbf{H}_i^H\mathbf{H}_i$ is computed. In the previous appendix we proved that $\mathbf{H}_i^H\mathbf{H}_i$ has a quasi-diagonal structure and can be expressed as:

$$\mathbf{H}_i^H\mathbf{H}_i=h_i^2\begin{bmatrix} \mathbf{I}_2 & X_i\mathbf{J} \\ -X_i\mathbf{J} & \mathbf{I}_2 \end{bmatrix}$$

Then, by takin into consideration the contributions of all the receive antennas we have that:

$$\sum_{i=1}^N\mathbf{H}_i^H\mathbf{H}_i=h_{tot}^2\begin{bmatrix} \mathbf{I}_2 & X_{tot}\mathbf{J} \\ -X_{tot}\mathbf{J} & \mathbf{I}_2 \end{bmatrix}$$

with $h_{tot}^2 = \sum_{i=1}^{N} h_i^2$, and $X_t = \left(\sum_{i=1}^{N} X_i h_i^2\right) / h_{tot}^2$.

By invoking the block matrix inversion lemma [Mag02], the inverse of the above matrix expression can be easily obtained:

$$\begin{aligned}\left[\sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{H}_i\right]^{-1} &= \frac{1}{h_{tot}^2} \begin{bmatrix} \mathbf{I}_2 & X_{tot}\mathbf{J} \\ -X_{tot}\mathbf{J} & \mathbf{I}_2 \end{bmatrix}^{-1} = \frac{1}{h_{tot}^2} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} \\ &= \frac{1}{h_{tot}^2} \begin{bmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{D}^{-1} \end{bmatrix}\end{aligned} \tag{3.29}$$

where:

$$\begin{aligned}\mathbf{A}_{11} &= \mathbf{I}_2 \\ \mathbf{A}_{12} &= X_{tot}\mathbf{J} \\ \mathbf{A}_{21} &= -X_{tot}\mathbf{J} \\ \mathbf{A}_{22} &= \mathbf{I}_2 \\ \mathbf{D}^{-1} &= \left[\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{I}_2\right]^{-1} = \left[\mathbf{I}_2 + X_{tot}\mathbf{J}\mathbf{I}_2 X_{tot}\mathbf{J}\right]^{-1} \\ &= \left[(1 - X_{tot}^2)\mathbf{I}_2\right]^{-1} = \frac{1}{1 - X_{tot}^2}\mathbf{I}_2 \\ \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} &= \mathbf{I}_2 + \mathbf{I}_2 X_{tot}\mathbf{J}\frac{1}{1 - X_{tot}^2}\mathbf{I}_2 \left(-X_{tot}\mathbf{J}\right)\mathbf{I}_2 \\ &= \mathbf{I}_2\left(1 + \frac{X_{tot}^2}{1 - X_{tot}^2}\right) = \frac{1}{1 - X_{tot}^2}\mathbf{I}_2 \\ -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1} &= -\mathbf{I}_2 X_{tot}\mathbf{J}\frac{1}{1 - X_{tot}^2}\mathbf{I}_2 = -\frac{X_{tot}}{1 - X_{tot}^2}\mathbf{J} \\ -\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} &= -\frac{1}{1 - X_{tot}^2}\mathbf{I}_2 \left(-X_{tot}\mathbf{J}\right)\mathbf{I}_2 = \frac{X_{tot}}{1 - X_{tot}^2}\mathbf{J}\end{aligned}$$

Then, one can express the matrix inversion in Eq. (3.29) as:

$$\left[\sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{H}_i\right]^{-1} = \frac{1}{h_{tot}^2\left(1 - X_{tot}^2\right)} \begin{bmatrix} \mathbf{I}_2 & -X_{tot}\mathbf{J} \\ X_{tot}\mathbf{J} & \mathbf{I}_2 \end{bmatrix} \tag{3.30}$$

Now, by plugging Eq. (3.30) into Eq. (3.28) and bearing in mind the block representation of matrix $\mathbf{H}_i$ defined in Eq. (3.24), we can rewrite $\mathbf{W}$ as a block product of matrices:

$$\begin{aligned}\mathbf{W} &= \left[\sum_{i=1}^{N} \mathbf{H}_i^H \mathbf{H}_i\right]^{-1} \sum_{i=1}^{N} \mathbf{H}_i^H = \frac{1}{h_{tot}^2\left(1 - X_{tot}^2\right)} \begin{bmatrix} \mathbf{I}_2 & -X_{tot}\mathbf{J} \\ X_{tot}\mathbf{J} & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{H}_{k_1}'^H & -\mathbf{H}_{k_2}'^T \\ \mathbf{H}_{k_2}'^H & \mathbf{H}_{k_1}'^T \end{bmatrix} \\ &= \frac{1}{h_{tot}^2\left(1 - X_{tot}^2\right)} \begin{bmatrix} \mathbf{H}_{k_1}'^H - X_{tot}\mathbf{J}\mathbf{H}_{k_2}'^H & -\mathbf{H}_{k_2}'^T - X_{tot}\mathbf{J}\mathbf{H}_{k_1}'^T \\ \mathbf{H}_{k_2}'^H + X_{tot}\mathbf{J}\mathbf{H}_{k_1}'^H & \mathbf{H}_{k_1}'^T - X_{tot}\mathbf{J}\mathbf{H}_{k_2}'^T \end{bmatrix} \\ &= \frac{1}{h_{tot}^2\left(1 - X_{tot}^2\right)} \begin{bmatrix} \mathbf{G}_{k_1} & \mathbf{G}_{k_2} \\ -\mathbf{G}_{k_2}^* & \mathbf{G}_{k_1}^* \end{bmatrix}\end{aligned} \tag{3.31}$$

with $\mathbf{H}'_{k_j} = \sum_{i=1}^{N} \mathbf{H}_{i,k_j}$. Notice that, the symmetrical structure observed in the channel matrices $\mathbf{H}_i$ appears when the ZF linear receiver is constructed. The same effect is observed for elements in matrices $\mathbf{G}_{k_1}$ and $\mathbf{G}_{k_2}$:

$$\begin{aligned}\mathbf{G}_{k_1} &= \mathbf{H}'^{H}_{k_1} - X_{tot}\mathbf{J}\mathbf{H}'^{H}_{k_2} = \begin{bmatrix} h'^{*}_1 & -h'_2 \\ h'^{*}_2 & h'_1 \end{bmatrix} - X_{tot}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} h'^{*}_3 & -h'_4 \\ h'^{*}_4 & h'_3 \end{bmatrix} \\ &= \begin{bmatrix} h'^{*}_1 & -h'_2 \\ h'^{*}_2 & h'_1 \end{bmatrix} - X_{tot}\begin{bmatrix} h'^{*}_4 & h'_3 \\ -h'^{*}_3 & h'_4 \end{bmatrix} = \begin{bmatrix} h'^{*}_1 - Xh'^{*}_4 & -h'_2 - X_{tot}h'_3 \\ h'^{*}_2 + X_{tot}h'^{*}_3 & h'_1 - X_{tot}h'_4 \end{bmatrix} \\ &= \begin{bmatrix} g_{11} & g_{12} \\ -g^{*}_{12} & g^{*}_{11} \end{bmatrix}\end{aligned}$$

$$\begin{aligned}\mathbf{G}_{k_2} &= \mathbf{H}'^{H}_{k_2} + X_{tot}\mathbf{J}\mathbf{H}'^{H}_{k_1} = \begin{bmatrix} h'^{*}_3 & -h'_4 \\ h'^{*}_4 & h'_3 \end{bmatrix} + X_{tot}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} h'^{*}_1 & -h'_2 \\ h'^{*}_2 & h'_1 \end{bmatrix} \\ &= \begin{bmatrix} h'^{*}_3 & -h'_4 \\ h'^{*}_4 & h'_3 \end{bmatrix} + X_{tot}\begin{bmatrix} h'^{*}_2 & h'_1 \\ -h'^{*}_1 & h'_2 \end{bmatrix} = \begin{bmatrix} h'^{*}_3 + X_{tot}h'^{*}_2 & -h'_4 + X_{tot}h'_1 \\ h'^{*}_4 - X_{tot}h'^{*}_1 & h'_3 + X_{tot}h'_2 \end{bmatrix} \\ &= \begin{bmatrix} g_{21} & g_{22} \\ -g^{*}_{22} & g^{*}_{21} \end{bmatrix}\end{aligned}$$

where $h'_j = \sum_{i=1}^{N} h_{i_j}$ stand for the elements of matrix $\mathbf{H}'_{k_j}$. Indeed, due to this structure we have that matrices $\mathbf{G}_{k_1}$ and $\mathbf{G}_{k_2}$ are orthogonal:

$$\mathbf{G}^{H}_{k_1}\mathbf{G}_{k_1} = \mathbf{G}_{k_1}\mathbf{G}^{H}_{k_1} = \begin{bmatrix} |g_{11}|^2 + |g_{12}|^2 & 0 \\ 0 & |g_{11}|^2 + |g_{12}|^2 \end{bmatrix} \tag{3.32}$$

$$\mathbf{G}^{H}_{k_2}\mathbf{G}_{k_2} = \mathbf{G}_{k_2}\mathbf{G}^{H}_{k_2} = \begin{bmatrix} |g_{21}|^2 + |g_{22}|^2 & 0 \\ 0 & |g_{21}|^2 + |g_{22}|^2 \end{bmatrix} \tag{3.33}$$

where the following equalities hold:

$$\begin{aligned}|g_{11}|^2 + |g_{12}|^2 &= |h'_1|^2 + |h'_2|^2 + X^2_{tot}\left(|h'_3|^2 + |h'_4|^2\right) - 2X_{tot}\mathrm{Re}\left\{h'_1h'^{*}_4 - h'_2h'^{*}_3\right\} \\ |g_{21}|^2 + |g_{22}|^2 &= |h'_3|^2 + |h'_4|^2 + X^2_{tot}\left(|h'_1|^2 + |h'_2|^2\right) - 2X_{tot}\mathrm{Re}\left\{h'_1h'^{*}_4 - h'_2h'^{*}_3\right\}\end{aligned}$$

Finally, in order to obtain the norms associated to the row vectors of the matrix $\mathbf{W}$, i.e.:

$$\|w_k\| = \sqrt{w^H_k w_k} = \sqrt{[\mathbf{W}\mathbf{W}^H]_{k,k}}$$

we use the block representation given by Eq. (3.31) as follows:

$$\begin{aligned}\mathbf{W}\mathbf{W}^H &= \frac{1}{\left[h^2_{tot}\left(1 - X^2_{tot}\right)\right]^2}\begin{bmatrix} \mathbf{G}_{k_1} & \mathbf{G}_{k_2} \\ -\mathbf{G}^{*}_{k_2} & \mathbf{G}^{*}_{k_1} \end{bmatrix}\begin{bmatrix} \mathbf{G}^{H}_{k_1} & -\mathbf{G}^{T}_{k_2} \\ \mathbf{G}^{H}_{k_2} & \mathbf{G}^{T}_{k_1} \end{bmatrix} \\ &= \frac{1}{\left[h^2_{tot}\left(1 - X^2_{tot}\right)\right]^2}\begin{bmatrix} \mathbf{G}_{k_1}\mathbf{G}^{H}_{k_1} + \mathbf{G}_{k_2}\mathbf{G}^{H}_{k_2} & -\mathbf{G}_{k_1}\mathbf{G}^{T}_{k_2} + \mathbf{G}_{k_2}\mathbf{G}^{T}_{k_1} \\ -\mathbf{G}^{*}_{k_2}\mathbf{G}^{H}_{k_1} + \mathbf{G}^{*}_{k_1}\mathbf{G}^{H}_{k_2} & \mathbf{G}^{*}_{k_1}\mathbf{G}^{T}_{k_1} + \mathbf{G}^{*}_{k_2}\mathbf{G}^{T}_{k_2} \end{bmatrix}\end{aligned}$$

Notice that:

$$\mathbf{G}_{k_1}\mathbf{G}^{H}_{k_1} + \mathbf{G}_{k_2}\mathbf{G}^{H}_{k_2} = \left(\mathbf{G}^{*}_{k_1}\mathbf{G}^{T}_{k_1} + \mathbf{G}^{*}_{k_2}\mathbf{G}^{T}_{k_2}\right)^{*} \tag{3.34}$$

Then, all the matrices in the diagonal of $\mathbf{W}\mathbf{W}^H$ are the same matrix. This is because all the elements of the matrices $\mathbf{G}_{k_1}\mathbf{G}_{k_1}^H$ and $\mathbf{G}_{k_2}\mathbf{G}_{k_2}^H$ are real-valued. Furthermore, by inserting Eq. (3.32) and Eq. (3.33) into Eq. (3.34), one can notice that all the ZF vector norms are equal since all the elements $[\mathbf{W}\mathbf{W}^H]_{k,k}$ can be written according to the following expression:

$$\begin{aligned}\left[\mathbf{W}\mathbf{W}^H\right]_{k,k} =& \frac{1}{\left[h_{tot}^2\left(1-X_{tot}^2\right)\right]^2}\left(|g_{11}|^2+|g_{12}|^2+|g_{21}|^2+|g_{22}|^2\right)\\ =& \frac{1}{\left[h_{tot}^2\left(1-X_{tot}^2\right)\right]^2}\Big[\left(|h'_1|^2+|h'_2|^2+|h'_3|^2+|h'_4|^2\right)\left(1+X_{tot}^2\right)\\ & -4X_{tot}\mathrm{Re}\left\{h'_1h'^*_4-h'_2h'^*_3\right\}\Big] \qquad \text{for } k=1,...,M\end{aligned}$$

3.C.2 D-SSTD

A similar procedure to the conducted in the previous Appendix can be carried out. First, we represent Eq. (3.20) as a block matrix:

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{H}_{i,k_1} & \mathbf{H}_{i,k_2} \end{bmatrix}$$

By doing so, we can express $\sum_{i=1}^{N}\mathbf{H}_i^H\mathbf{H}_i$ as:

$$\sum_{i=1}^{N}\mathbf{H}_i^H\mathbf{H}_i = \begin{bmatrix} \alpha\mathbf{I}_2 & \mathbf{B} \\ \mathbf{B}^H & \beta\mathbf{I}_2 \end{bmatrix}$$

with $\mathbf{B}=\sum_{i=1}^{N}\mathbf{H}_{i,k_1}^H\mathbf{H}_{i,k_2}$, $\alpha=\sum_{i=1}^{N}|h_{i_1}|^2+|h_{i_2}|^2$ and $\beta=\sum_{i=1}^{N}|h_{i_3}|^2+|h_{i_4}|^2$.

Then, the block matrix inversion lemma can be invoked again to obtain the inverse of $\sum_{i=1}^{N}\mathbf{H}_i^H\mathbf{H}_i$:

$$\begin{aligned}\left[\sum_{i=1}^{N}\mathbf{H}_i^H\mathbf{H}_i\right]^{-1} &= \begin{bmatrix} \alpha\mathbf{I}_2 & \mathbf{B} \\ \mathbf{B}^H & \beta\mathbf{I}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1}\\ &= \begin{bmatrix} \mathbf{A}_{11}^{-1}+\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{D}^{-1} \end{bmatrix}\end{aligned}$$

where, in this case, the following equalities are used:

$$\begin{aligned}\mathbf{A}_{11} &= \alpha\mathbf{I}_2\\ \mathbf{A}_{12} &= \mathbf{B}\\ \mathbf{A}_{21} &= \mathbf{B}^H\\ \mathbf{A}_{22} &= \beta\mathbf{I}_2\\ \mathbf{D}^{-1} &= \left[\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{I}_2\right]^{-1} = \left[\beta\mathbf{I}_2-\alpha^{-1}\mathbf{B}^H\mathbf{B}\right]^{-1}\end{aligned}$$

By taking into consideration expression Eq. (3.26), $\mathbf{B^H B}$ can be easily obtained as:

$$
\begin{aligned}
&\mathbf{B}^H\mathbf{B} \\
&= \left(\sum_{i=1}^{N} \mathbf{H}_{i,i,k_1}^H \mathbf{H}_{i,k_2}\right)^H \left(\sum_{i=1}^{N} \mathbf{H}_{i,i,k_1}^H \mathbf{H}_{i,k_2}\right) \\
&= \left(\sum_{i=1}^{N} \begin{bmatrix} h_{i_1}^* h_{i_3} + h_{i_2} h_{i_4}^* & h_{i_1}^* h_{i_4} - h_{i_2} h_{i_3}^* \\ h_{i_2}^* h_{i_3} - h_{i_1} h_{i_4}^* & h_{i_2}^* h_{i_4} + h_{i_1} h_{i_3}^* \end{bmatrix}\right)^H \sum_{i=1}^{N} \begin{bmatrix} h_{i_1}^* h_{i_3} + h_{i_2} h_{i_4}^* & h_{i_1}^* h_{i_4} - h_{i_2} h_{i_3}^* \\ h_{i_2}^* h_{i_3} - h_{i_1} h_{i_4}^* & h_{i_2}^* h_{i_4} + h_{i_1} h_{i_3}^* \end{bmatrix} \\
&= \left(\sum_{i=1}^{N} \begin{bmatrix} b_{i,1} & b_{i,2} \\ -b_{i,2}^* & b_{i,1}^* \end{bmatrix}\right)^H \sum_{i=1}^{N} \begin{bmatrix} b_{i,1} & b_{i,2} \\ -b_{i,2}^* & b_{i,1}^* \end{bmatrix} \\
&= \begin{bmatrix} b_1' & b_2' \\ -b_2'^* & b_1'^* \end{bmatrix}^H \begin{bmatrix} b_1' & b_2' \\ -b_2'^* & b_1'^* \end{bmatrix} = \begin{bmatrix} |b_1'|^2 + |b_2'|^2 & 0 \\ 0 & |b_1'|^2 + |b_2'|^2 \end{bmatrix}
\end{aligned} \tag{3.35}
$$

with $b_{i,1} = h_{i_1}^* h_{i_3} + h_{i_2} h_{i_4}^*$, $b_{i_2} = h_{i_1}^* h_{i_4} - h_{i_2} h_{i_3}^*$ and $b_j' = \sum_{i=1}^{N} b_{i,j}$. As a consequence, we can obtain the following expressions:

$$
\begin{aligned}
\mathbf{D}^{-1} &= \left[\left(\beta - \frac{|b_1'|^2 + |b_2'|^2}{\alpha}\right)\mathbf{I}_2\right]^{-1} = \frac{\alpha}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\mathbf{I}_2 \\
\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} &= \alpha^{-1}\mathbf{I}_2 + \alpha^{-1}\mathbf{I}_2\mathbf{B}\left(\frac{\alpha}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\right)\mathbf{I}_2\mathbf{B}^H\alpha^{-1}\mathbf{I}_2 \\
&= \alpha^{-1}\mathbf{I}_2 + \alpha^{-1}\left(\frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\right)\mathbf{B}\mathbf{B}^H \\
&= \alpha^{-1}\mathbf{I}_2 + \alpha^{-1}\left(\frac{|b_1'|^2 + |b_2'|^2}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\right)\mathbf{I}_2 = \frac{\beta}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\mathbf{I}_2
\end{aligned}
$$

where in the latter expression we have also used Eq. (3.35) given that $\mathbf{B}\mathbf{B}^H = \mathbf{B}^H\mathbf{B}$. In addition, we also have that:

$$
\begin{aligned}
-\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1} &= -\alpha^{-1}\mathbf{I}_2\mathbf{B}\frac{\alpha}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\mathbf{I}_2 = -\frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\mathbf{B} \\
-\mathbf{D}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} &= -\left(\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{D}^{-1}\right)^H = -\frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\mathbf{B}^H
\end{aligned}
$$

By collecting the results obtained above, the inverse of $\sum_{i=1}^{N} \mathbf{H}_i^H\mathbf{H}_i$ can be written as:

$$
\left[\sum_{i=1}^{N} \mathbf{H}_i^H\mathbf{H}_i\right]^{-1} = \frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\begin{bmatrix} \beta\mathbf{I}_2 & -\mathbf{B} \\ -\mathbf{B}^H & \alpha\mathbf{I}_2 \end{bmatrix} \tag{3.36}
$$

Now, we can write the ZF receive filter by plugging Eq. (3.36) into Eq. (3.28):

$$
\begin{aligned}
\mathbf{W} &= \left[\sum_{i=1}^{N} \mathbf{H}_i^H\mathbf{H}_i\right]^{-1} \sum_{i=1}^{N} \mathbf{H}_i^H = \frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\begin{bmatrix} \beta\mathbf{I}_2 & -\mathbf{B} \\ -\mathbf{B}^H & \alpha\mathbf{I}_2 \end{bmatrix}\begin{bmatrix} \mathbf{H}_{k_1}'^H \\ \mathbf{H}_{k_2}'^H \end{bmatrix} \\
&= \frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\begin{bmatrix} \beta\mathbf{H}_{k_1}'^H - \mathbf{B}\mathbf{H}_{k_2}'^H \\ \alpha\mathbf{H}_{k_2}'^H - \mathbf{B}^H\mathbf{H}_{k_1}'^H \end{bmatrix} = \frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2}\begin{bmatrix} \mathbf{G}_{k_1} \\ \mathbf{G}_{k_2} \end{bmatrix}
\end{aligned}
$$

In this case, the symmetrical structure is observed in matrices $\mathbf{G}_{k_1}$ and $\mathbf{G}_{k_2}$, as well. More precisely, by bearing in mind that α and β are real numbers we have that:

$$\begin{aligned}
\mathbf{G}_{k_1} = \beta\mathbf{H}_{k_1}'^H - \mathbf{B}\mathbf{H}_{k_2}'^H &= \beta \begin{bmatrix} h_1'^* & -h_2' \\ h_2'^* & h_1' \end{bmatrix} - \begin{bmatrix} b_1' & b_2' \\ -b_2'^* & b_1'^* \end{bmatrix} \begin{bmatrix} h_3'^* & -h_4' \\ h_4'^* & h_3' \end{bmatrix} \\
&= \beta \begin{bmatrix} h_1'^* & -h_2' \\ h_2'^* & h_1' \end{bmatrix} - \begin{bmatrix} b_1'h_3'^* + b_2'h_4'^* & -b_1'h_4' + b_2'h_3' \\ -(-b_1'h_4' + b_2'h_3')^* & (b_1'h_3'^* + b_2'h_4'^*)^* \end{bmatrix} \\
&= \begin{bmatrix} \beta h_1'^* - b_1'h_3'^* - b_2'h_4'^* & -\beta h_2' + b_1'h_4' - b_2'h_3' \\ -(-\beta h_2' + b_1'h_4' - b_2'h_3')^* & (\beta h_1'^* - b_1'h_3'^* - b_2'h_4'^*)^* \end{bmatrix} \\
&= \begin{bmatrix} g_{11} & g_{12} \\ -g_{12}^* & g_{11}^* \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{G}_{k_2} = \beta\mathbf{H}_{k_1}'^H - \mathbf{B}\mathbf{H}_{k_2}'^H &= \alpha \begin{bmatrix} h_3'^* & -h_4' \\ h_4'^* & h_3' \end{bmatrix} - \begin{bmatrix} b_1'^* & -b_2' \\ b_2'^* & b_1' \end{bmatrix} \begin{bmatrix} h_1'^* & -h_2' \\ h_2'^* & h_1' \end{bmatrix} \\
&= \alpha \begin{bmatrix} h_3'^* & -h_4' \\ h_4'^* & h_3' \end{bmatrix} - \begin{bmatrix} b_1'^*h_1'^* - b_2'h_2'^* & -b_1'^*h_2' - b_2'h_1' \\ -(-b_1'^*h_2' - b_2'h_1')^* & (b_1'^*h_1'^* - b_2'h_2'^*)^* \end{bmatrix} \\
&= \begin{bmatrix} \alpha h_3'^* - b_1'^*h_1'^* + b_2'h_2'^* & -\alpha h_4' + b_1'^*h_2 + b_2'h_1' \\ -(-\alpha h_4' + b_1'^*h_2 + b_2'h_1')^* & (\alpha h_3'^* - b_1'^*h_1'^* + b_2'h_2'^*) \end{bmatrix} \\
&= \begin{bmatrix} g_{21} & g_{22} \\ -g_{22}^* & g_{21}^* \end{bmatrix}
\end{aligned}$$

Finally, given that matrix $\mathbf{W}$ presents the following structure:

$$\mathbf{W} = \frac{1}{\alpha\beta - |b_1'|^2 - |b_2'|^2} \begin{bmatrix} g_{11} & g_{12} \\ -g_{12}^* & g_{11}^* \\ g_{21} & g_{22} \\ -g_{22}^* & g_{21}^* \end{bmatrix}$$

it can be easily shown that the diagonal of the $\mathbf{W}\mathbf{W}^H$ can be written as follows:

$$\mathrm{diag}(\mathbf{W}\mathbf{W}^H) = \frac{1}{\left[\alpha\beta - |b_1'|^2 - |b_2'|^2\right]^2} \begin{bmatrix} |g_{11}|^2 + |g_{12}|^2 \\ |g_{11}|^2 + |g_{12}|^2 \\ |g_{21}|^2 + |g_{22}|^2 \\ |g_{21}|^2 + |g_{22}|^2 \end{bmatrix}$$

In other words, vector norms take the same value by pairs. This is to say:

$$\begin{aligned}
\|\mathbf{w}_1\|^2 = \|\mathbf{w}_2\|^2 &= \frac{1}{\left[\alpha\beta - |b_1'|^2 - |b_2'|^2\right]^2} \left(|g_{11}|^2 + |g_{12}|^2\right) \\
\|\mathbf{w}_3\|^2 = \|\mathbf{w}_4\|^2 &= \frac{1}{\left[\alpha\beta - |b_1'|^2 - |b_2'|^2\right]^2} \left(|g_{21}|^2 + |g_{22}|^2\right)
\end{aligned}$$

with:

$$\begin{aligned}
g_{11} &= \beta h_1'^* - b_1'h_3'^* - b_2'h_4'^* \\
g_{12} &= -\beta h_2' + b_1'h_4' - b_2'h_3' \\
g_{21} &= \alpha h_3'^* - b_1'^*h_1'^* + b_2'h_2'^* \\
g_{22} &= -\alpha h_4' + b_1'^*h_2 + b_2'h_1'
\end{aligned}$$

Chapter 4

Joint Exploitation of Spatial and Multi-user Diversity

In this chapter, we extend the study to a multi-user scenario where only one user is scheduled for transmission in each time-slot by means of an opportunistic scheduler. In such a context, antenna selection mechanisms are adopted with the aim of jointly exploiting spatial and multi-user diversity with a low-rate feedback channel. The analysis is conducted in several scenarios with different source of impairments in the channel state information available at the scheduler. In order to improve robustness against such impairments, schemes based on the combination of OSTBC with transmit antenna selection are also considered.

4.1 Introduction

As explained in Chapter 2, the average cell throughput of a SISO multi-user system can be maximized when in each slot the user with the best channel conditions is scheduled [Kno95, Tse97]. Such an effect is referred to as *multi-user diversity* and relies on the assumption that different users experience independent fading processes. On the other hand, transmit *spatial diversity* provides an effective means to combat fading and, thus, have a more reliable transmission. In particular, OSTBC schemes provide full diversity order while using low complexity receivers [Tar99, Ala98]. Both schemes, aimed at exploiting either spatial or multi-user diversity, have been proposed for packet data services in 3G wireless networks. For that reason, much attention has been recently paid to the cross-layer interaction between transmit spatial diversity (PHY layer) and multi-user diversity (MAC functionality at data link layer).

For instance, in e.g., [Goz03], [Jia04] and [Hoc04], the inclusion of OSTBC in multiuser schemes was analyzed. It was shown that, in a multi-user context, SISO schemes outperform OSTBC-based ones in terms of aggregated cell capacity. Certainly, spatial diversity helps reduce the probability of deep fades but, by averaging over different diversity branches, SNR peaks (those

that multi-user diversity can exploit) are suppressed as well. As a result, the resulting system capacity is lower. Similar conclusions were drawn for the case of Nakagami fading channels in [Che06].

Nonetheless, it was proven in [Lar04, Che06] that with perfect and complete CSI at the transmitter, spatial diversity can be efficiently exploited in a multi-user context via optimal transmit beamforming. To do so, though, a high-rate return channel is required since, in principle, all the channel gains (amplitude and phase) must be fed back to the scheduler. Alternatively, a transmit antenna selection scheme where transmit power is concentrated in the antenna with the largest channel gain was also proposed. For this second approach, only a limited amount of CSI must be conveyed to the transmitter and, consequently, a low-rate signalling channel suffices. However, gains resulting from spatial processing decrease as well.

Recently, several studies have been carried out in scenarios where the CSI available at the scheduler is subject to impairments. In [Ber03] and [Kob04], the authors analyzed the impact of *delays* in the feedback channel by means of system-level computer simulations and numerical integration, respectively. A comparative study between the use of OSTBC or optimal beamforming was carried out in [Ma05]. More precisely, the effect of delayed feedback channel was analyzed in a practical system where an adaptive modulation strategy subject to BER constraints is encompassed. Basically, all these works show that the increased robustness of OSTBC schemes against imperfect CSI provides significant capacity gains with respect to those of SISO (and optimal beamforming) approaches.

In this chapter, we also assess the existing trade-offs in the combined use of multi-user and transmit spatial diversity. To do so, we analyze and compare different transmission schemes based on SISO and OSTBC approaches with and without antenna selection capabilities. We start by considering an ideal scenario where *all* users permanently report their channel status to the centralized scheduler over a error- and delay-free feedback channel. After that, different impairments are introduced: bandwidth restrictions in the feedback channel and imperfect CSI available at the scheduler. In the former case, we adopt a selective-feedback scheduler [Ges03, Ges04, Yan04], which consists in exclusively letting users report CSI when channel quality exceeds a pre-defined threshold (i.e. when there is a chance for that user to be eventually scheduled). When the number of active users increases, their individual contribution to feedback channels should be necessarily decreased. Therefore, it is important to study what is the impact of such bandwidth restrictions on the different transmission schemes. As for the case where imperfect CSI is available, we do not restrict ourselves to a specific source of imperfections but, instead, we adopt a generic statistical approach to its modeling. In order to gain some insight, we then present two practical examples: delayed feedback channel and imperfect channel estimation.

As far as performance assessment is concerned, we conduct the analysis in terms of ergodic capacity measures and system throughput. More precisely, we obtain closed-form expressions

for the former case and accurate approximations for the latter one. To do so, we first derive an exact expression for the pdf of the post-scheduling signal-to-noise ratio associated with all the considered scenarios.

To conclude the chapter, we also assess spatial vs. multi-user trade-offs by using mean vs. standard deviation plots [Bar05], inspired by theory of modern *portfolio* [Mar52, Mar91]. By doing so, both the degree of robustness to short-term SNR fluctuations and its impact in terms of system performance can be easily quantified for the different transmission schemes.

In summary, the contributions of this chapter are the following:

- The analysis conducted in [Ges03, Ges04, Yan04] is extended to OSTBC and antenna selection mechanisms.
- A general statistical approach is derived to assess the impact of imperfect CSI on the scheduler.
- Closed-form expressions are obtained for the ergodic capacity and (tight) approximations are derived for the system throughput.
- A novel approach is adopted to illustrate the robustness of the different transmissions schemes against incomplete or imperfect CSI.

This chapter is organized as follows: in Section 4.2, the corresponding signal and system models are presented. A short description of the selected transmission schemes and the associated scheduling algorithms are provided as well. Closed-form expressions for the density functions (pdf and CDF) of the pre-scheduling SNRs are then derived in Section 4.3. Next, the different transmission schemes are analyzed and compared for a scenario with complete and non-degraded CSI at the scheduler in Section 4.4. After that, scenarios with restrictions in the feedback channel and degraded CSI are considered in Sections 4.5 and 4.6, respectively. In Section 4.7, the robustness of the different transmission schemes are illustrated by using mean vs. standard deviation plots. Finally, in Section 4.8, the summary and conclusions of this chapter are presented.

The results obtained in this chapter have been published in [Vic05a, Vic05c, Vic05d, Vic06d, Vic06e, Vic06f, Vic06g, Vic06h].

4.2 Signal and System Model

In this section, the multi-user scenario considered along this chapter is presented. First, the signal model is described. After that, the different transmission schemes are introduced and, finally, a description of the adopted scheduling algorithm is provided.

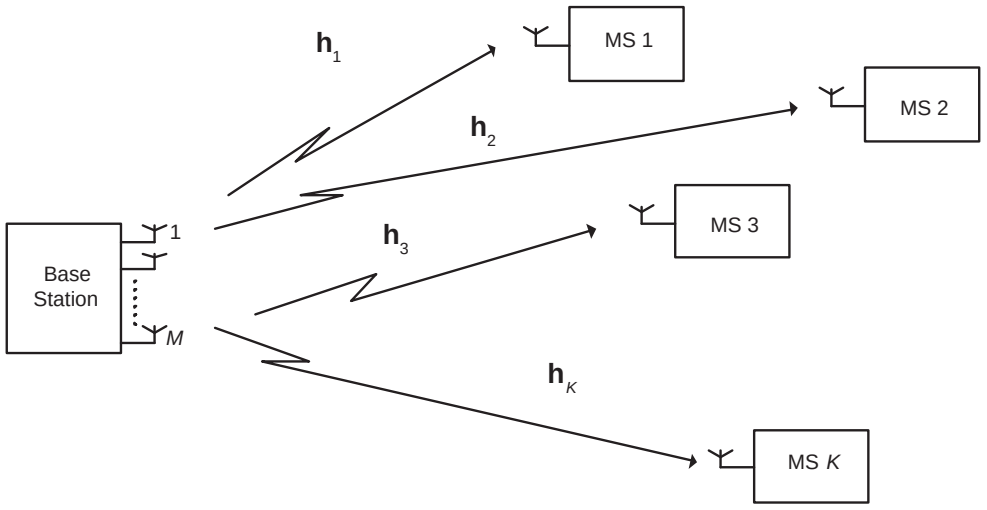


Figure 4.1: Block diagram of a multi-user communication system with a multi-antenna BS and single antenna MSs.

4.2.1 Signal Model

Consider the downlink of a cellular system with one BS equipped with multiple antennas (M), and K single-antenna MSs (see Fig. 4.1). The received signal at the k-th MS is given by:

$$r_k = \mathbf{h}_k^T \mathbf{s} + n_k \tag{4.1}$$

where the time index has been dropped for the ease of notation, $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ is the channel vector gain between the BS and the k-th terminal, for which each component is assumed to be independent and identically distributed, circularly symmetric Gaussian random variable with zero mean and user-dependent variance $\sigma_{h_k}^2$ ($\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \sigma_{h_k}^2 \mathbf{I}_M)$), $\mathbf{s} \in \mathbb{C}^{M\times 1}$ is the symbol vector broadcasted from the BS and $n_k \in \mathbb{C}$ denotes AWGN with zero mean and variance σ^2. The active users in the system are assumed to undergo independent Rayleigh fading processes and so does the signal being transmitted from different antennas in the BS. Further, we consider block fading, i.e, the channel response remains constant during one time-slot. We denote by $\gamma_k = \frac{P_t \|\mathbf{h}_k\|^2}{M\sigma^2}$ the *instantaneous* signal-to-noise ratio experienced by user k in a given time-slot and by $\bar{\gamma}_k = \frac{P_t \mathbb{E}[\|\mathbf{h}_k\|^2]}{M\sigma^2} = \frac{P_t \sigma_{h_k}^2}{\sigma^2}$ its long-term average SNR, with P_t standing for the total transmit power and where the near-far effect among users in the cell is represented by $\sigma_{h_k}^2$. Notice that the total transmitted power is constant and evenly distributed among transmit antennas. Concerning channel state information, we assume perfect CSI knowledge for *each* user at the receive side[1], and the availability of a low-rate feedback channel to convey partial CSI to the BS, in particular their instantaneous SNR. As for the partial CSI received at the scheduler, different assumptions will be made along this chapter.

[1] In Section 4.6 we will show an isolated case where perfect CSI is not perfectly known at the receiver.

4.2.2 Transmission Schemes

As far as the transmission schemes are concerned, we will consider four different cases:

- *SISO (Single-Input Single-Output):*

This constitutes the first baseline case where the BS uses one single antenna at all the times. In this situation, $M_{TX} = M = 1$ where M_{TX} and M stand for the number of *active* and *available* antennas at the BS, respectively. By denoting the complex channel gains between the *i-th* transmit antenna in the BS and the *k-th* user with $h_{i,k}$, the received SNR for user k becomes

$$\gamma_{k,SISO} = \bar{\gamma}_k \left|h_{1,k}\right|^2 \tag{4.2}$$

- *SISO-AS (SISO with Antenna Selection):*

In this scheme, the *best* antenna ($M_{TX} = 1$) out of the $M > M_{TX}$ antennas available in the BS will be selected for data transmission, more precisely, the one that maximizes the received SNR for user k

$$\gamma_{k,SISO-AS} = \bar{\gamma}_k \max_{1 \leq i \leq M} \left\{ \left|h_{i,k}\right|^2 \right\} \tag{4.3}$$

This approach requires the antenna index to be fed back along with SNR information. For that purpose, $\log_2 M$ additional signalling bits are needed.

- *OSTBC (Orthogonal Space-Time Block Coding),*

Besides from the channel hardening effect associated with the increase in the number of transmit antennas in an OSTBC scheme [Hoc04], full rate is only achieved with $M_{TX} = M = 2$ transmit antennas. For that reason, although the analysis can be easily extended to the general case, we restrict ourselves to consider an OSTBC scheme with $M_{TX} = M = 2$ transmit antennas (i.e., the Alamouti scheme [Ala98]). In this case, the received SNR for user k turns out to be

$$\gamma_{k,OSTBC} = \frac{\bar{\gamma}_k}{2} \left(\left|h_{1,k}\right|^2 + \left|h_{2,k}\right|^2 \right) = \frac{\bar{\gamma}_k}{2} \left|\left|\mathbf{h}_k\right|\right|^2 \tag{4.4}$$

since power is evenly allocated to transmit antennas.

- *OSTBC-AS (OSTBC with Antenna Selection)*:

Now, the antenna subset with M_{TX} out of the M antennas available in the BS that maximizes the received SNR will be chosen and, thus, we have

$$\gamma_{k,OSTBC-AS} = \frac{\bar{\gamma}_k}{2} \max_{1 \leq v \leq V} \left\{ \left|\left|\mathbf{h}_k^{(v)}\right|\right|^2 \right\}, \tag{4.5}$$

where superscript v is an index to antenna subsets. As in the *SISO-AS* case, a number of additional signalling bits equal to $\log_2 V = \log_2 \binom{M}{M_{TX}}$ are needed in order to convey such antenna subset index over the feedback channel.

4.2.3 Centralized Scheduler

In the proposed system, where only one user is scheduled for transmission in each time-slot, the average cell throughput can be maximized if a *max-SNR* (*greedy*) scheduling strategy is adopted. That is, by selecting in each time-slot the active user k^* with the highest channel gain:

$$k^* = \arg\max_k \{\gamma_1, ..., \gamma_k, ..., \gamma_K\}$$

Such scheduler, though, may be unfair in the case that channel conditions are not statistically equal for all the users. In other words, the users with the worst channels are never scheduled. In order to circumvent that, an adapted version of the Proportional Fair Scheduling [Vis02] is adopted. In particular, in each time-slot the user with the maximum *normalized* SNR is scheduled, i.e.,

$$k^* = \arg\max_k \left\{ \frac{\gamma_1}{\bar{\gamma}_1}, ..., \frac{\gamma_k}{\bar{\gamma}_k}, ..., \frac{\gamma_K}{\bar{\gamma}_K} \right\}$$

By doing so, users are only allowed to transmit when the instantaneous SNR is near *its own peak* [Vis02], that is with respect to their average SNR. As a result, multi-user diversity can still be exploited and the scheduler will grant access probability of $1/K$ to each user.

4.3 Pre-scheduling SNR Statistics

In this section, we are interested in deriving analytical expressions for the statistics (pdf and CDF) of the pre-scheduling SNRs. We define *pre-scheduling* signal-to-noise ratio *for user* k as the SNR that is measured and reported by such user to the base station (the SNR expressions for the different transmission schemes can be found in Eqs. (4.2)-(4.5) above).

Certainly, the probability density function of the pre-scheduling SNR, $f_\gamma(\gamma)$, will strongly depend on the transmission scheme but, conversely, the cumulative density function unequivocally relates to the pdf through

$$F_{\gamma'}(\gamma) = \text{Prob}(\gamma' \leq \gamma) = \int_0^\gamma f_{\gamma'}(\gamma')\, d\gamma' \tag{4.6}$$

A case-by-case analysis follows:

- *SISO*:

Since a Rayleigh fading case is assumed, the SNR for user k follows a chi-square random variable with two degrees of freedom, χ^2_2. Hence, the pdf and CDF take the following expressions:

$$f_{\gamma_{k,SISO}}(\gamma) = \frac{1}{\bar{\gamma}_k} e^{-\frac{\gamma}{\bar{\gamma}_k}} \qquad F_{\gamma_{k,SISO}}(\gamma) = 1 - e^{-\frac{\gamma}{\bar{\gamma}_k}} \tag{4.7}$$

- *SISO-AS*:

Bearing in mind that channel coefficients corresponding to different BS antennas fade independently from each other, one can find that

$$\begin{gathered} F_{\gamma_{k,SISO-AS}}(\gamma) = \text{Prob}\left(\frac{P_t}{\sigma^2}\max_{1\le i\le M}\left\{|h_{i,k}|^2\right\} \le \gamma\right) = \left(F_{\gamma_{k,SISO}}(\gamma)\right)^M = \left(1 - e^{-\frac{\gamma}{\bar{\gamma}_k}}\right)^M \\ f_{\gamma_{k,SISO}}(\gamma) = \frac{M}{\bar{\gamma}_k} e^{-\frac{\gamma}{\bar{\gamma}_k}} \left(1 - e^{-\frac{\gamma}{\bar{\gamma}_k}}\right)^{M-1} \end{gathered} \tag{4.8}$$

- *OSTBC*:

Now, the signal-to-noise ratio becomes a chi-square random variable with $2M_{TX}$ degrees of freedom, $\chi^2_{2M_{TX}}$. For the two-antenna Alamouti scheme, we have

$$f_{\gamma_{k,OSTBC}}(\gamma) = \frac{4\gamma}{\bar{\gamma}_k^2} e^{-\frac{2\gamma}{\bar{\gamma}_k}} \tag{4.9}$$

$$F_{\gamma_{k,OSTBC}}(\gamma) = 1 - e^{-\frac{2\gamma}{\bar{\gamma}_k}}\left(\frac{2\gamma}{\bar{\gamma}_k} + 1\right) \tag{4.10}$$

- *OSTBC-AS*:

In the proposed scheme, the best two out of M BS antennas will be selected for data transmission. In the sequel, we will restrict ourselves to the case $M = 3$ in order to keep the analytical derivations tractable (in particular for the expressions derived in Section 4.5). Since the search over the available antenna subsets is aimed at maximizing the received SNR, the resulting problem can be readily solved by means of the so-called *order statistics* [Dav81]. By defining X_1, X_2 and X_3 as the channel gains associated to the different transmit antennas after applying a power normalization and arranging them in an increasing order ($X_1 \le X_2 \le X_3$), it is clear that the resulting SNR is maximized when the highest and second-highest elements are selected, that is

$$\begin{aligned} F_{\gamma_{k,OSTBC-AS}}(\gamma) &= \text{Prob}\left(\max_{1\le i,j\ne i\le M}\left\{\frac{\bar{\gamma}_k}{2}(X_i + X_j)\right\} \le \gamma\right) \\ &\underset{M=3}{=} \text{Prob}\left((X_2 + X_3) \le \frac{2\gamma}{\bar{\gamma}_k}\right) \end{aligned}$$

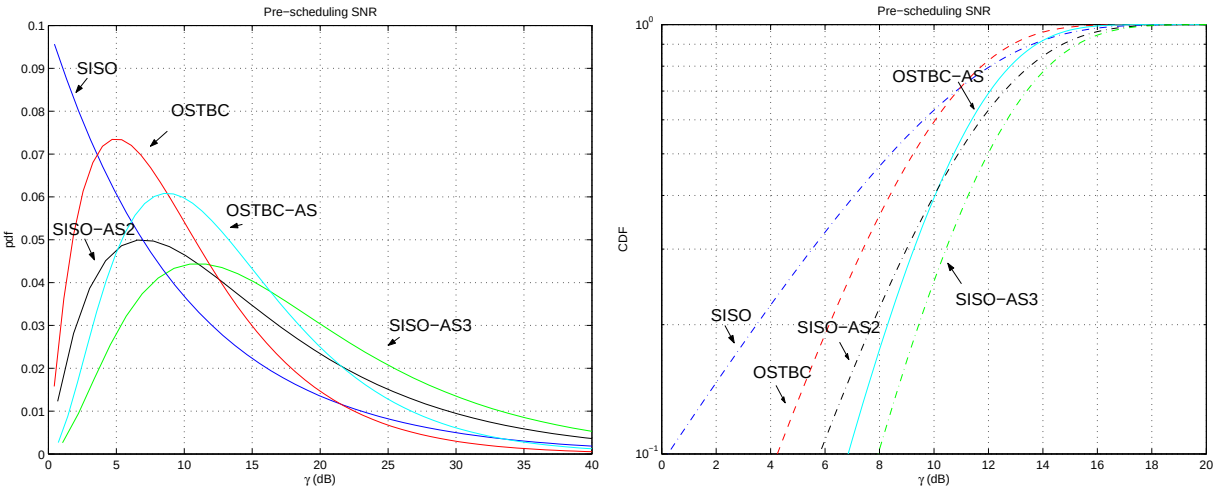


Figure 4.2: Density functions of the pre-scheduling SNR. Left, probability density functions. Right, cumulative density functions in log-scale. ($\bar{\gamma}_k = 10$ dB).

The joint density function of the (individually χ^2_2-distributed) *ordered* random variables X_2 and X_3 is given by:

$$f_{X_3X_2}(x_3, x_2) = \begin{cases} 3!\,(1 - e^{-x_2})\, e^{-(x_2+x_3)} & \text{for } x_2 \leq x_3 \\ 0 & \text{otherwise} \end{cases}$$

and, from that, we can obtain the corresponding closed-form expressions for both the CDF and pdf functions.

$$\begin{aligned} F_{\gamma_{k,OSTBC-AS}}(\gamma) &= \int_{x_3=0}^{\frac{\gamma}{\bar{\gamma}_k}} \int_{x_2=0}^{x_3} f_{X_3X_2}(x_3, x_2) dx_3 dx_2 + \int_{x_3=\frac{\gamma}{\bar{\gamma}_k}}^{\frac{2\gamma}{\bar{\gamma}_k}} \int_{x_2=0}^{\frac{2\gamma}{\bar{\gamma}_k}-x_3} f_{X_3X_2}(x_3, x_2) dx_3 dx_2 \\ &= 1 - e^{-\frac{2\gamma}{\bar{\gamma}_k}} \left(\frac{6\gamma}{\bar{\gamma}_k} + 4e^{-\frac{\gamma}{\bar{\gamma}_k}} - 3 \right) \\ f_{\gamma_{k,OSTBC-AS}}(\gamma) &= \frac{12}{\bar{\gamma}_k} e^{-\frac{2\gamma}{\bar{\gamma}_k}} \left(e^{-\frac{\gamma}{\bar{\gamma}_k}} + \frac{\gamma}{\bar{\gamma}_k} - 1 \right) \end{aligned} \tag{4.11}$$

In Fig. 4.2, we depict both the pdf and CDF of the pre-scheduling SNR for the different transmission schemes, respectively. When comparing the curves for the SISO and OSTBC schemes, one can observe that OSTBC substantially *stabilizes* the SNR distribution. This means that the probability that a deep fade occurs decreases but, unfortunately, the number of SNR peaks (those that multi-user diversity can exploit) is reduced, as well. However, this undesirable effect can be compensated by introducing antenna selection mechanisms (*OSTBC-AS*) which partly restore the *missing* SNR peaks (for $\gamma \geq 14$ dB) and, simultaneously, further suppress deep fades. As a result, the effective SNR experienced by users in the system increases.

Besides, one can also observe that an SISO-AS configuration with $M = 2$ (i.e. curve labelled with SISO-AS2) performs better than the OSTBC-AS scheme in terms of SNR peaks generation.

However, its capability of mitigating fades is far more limited (see crossing of CDF curves around $\gamma = 10$ dB), which can be improved by increasing the number of transmit antennas to $M = 3$ (SISO-AS3 configuration).

In summary, antenna selection strategies provide an effective means to shape the pre-scheduling SNR statistics exhibited by the OSTBC and SISO schemes in terms of fade occurrence and SNR peak generation. However, one should bear in mind that it is measures such as the ergodic capacity or system throughput (and not SNR distributions) that provide a more accurate view on the actual performance. Besides, we have not yet come to the point of analyzing robustness issues resulting from bandwidth constraints or impairments in the feedback channels. The following sections are devoted to conduct such an analysis.

4.4 Analysis in a Scenario with Ideal Feedback Channel

In this section, we analyze and compare performance for the different transmission schemes in a scenario where *all* users permanently monitor and report the *pre-scheduling* SNR to the centralized scheduler over a error- and delay-free feedback channel. Ultimately, we aim at conducting a capacity and throughput analysis but, to do so, we first have to obtain the closed-form expressions of the *post-scheduling* SNRs. We define *post-scheduling* SNR as the signal-to-noise ratio experienced by the scheduled user.

4.4.1 Post-scheduling SNR

By defining $\mathcal{A}_k$ as the event that user k is the scheduled user and by applying Bayes theorem, one can readily obtain the CDF of the post-scheduling SNR, γ^*, as:

$$F_{\gamma^*}(\gamma) = \sum_{k=1}^{K} \text{Prob}\left(\gamma_k \leq \gamma | \mathcal{A}_k\right) \text{Prob}\left(\mathcal{A}_k\right)$$

where $\text{Prob}(\mathcal{A}_k) = 1/K$, $k = 1, ..., K$, due to the properties of the adopted scheduler [Vis02]. Now, by focusing on user k, we can easily derive the probability $\text{Prob}\left(\gamma_k \leq \gamma | \mathcal{A}_k\right)$. In particular, by recalling that all users experience independently-distributed fading, we have:

$$\begin{aligned}\text{Prob}\left(\gamma_k \leq \gamma | \mathcal{A}_k\right) &= \text{Prob}(\gamma_k \leq \gamma) \text{Prob}\left(\frac{\gamma_i}{\bar{\gamma}_i} \leq \frac{\gamma}{\bar{\gamma}_k} \text{ for all } i \neq k\right) \\ &= F_{\gamma_k}(\gamma) \prod_{\substack{i=1 \\ i \neq k}}^{K} F_{\gamma_i}\left(\frac{\gamma}{\bar{\gamma}_k} \bar{\gamma}_i\right) = \left(F_{\gamma_k}(\gamma)\right)^K\end{aligned}$$

As a result, the CDF of the post-scheduling SNR can be written as:

$$F_{\gamma^*}(\gamma) = \frac{1}{K} \sum_{k=1}^{K} \left(F_{\gamma_k}(\gamma)\right)^K$$

By differentiating with respect to γ, the corresponding pdf expression can be found

$$f_{\gamma^*}(\gamma) = \sum_{k=1}^{K} \left(F_{\gamma_k}(\gamma)\right)^{K-1} f_{\gamma_k}(\gamma)$$

where $F_{\gamma_k}(\gamma)$ and $f_{\gamma_k}(\gamma)$ in the above equations for the different transmission schemes can be found in Eqs. (4.7)-(4.11). In summary, the closed-form expressions for the pdf of the post-scheduling SNRs are

$$f_{\gamma^*_{SISO}}(\gamma) = \sum_{k=1}^{K} \frac{e^{-\frac{\gamma}{\bar{\gamma}_k}}}{\bar{\gamma}_k} \left(1 - e^{-\frac{\gamma}{\bar{\gamma}_k}}\right)^{K-1} \tag{4.12}$$

$$f_{\gamma^*_{SISO-AS}}(\gamma) = \sum_{k=1}^{K} M \frac{e^{-\frac{\gamma}{\bar{\gamma}_k}}}{\bar{\gamma}_k} \left(1 - e^{-\frac{\gamma}{\bar{\gamma}_k}}\right)^{KM-1} \tag{4.13}$$

$$f_{\gamma^*_{OSTBC}}(\gamma) = \sum_{k=1}^{K} \frac{4\gamma}{\bar{\gamma}_k^2} e^{-\frac{2\gamma}{\bar{\gamma}_k}} \left(1 - e^{-\frac{2\gamma}{\bar{\gamma}_k}} \left(\frac{2\gamma}{\bar{\gamma}_k} + 1\right)\right)^{K-1} \tag{4.14}$$

$$f_{\gamma^*_{OSTBC-AS}}(\gamma) = \sum_{k=1}^{K} \frac{12}{\bar{\gamma}_k} e^{-\frac{2\gamma}{\bar{\gamma}_k}} \left(e^{-\frac{\gamma}{\bar{\gamma}_k}} + \frac{\gamma}{\bar{\gamma}_k} - 1\right) \left(1 - e^{-\frac{2\gamma}{\bar{\gamma}_k}} \left(\frac{6\gamma}{\bar{\gamma}_k} + 4e^{-\frac{\gamma}{\bar{\gamma}_k}} - 3\right)\right)^{K-1} \tag{4.15}$$

In the sequel, since the proposed scheduler is fair irrespective of the average SNRs, we will restrict ourselves to only show results corresponding to homogeneous scenarios (i.e., $\bar{\gamma}_k = \bar{\gamma}$ for $k = 1, ..., K$). Figure 4.3 shows the CDF function of the post-scheduling SNR in a system with K=1, 5 and 50 users. Notice that the case with K=1 users is essentially equivalent to that of *round-robin* scheduling, i.e., letting K users take turns in utilizing the channel, regardless of their SNR. From the relative ordering in each group of curves, it becomes clear that in a OSTBC-AS scheme the additional *spatial* diversity provided by the antenna selection mechanism is better exploited in a scenario with a low to moderate number of users (K=1,5). The approximate gain with respect to the SISO configuration turns out to be 4 and 2 dB, respectively. Conversely, as the number of users increases (K=50) the beneficial effect of multi-user diversity on a SISO scheme exceeds spatial diversity gains.

Apart from that, one can also observe that the SISO-AS3 scheme is far more successful than its OSTBC-AS counterpart in shifting the SNR curves towards higher values. In other words, the joint exploitation of MUD and the spatial diversity component provided by the antenna selection mechanisms is more effectively done when no space-time block coding scheme is used. Actually, the introduction of antenna selection mechanisms is often interpreted as an increase in the number of active users since, in the end, SNRs are measured for a higher number of antenna pairs. In the OSTBC-AS case, though, the SNRs for the different antenna subsets are correlated due to the fact that some antenna subsets have one or more elements in common. As a result the effective increase in the number of users is less than V-fold (to recall, $V = \binom{M}{M_{TX}}$ is the number of antenna subsets).

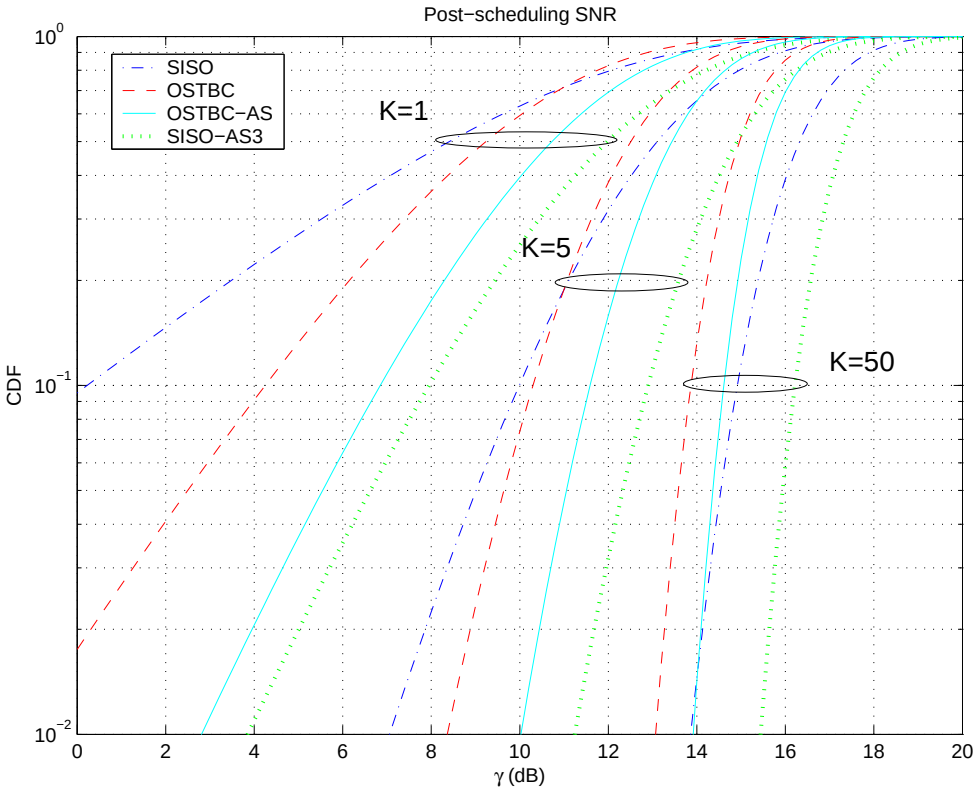


Figure 4.3: CDF of the received SNR for the different transmission schemes, for $K = 1, 5, 50$ users (left to right, $\bar{\gamma} = 10$ dB).

4.4.2 Ergodic Capacity Analysis

So far, the different transmission schemes have been compared in terms of post-scheduling SNR densities only. For that reason a complementary study addressing system capacity and throughput aspects becomes necessary. The Shannon capacity achievable by the scheduled user k^* over the equivalent single-input single-output channel is given by

$$C^* = \log_2 (1 + \gamma^*)$$

and, consequently, the ergodic system capacity achievable with a max-normalized SNR scheduling policy can be expressed as

$$\overline{C}(K) = \mathbb{E}_\gamma [C^*] = \int_{\gamma=0}^{\infty} \log_2 (1 + \gamma) f_{\gamma^*} (\gamma) d\gamma \tag{4.16}$$

By plugging expressions (4.12)-(4.15) into the above equation, one can obtain the corresponding capacities for the SISO, SISO-AS, OSTBC and OSTBC-AS schemes.

For the SISO case, one should resort to the binomial expansion of Eq. (4.12) and find out

that the following integral must be solved:

$$\overline{C}_{SISO}(K) = \log_2(e) \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} \frac{(-1)^i}{\bar{\gamma}_k} \int_{\gamma=0}^{\infty} \ln(1+\gamma) e^{-\frac{\gamma}{\bar{\gamma}_k}(i+1)} d\gamma$$

The above integral can be easily computed with the help of [Gra65, Eq. 4.331.2]:

$$\overline{C}_{SISO}(K) = -\log_2(e) \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} \frac{(-1)^i}{i+1} e^{\frac{i+1}{\bar{\gamma}_k}} E_i\left(-\frac{i+1}{\bar{\gamma}_k}\right)$$

with $E_i(x)$ standing for the exponential integral function ($E_i(x) \triangleq -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$, for $x < 0$). From the above expression, we can also obtain the ergodic capacity for the SISO-AS approach:

$$\overline{C}_{SISO-AS}(K) = -\log_2(e) M \sum_{k=1}^{K} \sum_{i=0}^{MK-1} \binom{MK-1}{i} \frac{(-1)^i}{i+1} e^{\frac{i+1}{\bar{\gamma}_k}} E_i\left(-\frac{i+1}{\bar{\gamma}_k}\right)$$

As for the OSTBC case, the following expression can be obtained:

$$\overline{C}_{OSTBC}(K) = \log_2(e) \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{k} (-1)^i \sum_{n=0}^{i} \binom{i}{n} \left(\frac{2}{\bar{\gamma}_k}\right)^{n+2} \Upsilon\left(0, n+2, \frac{2}{\bar{\gamma}_k}(i+1)\right)$$

where $\Upsilon(a, m, \mu)$ is an integral equation. Details about this equation can be found in Appendix 4.A.

Finally, by repeatedly applying the binomial expansion, one can verify that the following expression holds for the OSTBC-AS case:

$$\overline{C}_{OSTBC-AS}(K) = \log_2(e) 12 \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} 3^i \sum_{n=0}^{i} \binom{i}{n} \left(-\frac{1}{3}\right)^n \sum_{s=0}^{n} \binom{n}{s} \frac{6^{n-s} 4^s}{\bar{\gamma}_k^{n-s+1}}$$
$$\times \left[\Upsilon\left(0, n-s+1, \frac{2i+s+3}{\bar{\gamma}_k}\right) + \frac{1}{\bar{\gamma}_k} \Upsilon\left(0, n-s+2, \frac{2i+s+2}{\bar{\gamma}_k}\right) - \Upsilon\left(0, n-s+1, \frac{2i+s+2}{\bar{\gamma}_k}\right) \right]$$

4.4.3 System Throughput Analysis

Ergodic capacity measures obtained in the previous subsection simply provide a rough idea on how spectrally efficient the system can be when an *infinite* number of modulation and ideal coding schemes are available. In practical cases, though, coding schemes may be far from being ideal and the number of AMC schemes can be rather limited. For instance, in HSDPA two modulation schemes are available: QPSK and 16-QAM. In that scenario, it is also worth investigating to what extent can MUD and spatial diversity be jointly exploited. In particular, we will restrict ourselves to consider uncoded transmissions[2] with two modulation schemes in the PHY layer and, then, assess the associated average system throughput.

[2] As commented in the previous chapter, we restrict the analysis of link layer throughput to uncoded transmissions in order to keep analytical derivations tractable.

First, we derive expression of the (post-scheduling) instantaneous throughput for a given modulation scheme m:

$$\eta_m^* = b_m \cdot (1 - \mathrm{PER}_m(\gamma^*)) = b_m \cdot (1 - \mathrm{SER}_m(\gamma^*))^L$$

where L stands for the number of symbols in the transmitted burst and b_m is the number of bits per symbol according to the selected modulation scheme[3]. By averaging over all the realizations of γ^*, the average throughput can be obtained:

$$\overline{\eta}_m(K) = \mathbb{E}_\gamma[\eta_m^*] = b_m \int_{\gamma=0}^{\infty} (1 - \mathrm{SER}_m(\gamma))^L f_{\gamma^*}(\gamma) d\gamma \tag{4.17}$$

With the accustomed bound of the SER [Pro01], the expression above is barely integrable. Alternatively, we will adopt the approximation for M-QAM modulation schemes presented in [Chu01]:

$$\mathrm{SER}_m(\gamma^*) \approx b_m 0.2 e^{-1.6\frac{\gamma^*}{2^{b_m}-1}} = \alpha_m e^{-\beta_m \gamma^*} \tag{4.18}$$

Now, an adaptive modulation mechanism is derived. We propose a strategy where the constellation size maximizing the throughput expression is chosen, that is:

$$m^* = \arg\max_{m \in \mathcal{M}} b_m (1 - \alpha_m e^{-\beta_m \gamma^*})^L$$

where $\mathcal{M}$ stands for the set of available modulation schemes $\mathcal{M} = \{\text{QPSK, 16-QAM}\}$. From the above expression, it is straightforward to obtain the corresponding AMC thresholds, $\gamma_{th,m}$ (see Fig. 4.4). Consequently, the constellation size associated with the measured post-scheduling SNR can be determined according to the following rule:

$$m^* = m \Longleftrightarrow \gamma_{th,m} \leq \gamma^* < \gamma_{th,m+1}$$

with $\gamma_{th,1} = 0$ and $\gamma_{th,\mathrm{card}(\mathcal{M})+1} = \infty$. Then, by taking into consideration Eq. (4.18) and the proposed adaptive modulation mechanism, the average throughput can be expressed as:

$$\overline{\eta}(K) \approx \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \int_{\gamma=\gamma_{th,m}}^{\gamma_{th,m+1}} (1 - \alpha_m e^{-\beta_m \gamma})^L f_{\gamma^*}(\gamma) d\gamma$$

The above expression can be particularized for the different transmission schemes with the help of Eqs.(4.12)-(4.15). By using the binomial expansion and integrating by parts one can derive the following closed-form expressions for the SISO and SISO-AS approaches:

$$\overline{\eta}_{SISO}(K) \approx \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} (-1)^i \sum_{l=0}^{L} \binom{L}{l} \frac{(-\alpha_m)^l}{\bar{\gamma}_k \beta_m l + i + 1} \times \left[e^{-\frac{\gamma_{th,m}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + i + 1)} - e^{-\frac{\gamma_{th,m+1}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + i + 1)} \right]$$

[3]Notice that such an expression corresponds to the link layer throughput of an N-SAW ARQ protocol with as many round-trip delay slots as SAW processes in parallel ($W = N_{SAW}$) [Lin84, UTR01].

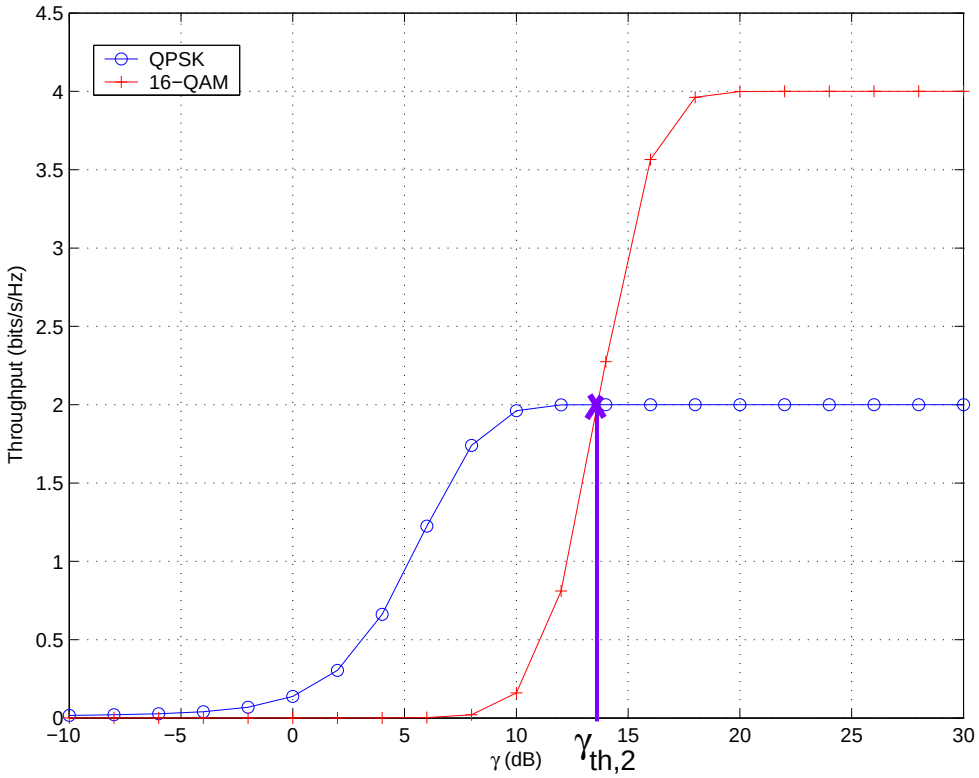


Figure 4.4: Throughput vs. SNR for the different modulation schemes (L=10 symbols).

$$\overline{\eta}_{SISO-AS}(K) \approx M \sum_{m=1}^{\text{card}(\mathcal{M})} b_m \sum_{k=1}^{K} \sum_{i=0}^{MK-1} \binom{MK-1}{i} (-1)^i \sum_{l=0}^{L} \binom{L}{l} \frac{(-\alpha_m)^l}{\bar{\gamma}_k \beta_m l + i + 1} \times \left[e^{-\frac{\gamma_{th,m}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + i + 1)} - e^{-\frac{\gamma_{th,m+1}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + i + 1)} \right]$$

Concerning the OSTBC-based schemes, the integrations involved in the resolution of their system throughput expressions are of the type $\int_c^d t^m e^{-t\mu} dt = \int_c^\infty t^m e^{-t\mu} dt - \int_d^\infty t^m e^{-t\mu} dt$, which can be easily solved by means of [Gra65, Eq. 3.351.1]. As a result, the average throughput for both the OSTBC and OSTBC-AS approaches can be written in terms of the complementary incomplete gamma function ($\Gamma(\alpha, x) \triangleq \int_x^\infty e^{-t} t^{\alpha-1} dt$) as follows:

$$\overline{\eta}_{OSTBC}(K) \approx 4 \sum_{m=1}^{\text{card}(\mathcal{M})} b_m \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} (-1)^i \sum_{l=0}^{L} \binom{L}{l} (-\alpha_m)^l \sum_{n=0}^{i} \binom{i}{n} \frac{2^n}{(\bar{\gamma}_k \beta_m l + 2i + 2)^{n+2}} \times \left[\Gamma(n+2, \frac{\gamma_{th,m}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + 2)) - \Gamma(n+2, \frac{\gamma_{th,m+1}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + 2)) \right]$$

$$\overline{\eta}_{OSTBC-AS}(K) \approx 12 \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} 3^i \sum_{l=0}^{L} \binom{L}{l} (-\alpha_m)^l \sum_{n=0}^{i} \binom{i}{n} \left(-\frac{1}{3}\right)^n \sum_{s=0}^{n} \binom{n}{s} 6^{n-s} 4^s$$
$$\times \left[\frac{\Gamma(n-s+1, \frac{\gamma_{th,m}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + s + 3)) - \Gamma(n-s+1, \frac{\gamma_{th,m+1}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + s + 3))}{(\bar{\gamma}_k \beta_m l + 2i + s + 3)^{n-s+1}} \right.$$
$$+ \frac{1}{(\bar{\gamma}_k \beta_m l + 2i + s + 2)^{n-s+1}} \left(\frac{\Gamma(n-s+2, \frac{\gamma_{th,m}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + s + 2)) - \Gamma(n-s+2, \frac{\gamma_{th,m+1}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + s + 2))}{\bar{\gamma}_k \beta_m l + 2i + s + 2} \right.$$
$$\left. \left. - \Gamma(n-s+1, \frac{\gamma_{th,m}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + s + 2)) + \Gamma(n-s+1, \frac{\gamma_{th,m+1}}{\bar{\gamma}_k}(\bar{\gamma}_k \beta_m l + 2i + s + 2)) \right) \right]$$

It is worth noting that when the above expressions are evaluated, one should bear in mind that $\lim_{x \to \infty} \Gamma(\alpha, x) = 0$ and $\Gamma(\alpha, 0) = \Gamma(\alpha)$, with $\Gamma(\alpha)$ standing for the Gamma function. The Gamma function is defined as $\Gamma(\alpha) \triangleq \int_0^\infty e^{-t} t^{\alpha-1} dt$. For α a natural number, this function can be easily computed as $\Gamma(\alpha) = (\alpha - 1)!$.

4.4.4 Numerical Results and Discussion

Throughout this section, we will consider a system with a number of active users in the range $K = 1, ..., 30$ transmitting data packets with $L = 10$ symbols in each. The average SNR is set to $\bar{\gamma} = 10$ dB (recall that we consider homogeneous scenarios due to the fair behavior of the adopted scheduler).

In Fig. 4.5, we plot the ergodic *capacity* as a function of the number of active users. One can observe that OSTBC systems without antenna selection mechanisms exclusively outperform SISO schemes in the K=1 case, that is, when multi-user diversity cannot be exploited at all. Conversely, in the presence of MUD ($K \geq 2$) the spatial diversity component associated to the OSTBC scheme penalizes (rather than boosts) system performance. However, the system exhibits a different behavior when an additional diversity component is introduced via antenna selection mechanisms. On the one hand, the OSTBC-AS scheme efficiently combines antenna diversity and multi-user diversity and, thus, performs better than the SISO approach in the $K = 1..8$ users range. This can be very beneficial, for instance, in the early phases of deployment of packet data systems or, perhaps, in scarcely populated areas where the number of users per cell is potentially low. On the other hand, a remarkable shift can be observed between OSTBC-AS and OSTBC curves for the whole range of users. Due to correlation effects among antenna subsets, the effective increase in the number of users is a factor of 2 (e.g. 5 vs. 10 users @ C=4.5 bits/s/Hz) whereas the number of subsets is actually V=3. As for the comparison between the SISO-AS and OSTBC-AS schemes in terms of channel capacity, the former outperforms the latter for the whole range of users (as pointed out in the SNR analysis above). However, this analysis has been conducted in a rather idealized scenario: with an infinite number of AMC levels and for an error- and delay-free infinite-bandwidth feedback channel. The throughput analysis

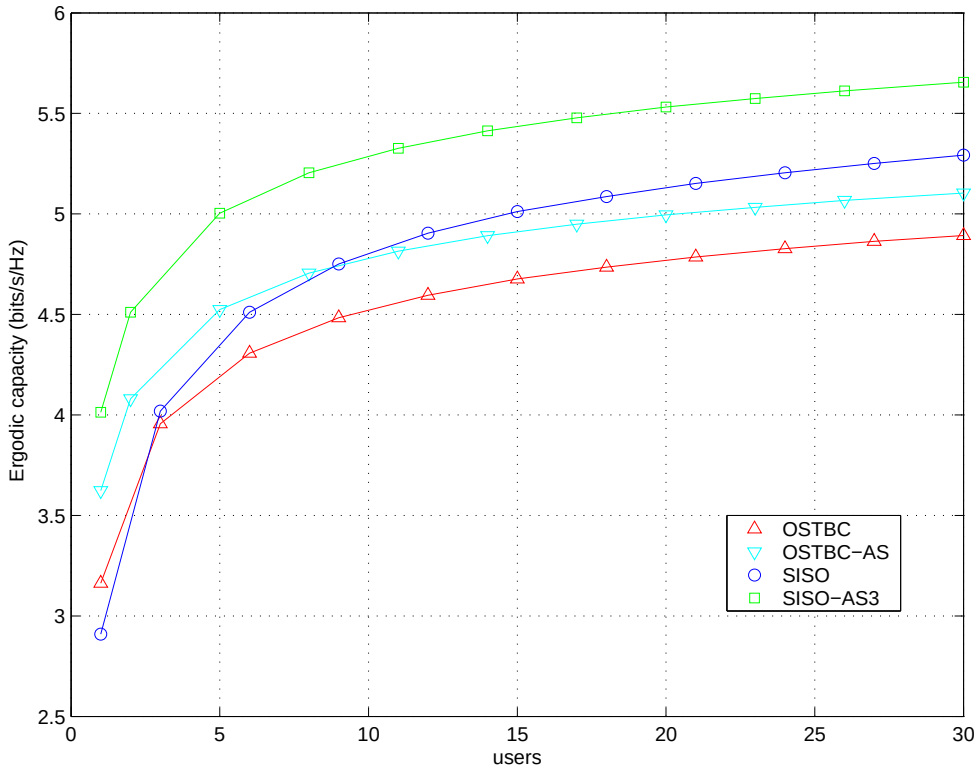


Figure 4.5: Ergodic capacity vs. the number of active users for the different transmission schemes in a scenario with ideal feedback channel (symbols: simulated results, curves: analytical expressions, $\bar{\gamma}$=10 dB).

to follow will bring things closer to the real world.

The average system *throughput* curves depicted in Fig. 4.6 show that, for a given modulation scheme (QPSK), MUD and spatial diversity can still be effectively combined in the OSTBC-AS scheme, in particular for a low to moderate number of users. As opposed to ergodic capacity, throughput curves for SISO-AS3 exhibit slightly better performance with respect to the OSTBC-AS approach and OSTBC outperforms SISO for the whole range of users. The reason for that is that the system cannot take advantage of the instantaneous SNR peaks resulting from MUD gains due to bit cap effects [Sta99]. As a consequence, the reduced probability of deep fades associated with the OSTBC and AS-based approaches prevails against peak generation capability when only QPSK is considered. By introducing an additional degree of freedom in the system (adaptive modulation with QPSK and 16-QAM), the possible transmission rate is not saturated and the peak generation capability can be better exploited.

In summary, according to the number of users and the set of AMC levels taken into consideration at the base station, the trade-offs among transmission schemes, MUD and antenna selection mechanisms must be carefully established.

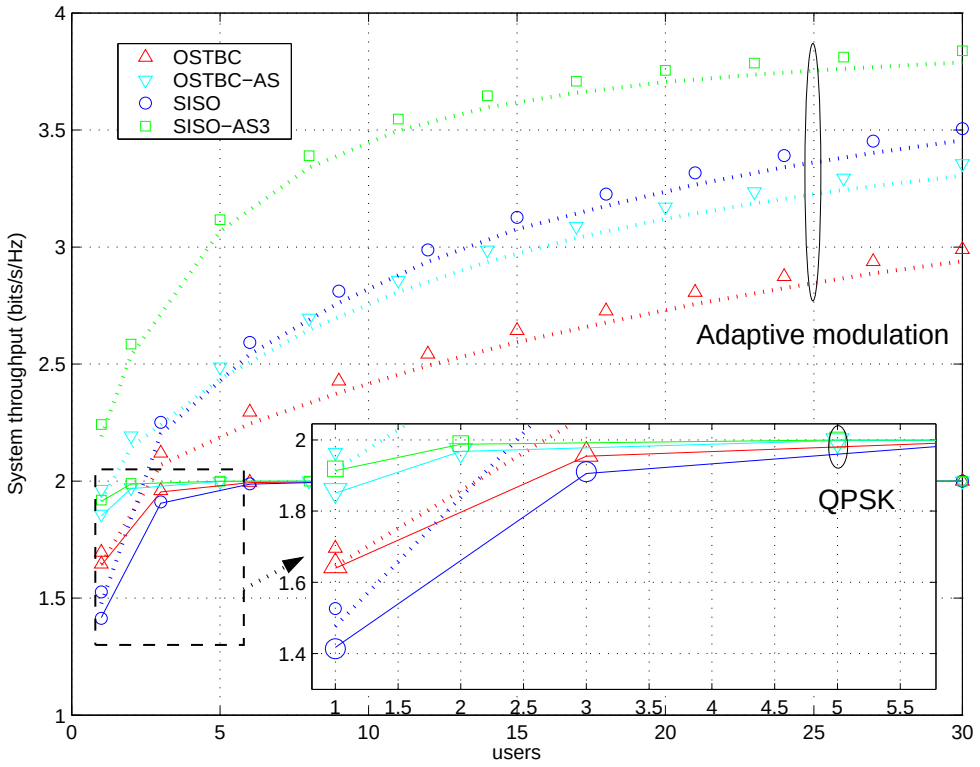


Figure 4.6: System throughput vs. the number of users for the different transmission schemes in a scenario with ideal feedback channel (solid line: QPSK, dotted line: adaptive modulation, symbols: simulated results, curves: analytical expressions, L=10 symbols, $\bar{\gamma} = 10$ dB). A close match is observed between the derived approximations and the simulated results.

4.5 Analysis in a Scenario with Selective-feedback Scheduling

In this section, we extend the analysis conducted by Gesbert et al in [Ges03, Ges04, Yan04] on Selective MUD (SMUD) schemes to encompass OSTBC and antenna selection mechanisms. The rationale behind the analysis of the different transmission schemes (SISO, SISO-AS, OSTBC and OSTBC-AS) in a Selective MUD scenario lies in the fact that, as the number of active users increases, their individual contribution to feedback channels should be necessarily decreased. In those conditions, chances are higher for the centralized scheduler in the BS to be forced to make a blind decision (i.e. when no mobile station conveys CSI to the scheduler for a specific time-slot). The robustness provided by the OSTBC-based approaches against deep fades associated to such random user picks could then pay off.

In a SMUD scenario, only terminals experiencing normalized-SNRs above a pre-defined threshold (ξ_{th}) are allowed to report their channel state information to the BS. Hence, the max normalized-SNR scheduler proposed in Section 4.2.3 will conduct the search over such a subset of the active users only. In summary, the scheduler is driven according to the following

rule:

$$k^* = \arg\max_k \left\{ \frac{\gamma_k}{\bar{\gamma}_k} \text{ s.t. } \frac{\gamma_k}{\bar{\gamma}_k} > \xi_{th} \right\}$$

when, at least, one user is reported. Conversely, when all the users remain silent (i.e. in the event of a *scheduling outage*) the scheduling rule amounts to:

$$k^* = \text{rand}\{1, ..., k, ..., K\}$$

where users are randomly selected.

Notice that, with the proposed scheduler both the max normalized-SNR and random schedulers will grant access probability of $1/K$ to each user. Clearly, since we consider a selective scheduling strategy with normalized threshold [Yan04], the probability of staying in either of the two scheduling regions (max normalized-SNR or random scheduling) is the same for all users (i.e., the scheduler is fair).

4.5.1 Post-scheduling SNR

In [Yan04], a expression was already derived for the formulas associated to the post-scheduling SNR statistics of the SISO approach (see [Yan04, Section IV] for more details). In this section, we derive a simplified version of that expression by taking advantage of the behavior of the proposed selective scheduler. In other words, by recalling that $\mathcal{A}_k$ is the event that user k is the scheduled user and expressing the CDF of the post-scheduling SNR as:

$$F_{\gamma^*}(\gamma) = \sum_{k=1}^{K} \text{Prob}\left(\gamma_k \leq \gamma | \mathcal{A}_k\right) \text{Prob}\left(\mathcal{A}_k\right)$$

we can obtain the CDF expression in a simpler way. This is because, the CDF is expressed in terms of the probabilities $\text{Prob}\left(\gamma_k \leq \gamma | \mathcal{A}_k\right), k = 1, ..., K$, which can be easily computed due to the normalization effect provided by the max-normalized SNR scheduler.

As in the previous section, Prob($\mathcal{A}_k$) is equal to 1/K and the analysis is focused on user k. However, two different SNR regions must be considered in this case: (1) the random scheduling region ($\gamma_k/\bar{\gamma}_k \leq \xi_{th}$), where all users remain silent[4]; and (2) the max-normalized scheduling region ($\gamma_k/\bar{\gamma}_k > \xi_{th}$), where at least user k reports back its SNR.

For the $\gamma_k/\bar{\gamma}_k \leq \xi_{th}$ case and by recalling that all users experience independently distributed

[4]In this region, user k is not allowed to report its SNR status to the BS. Therefore, this user can only be randomly selected for transmission when the scheduling is in random scheduling configuration, i.e., all the users are silent.

fading, we have:

$$\begin{aligned}\text{Prob}\left(\gamma_k \leq \gamma | \mathcal{A}_k\right) &= F_{\gamma_k|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right) \\ &= \text{Prob}(\gamma_k \leq \gamma)\text{Prob}\left(\frac{\gamma_i}{\bar{\gamma}_i} \leq \xi_{th} \text{ for all } i \neq k\right) \\ &= \text{Prob}(\gamma_k \leq \gamma)\prod_{\substack{i=1 \\ i\neq k}}^{K} F_{\gamma_i}\left(\xi_{th}\bar{\gamma}_i\right) = \text{Prob}(\gamma_k \leq \gamma)\left(F_x\left(\xi_{th}\right)\right)^{K-1} \end{aligned} \tag{4.19}$$

and hence:

$$f_{\gamma_k|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right) = \frac{d\text{Prob}(\gamma_k \leq \gamma)}{d\gamma}\left(F_x\left(\xi_{th}\right)\right)^{K-1}$$

where $F_x(\cdot)$ is the CDF of the random variable $\frac{\gamma}{\bar{\gamma}_k}$, which can be easily computed by setting $\bar{\gamma}_k = 1$ in expressions (4.7)-(4.11). It is worth noting that in this case, the probabilities $\text{Prob}(\gamma_k \leq \gamma)$ associated with the AS-based schemes are not equal to the pre-scheduling CDF expressions, $F_{\gamma_k}(\gamma)$, obtained in Section 4.3. This is because transmit antenna subsets are randomly selected in the scheduling outage region.

On the other hand, for $\gamma_k/\bar{\gamma}_k > \xi_{th}$ the conditional CDF/pdf functions are given by:

$$\begin{aligned}\text{Prob}\left(\gamma_k \leq \gamma | \mathcal{A}_k\right) &= F_{\gamma_k}\left(\gamma|\mathcal{A}_k\right) \\ &= \text{Prob}(\gamma_k \leq \gamma)\text{Prob}\left(\frac{\gamma_i}{\bar{\gamma}_i} \leq \frac{\gamma}{\bar{\gamma}_k} \text{ for all } i \neq k\right) \\ &= F_{\gamma_k}(\gamma)\prod_{\substack{i=1 \\ i\neq k}}^{K} F_{\gamma_i}\left(\frac{\gamma}{\bar{\gamma}_k}\bar{\gamma}_i\right) = \left(F_{\gamma_k}(\gamma)\right)^K \\ f_{\gamma_k|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right) &= K f_{\gamma_k}(\gamma)\left(F_{\gamma_k}\left(\gamma\right)\right)^{K-1}\end{aligned}$$

Therefore, the expressions derived in Section 4.4.1 are still valid for this second region.

Next, since the post-scheduling pdf can be written as

$$f_{\gamma^*}(\gamma) = \frac{1}{K}\sum_{k=1}^{K} f_{\gamma_k|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right)$$

the conditioned pdf expressions, $f_{\gamma_k|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right)$, are obtained for the different transmission schemes:

- *SISO:*

For the $\gamma_k/\bar{\gamma}_k \leq \xi_{th}$ case, we have:

$$F_{\gamma_{k,SISO}|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right) = F_{\gamma_{k,SISO}}(\gamma)\left(F_{x,SISO}\left(\xi_{th}\right)\right)^{K-1}$$

and by differentiating the expression above:

$$f_{\gamma_k,SISO|\mathcal{A}_k}(\gamma|\mathcal{A}_k) = \frac{e^{-\frac{\gamma}{\bar{\gamma}_k}}}{\bar{\gamma}_k}\left(1-e^{-\xi_{th}}\right)^{K-1} \tag{4.20}$$

whereas for $\gamma_k/\bar{\gamma}_k > \xi_{th}$, the conditional CDF/pdf functions are given by:

$$F_{\gamma_k,SISO|\mathcal{A}_k}(\gamma|\mathcal{A}_k) = \left(F_{\gamma_k,SISO}(\gamma)\right)^K$$

$$f_{\gamma_k,SISO|\mathcal{A}_k}(\gamma|\mathcal{A}_k) = K\frac{e^{-\frac{\gamma}{\bar{\gamma}_k}}}{\bar{\gamma}_k}\left(1-e^{-\frac{\gamma}{\bar{\gamma}_k}}\right)^{K-1} \tag{4.21}$$

- *SISO-AS:*

In this case, the conditional pdf and CDF expressions can be readily obtained by simply replacing the actual number of users K by KM in the SISO expressions. This holds true for both SNR regions, $\gamma_k/\bar{\gamma}_k \leq \xi_{th}$ and $\gamma_k/\bar{\gamma}_k > \xi_{th}$ since, in the former case, both the scheduled user *and* the BS transmit antenna are randomly selected.

- *OSTBC:*

It can be readily shown that, for the Alamouti scheme the conditional pdf expressions can be written as:

$$f_{\gamma_k,OSTBC|\mathcal{A}_k}(\gamma|\mathcal{A}_k) = \frac{4\gamma}{\bar{\gamma}_k^2}e^{-\frac{2\gamma}{\bar{\gamma}_k}}\left(1-e^{-2\xi_{th}}(2\xi_{th}+1)\right)^{K-1} \qquad \gamma_k/\bar{\gamma}_k \leq \xi_{th} \tag{4.22}$$

$$f_{\gamma_k,OSTBC|\mathcal{A}_k}(\gamma|\mathcal{A}_k) = K\frac{4\gamma}{\bar{\gamma}_k^2}e^{-\frac{2\gamma}{\bar{\gamma}_k}}\left(1-e^{-\frac{2\gamma}{\bar{\gamma}_k}}\left(\frac{2\gamma}{\bar{\gamma}_k}+1\right)\right)^{K-1} \qquad \gamma_k/\bar{\gamma}_k > \xi_{th} \tag{4.23}$$

- *OSTBC-AS:*

Since transmit antenna subsets are no longer statistically independent, the *post-scheduling* SNR analysis (for the $\gamma_k/\bar{\gamma}_k \leq \xi_{th}$ region) is somewhat involved and, thus, will be omitted (see Appendix 4.B for details). The resulting expressions for the pdf function in both regions are:

$$f_{\gamma_k,OSTBC-AS|\mathcal{A}_k}(\gamma|\mathcal{A}_k) = \frac{4}{\bar{\gamma}_k}e^{-2\xi_{th}}\left(e^{-\frac{\gamma}{\bar{\gamma}_k}}+\frac{\gamma}{\bar{\gamma}_k}e^{-2\frac{\gamma-\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k}}-1\right) \times\left(1-e^{-2\xi_{th}}\left(6\xi_{th}+4e^{-\xi_{th}}-3\right)\right)^{K-1} \qquad \gamma_k/\bar{\gamma}_k \leq \xi_{th} \tag{4.24}$$

$$f_{\gamma_k,OSTBC-AS|\mathcal{A}_k}(\gamma|\mathcal{A}_k) = K\frac{12}{\bar{\gamma}_k}e^{-\frac{2\gamma}{\bar{\gamma}_k}}\left(e^{-\frac{\gamma}{\bar{\gamma}_k}}+\frac{\gamma}{\bar{\gamma}_k}-1\right) \times\left(1-e^{-\frac{2\gamma}{\bar{\gamma}_k}}\left(\frac{6\gamma}{\bar{\gamma}_k}+4e^{-\frac{\gamma}{\bar{\gamma}_k}}-3\right)\right)^{K-1} \qquad \gamma_k/\bar{\gamma}_k > \xi_{th} \tag{4.25}$$

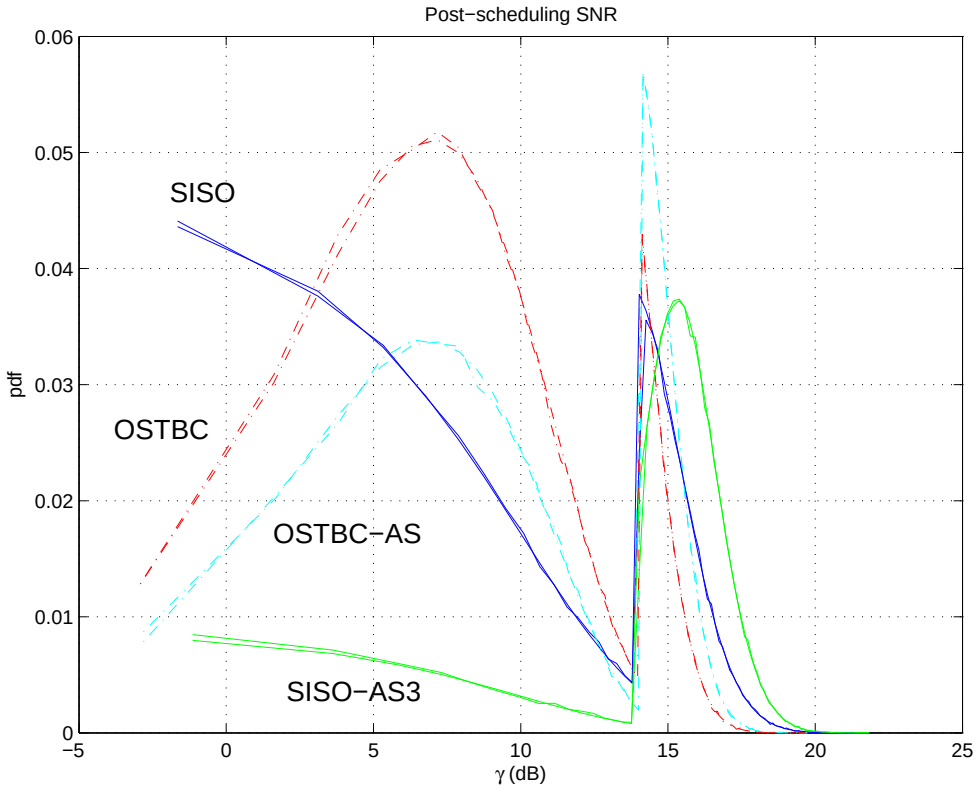


Figure 4.7: Analytical and simulated pdf of the post-scheduling SNR for the different transmission schemes in a selective-feedback scenario ($\bar{\gamma} = 10$ dB, $10\log_{10}(\xi_{th}\bar{\gamma}) = 14$ dB, K=10 users).

In Fig. 4.7, we depict the pdf of the post-scheduling SNR for the different transmission schemes. Apart from a close matching between analytical and simulated curves, we can appreciate in the pdf behavior a well-defined transition between random ($\gamma/\bar{\gamma} \leq \xi_{th}$) and max-normalized SNR scheduling regions. In the $\gamma/\bar{\gamma} \leq \xi_{th}$ region, one would emphasize the enhanced fading mitigation capabilities exhibited by the OSTBC-based schemes (reduced probability at low-SNR values). In the max-SNR scheduling region, instead, it is worth noting the improvement in terms of SNR peak generation (increased probability at high-SNR values) obtained with the SISO-based approaches.

4.5.2 Normalized Average Feedback Load

In [Ges04], the authors define *normalized average feedback load*, $\bar{F}$, as the usage ratio per time unit averaged over the total number of active users. This measure can also be interpreted as the probability for a given user to effectively signal its CSI over the feedback channel. Thus, in the proposed scheme where users report CSI when $\gamma_k/\bar{\gamma}_k > \xi_{th}$, $\bar{F}$ can be readily expressed as:

$$\bar{F} = \frac{1}{K}\sum_{k=1}^{K}(1 - F_{\gamma_k}(\xi_{th}\bar{\gamma}_k)) = 1 - F_x(\xi_{th})$$

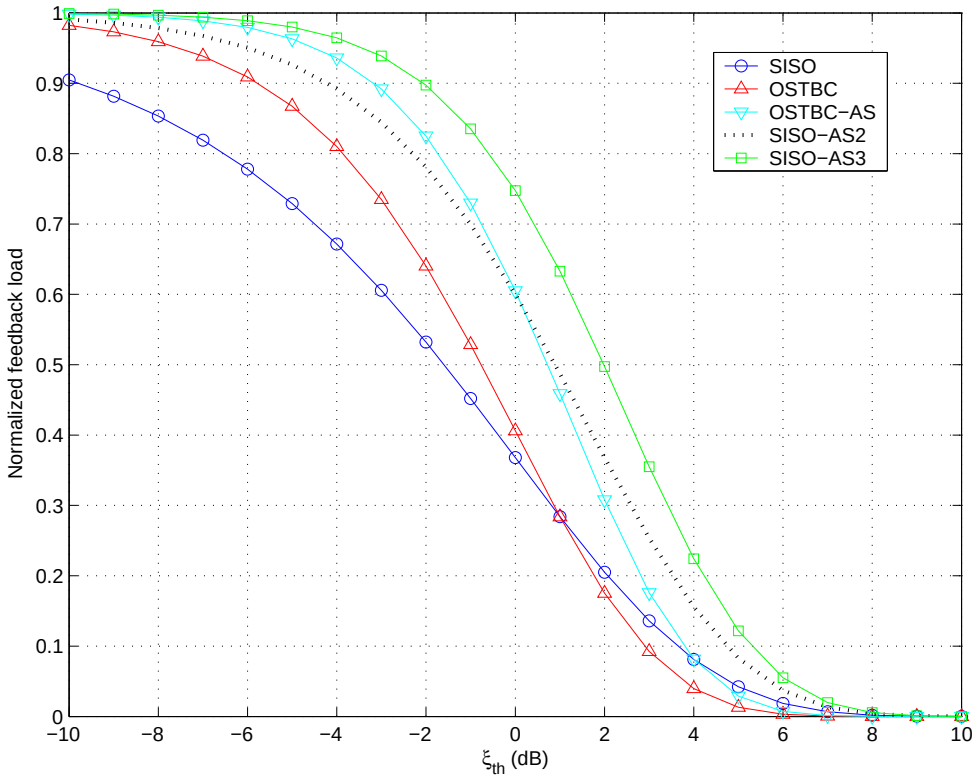


Figure 4.8: Normalized feedback load vs. normalized threshold ξ_{th} for the different transmission schemes.

For a given feedback load, the associated SNR thresholds, that will ultimately depend on the selected transmission scheme, can then be obtained from Eqs. (4.7), (4.8), (4.10) and (4.11) in Section 4.3. This is straightforward in the SISO case ($\xi_{th} = -\ln(\bar{F})$) but for the other cases one should resort to numerical methods.

As shown in Fig. 4.8, the stabilizing effect of OSTBC with respect to a SISO configuration has a clear impact on the determination of the normalized threshold. For a significantly high feedback load ($\bar{F} \geq 0.28$) the threshold is higher in the OSTBC case because of the reduction in the number of deep fades (e.g. the effective SNR is higher and, for a given feedback load, so is the threshold). Conversely, for low feedback loads where only the highest SNR peaks can be exploited by the Selective MUD schemes, the threshold associated to OSTBC must be lower (w.r.t. that of SISO) since the number of such peaks is substantially lower too. By introducing AS mechanisms (OSTBC-AS) the normalized thresholds can be increased again due to the associated SNR peak restoration (for $\bar{F} \geq 0.08$) and fade suppression effects. Finally, one can also observe how the fade mitigation capability of the SISO scheme can be improved with the introduction of antenna selection mechanisms. As the number of available antennas for selection increases (SISO-AS2, SISO-AS3) so do the associated normalized thresholds.

4.5.3 Ergodic Capacity Analysis

In the proposed selective-feedback scenario, the partial CSI is only available at the transmitter when at least one user reports its channel status. Then, rate adaptation can only be carried out in the max-normalized SNR region. Nevertheless, if long codewords are considered, the fading effect associated to the *outage scheduling* slots can be averaged by means of constant rate transmission. In order to obtain reliable transmissions, codewords must be long enough to average all the possible fading realizations and must be transmitted with a rate lower or equal to the ergodic capacity, which is computed as follows in this case:

$$\overline{C} = \frac{1}{K}\sum_{k=1}^{K}\left[\int_{\gamma=0}^{\xi_{th}\bar{\gamma}_k}\log_2\left(1+\gamma\right)f_{\gamma_k|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right)d\gamma + \int_{\gamma=\xi_{th}\bar{\gamma}_k}^{\infty}\log_2\left(1+\gamma\right)f_{\gamma_k|\mathcal{A}_k}\left(\gamma|\mathcal{A}_k\right)d\gamma\right]$$

Notice that, we have also considered the transmission of constant rate codewords in the max-normalized SNR region. This is motivated by the fact that with this simpler transmission design, the same average performance can be obtained with respect to the rate adaptation case as we consider that the system is not delay-constrained.

By plugging the pdf functions obtained in Section 4.5.1 into the above expression, the ergodic capacity for the different transmission schemes can be obtained. By resorting to the binomial expansion the following equation can be obtained for the SISO and SISO-AS cases:

$$\overline{C}_{SISO}\left(K\right) = \frac{\log_2(e)}{K}\sum_{k=1}^{K}\left[\left(1-e^{-\xi_{th}}\right)^{K-1}\left[e^{\frac{1}{\bar{\gamma}_k}}\left(E_i\left(-\frac{1+\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k}\right)-E_i\left(-\frac{1}{\bar{\gamma}_k}\right)\right) - e^{-\xi_{th}}\ln(1+\xi_{th}\bar{\gamma}_k)\right] + K\log_2(e)\sum_{i=0}^{K-1}\binom{K-1}{i}\frac{(-1)^i}{i+1} \times\left[e^{-\xi_{th}(i+1)}\ln(1+\xi_{th}\bar{\gamma}_k) - e^{\frac{i+1}{\bar{\gamma}_k}}E_i\left(-\frac{1+\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k}(i+1)\right)\right]\right]$$

$$\overline{C}_{SISO-AS}\left(K\right) = \frac{\log_2(e)}{K}\sum_{k=1}^{K}\left[\left(1-e^{-\xi_{th}}\right)^{MK-1}\left[e^{\frac{1}{\bar{\gamma}_k}}\left(E_i\left(-\frac{1+\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k}\right)-E_i\left(-\frac{1}{\bar{\gamma}_k}\right)\right) - e^{-\xi_{th}}\ln(1+\xi_{th}\bar{\gamma}_k)\right] + MK\log_2(e)\sum_{i=0}^{MK-1}\binom{MK-1}{i}\frac{(-1)^i}{i+1} \times\left[e^{-\xi_{th}(i+1)}\ln(1+\xi_{th}\bar{\gamma}_k) - e^{\frac{i+1}{\bar{\gamma}_k}}E_i\left(-\frac{1+\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k}(i+1)\right)\right]\right]$$

As for the OSTBC scheme, the following result holds:

$$\overline{C}_{OSTBC}(K) = \frac{\log_2(e)}{K}\sum_{k=1}^{K}\left[\frac{4}{\bar{\gamma}_k^2}\left(1-e^{-2\xi_{th}}(2\xi_{th}+1)\right)^{K-1}\Phi\left(\xi_{th}\bar{\gamma}_k,2,\frac{2}{\bar{\gamma}_k}\right)\right.$$
$$\left.+K\log_2(e)\sum_{i=0}^{K-1}\binom{K-1}{i}(-1)^i\sum_{n=0}^{i}\binom{i}{n}\left(\frac{2}{\bar{\gamma}_k}\right)^{n+2}\Upsilon\left(\xi_{th}\bar{\gamma}_k,n+2,\frac{2(i+1)}{\bar{\gamma}_k}\right)\right]$$

where $\Phi(a,m,\mu)$ is also an integral equation (further details can be found in Appendix 4.A). When AS mechanisms are introduced, by repeatedly applying the binomial expansion on Eqs.(4.24) and (4.25), one can verify that the ergodic capacity can be written as follows:

$$\overline{C}_{OSTBC-AS}(K) = \frac{\log_2(e)}{K}\sum_{k=1}^{K}\left[\frac{4}{\bar{\gamma}_k}e^{-2\xi_{th}}\left(1-e^{-2\xi_{th}}\left(6\xi_{th}+4e^{-\xi_{th}}-3\right)\right)^{K-1}\right.$$
$$\times\left[\Phi\left(\xi_{th}\bar{\gamma}_k,1,\frac{1}{\bar{\gamma}_k}\right)+\frac{e^{2\xi_{th}}}{\bar{\gamma}_k}\Phi\left(\xi_{th}\bar{\gamma}_k,2,\frac{2}{\bar{\gamma}_k}\right)-(1+\xi_{th}\bar{\gamma}_k)\ln(\xi_{th}\bar{\gamma}_k)+\xi_{th}\bar{\gamma}_k-\frac{\xi_{th}^2\bar{\gamma}_k^2}{2}\right]$$
$$+K12\sum_{i=0}^{K-1}\binom{K-1}{i}3^i\sum_{n=0}^{i}\binom{i}{n}\left(-\frac{1}{3}\right)^n\sum_{s=0}^{n}\binom{n}{s}\frac{6^{n-s}4^s}{\bar{\gamma}_k^{n-s+1}}\left[\Upsilon\left(\xi_{th}\bar{\gamma}_k,n-s+1,\frac{2i+s+3}{\bar{\gamma}_k}\right)\right.$$
$$\left.\left.+\frac{1}{\bar{\gamma}_k}\Upsilon\left(\xi_{th}\bar{\gamma}_k,n-s+2,\frac{2i+s+2}{\bar{\gamma}_k}\right)-\Upsilon\left(\xi_{th}\bar{\gamma}_k,n-s+1,\frac{2i+s+2}{\bar{\gamma}_k}\right)\right]\right]$$

Notice that in the derived expressions the first term in the summation accounts for contributions to capacity due to random scheduling (i.e. in the case of *scheduling outage*), whereas the second term reflects contributions coming from max-normalized SNR scheduling.

4.5.4 System Throughput Analysis

Concerning the throughput case, transmitting with a constant rate is not a useful strategy. As commented in the previous section, we consider a scenario with practical modulation schemes, finite complexity receivers and uncoded transmission. Then, we cannot apply the concept of long codewords transmission as in the ergodic capacity case. Instead, we adopt a transmission strategy where the lowest constellation size is selected when a scheduling outage occurs, whereas adaptive modulation is encompassed in the opposite case.

With the proposed transmission strategy, the average throughput expression can be re-written as follows:

$$\overline{\eta}(K) \approx \frac{1}{K}\sum_{k=1}^{K}\left[b_1\int_{\gamma=0}^{\xi_{th}\bar{\gamma}_k}(1-\alpha_1 e^{-\beta_1\gamma})^L f_{\gamma_k|\mathcal{A}_k}(\gamma|\mathcal{A}_k)\,d\gamma\right.$$
$$\left.+\sum_{m=1}^{\mathrm{card}(\mathcal{M})}b_m\int_{\gamma=\max(\xi_{th}\bar{\gamma}_k,\gamma_{th,m})}^{\max(\xi_{th}\bar{\gamma}_k,\gamma_{th,m+1})}(1-\alpha_m e^{-\beta_m\gamma})^L f_{\gamma_k|\mathcal{A}_k}(\gamma|\mathcal{A}_k)\,d\gamma\right] \qquad (4.26)$$

Again, by substituting Eqs.(4.20)-(4.25) into (4.26), one can particularize the above expression to the different transmission schemes. For the SISO approach we have:

$$\begin{aligned}\overline{\eta}_{SISO}(K) \approx \frac{1}{K}\sum_{k=1}^{K}\Bigg[& b_1\left(1-e^{-\xi_{th}}\right)^{K-1}\sum_{l=0}^{L}\binom{L}{l}\frac{(-\alpha_1)^l}{\bar{\gamma}_k\beta_1 l+1}\left[1-e^{-\xi_{th}(\bar{\gamma}_k\beta_1 l+1)}\right] \\ & + K\sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \sum_{i=0}^{K-1}\binom{K-1}{i}(-1)^i\sum_{l=0}^{L}\binom{L}{l}\frac{(-\alpha_m)^l}{\bar{\gamma}_k\beta_m l+i+1} \\ & \times\left[e^{-\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+i+1)}-e^{-\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+i+1)}\right]\Bigg]\end{aligned}$$

A similar expression is obtained for the SISO-AS scheme:

$$\begin{aligned}\overline{\eta}_{SISO-AS}(K) \approx \frac{1}{K}\sum_{k=1}^{K}\Bigg[& b_1\left(1-e^{-\xi_{th}}\right)^{MK-1}\sum_{l=0}^{L}\binom{L}{l}\frac{(-\alpha_1)^l}{\bar{\gamma}_k\beta_1 l+1}\left[1-e^{-\xi_{th}(\bar{\gamma}_k\beta_1 l+1)}\right] \\ & + MK\sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \sum_{i=0}^{MK-1}\binom{MK-1}{i}(-1)^i\sum_{l=0}^{L}\binom{L}{l}\frac{(-\alpha_m)^l}{\bar{\gamma}_k\beta_m l+i+1} \\ & \times\left[e^{-\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+i+1)}-e^{-\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+i+1)}\right]\Bigg]\end{aligned}$$

where we have defined the parameters c and d as:

$$\begin{aligned} c &= \max(\xi_{th}\bar{\gamma}_k, \gamma_{th,m}) \\ d &= \max(\xi_{th}\bar{\gamma}_k, \gamma_{th,m+1})\end{aligned}$$

Regarding the OSTBC transmission scheme, the following expressions hold:

$$\begin{aligned}\overline{\eta}_{OSTBC}(K) \approx \frac{1}{K}\sum_{k=1}^{K}\Bigg[& 4b_1\left(1-e^{-2\xi_{th}}(2\xi_{th}+1)\right)^{K-1}\sum_{l=0}^{L}\binom{L}{l}\frac{(-\alpha_1)^l}{(\bar{\gamma}_k\beta_1 l+2)^2} \\ & \times\left[1-e^{-\xi_{th}(\bar{\gamma}_k\beta_1 l+2)}(\xi_{th}(\bar{\gamma}_k\beta_1 l+2)+1)\right] \\ & + 4K\sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m\sum_{i=0}^{K-1}\binom{K-1}{i}(-1)^i\sum_{l=0}^{L}\binom{L}{l}(-\alpha_m)^l\sum_{n=0}^{i}\binom{i}{n}\frac{2^n}{(\bar{\gamma}_k\beta_m l+2i+2)^{n+2}} \\ & \times\left[\Gamma(n+2,\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+2))-\Gamma(n+2,\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+2))\right]\Bigg]\end{aligned}$$

whereas the following result is obtained for the OSTBC-AS case:

$$\overline{\eta}_{OSTBC-AS}(K) \approx \frac{1}{K}\sum_{k=1}^{K}\Bigg[-4b_1 e^{-2\xi_{th}}\left(1-e^{-2\xi_{th}}\left(6\xi_{th}+4e^{-\xi_{th}}-3\right)\right)^{K-1}\Bigg[\frac{d-c}{\bar{\gamma}_k}$$
$$+\sum_{l=1}^{L}\binom{L}{l}(-\alpha_1)^l\frac{e^{-c\beta_1 l}-e^{-d\beta_1 l}}{\bar{\gamma}_k\beta_1 l}-\sum_{l=0}^{L}\binom{L}{l}(-\alpha_1)^l\Bigg(\frac{e^{-\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_1 l+1)}-e^{-\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_1 l+1)}}{\bar{\gamma}_k\beta_1 l+1}$$
$$+\frac{e^{2\xi_{th}}}{(\bar{\gamma}_k\beta_1 l+2)^2}\left(e^{-\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_1 l+2)}\left(\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_1 l+2)+1\right)-e^{-\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_1 l+2)}\left(\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_1 l+2)+1\right)\right)\Bigg)\Bigg]$$
$$+12K\sum_{m=1}^{\mathrm{card}(\mathcal{M})}b_m\sum_{i=0}^{K-1}\binom{K-1}{i}3^i\sum_{l=0}^{L}\binom{L}{l}(-\alpha_m)^l\sum_{n=0}^{i}\binom{i}{n}\left(-\frac{1}{3}\right)^n\sum_{s=0}^{n}\binom{n}{s}6^{n-s}4^s$$
$$\Bigg[\frac{\Gamma(n-s+1,\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+s+3))-\Gamma(n-s+1,\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+s+3))}{(\bar{\gamma}_k\beta_m l+2i+s+3)^{n-s+1}}$$
$$+\frac{1}{(\bar{\gamma}_k\beta_m l+2i+s+2)^{n-s+1}}\Bigg(\frac{\Gamma(n-s+2,\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+s+2))-\Gamma(n-s+2,\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+s+2))}{\bar{\gamma}_k\beta_m l+2i+s+2}$$
$$-\Gamma(n-s+1,\frac{c}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+s+2))+\Gamma(n-s+1,\frac{d}{\bar{\gamma}_k}(\bar{\gamma}_k\beta_m l+2i+s+2))\Bigg)\Bigg]\Bigg]$$

Asymptotic Analysis

In order to gain some insight into the throughput expressions above, we will force the feedback channel bandwidth to be arbitrarily low (i.e. $\xi_{th}\to\infty$) and conduct the corresponding asymptotic analysis in an homogenous scenario. Bearing in mind that $\lim_{x\to\infty}\Gamma(n,x)=0$ we can easily prove that:

$$\lim_{\xi_{th}\to\infty}\overline{\eta}_{SISO-AS}=\lim_{\xi_{th}\to\infty}\overline{\eta}_{SISO}=b_1\sum_{l=0}^{L}\binom{L}{l}\frac{(-\alpha_1)^l}{\bar{\gamma}\beta_1 l+1} \tag{4.27}$$

$$\lim_{\xi_{th}\to\infty}\overline{\eta}_{OSTBC-AS}=\lim_{\xi_{th}\to\infty}\overline{\eta}_{OSTBC}=4b_1\sum_{l=0}^{L}\binom{L}{l}\frac{(-\alpha_1)^l}{(\bar{\gamma}\beta_1 l+2)^2} \tag{4.28}$$

Two main conclusions can be drawn from the expressions above. First, the inclusion of AS makes no difference in the asymptotic regime since, clearly, bandwidth constraints in the feedback channel will prevent the scheduler from being informed on the optimal antenna subset. Besides, MUD can be barely exploited in this context (the system will often resort to random scheduling) and, consistently, there is no dependence in Eqs. (4.27), (4.28) on the number of active users in the system.

Going one step further, we will assume a large enough average SNR, $\bar{\gamma}$, and will approximate

the summations in Eq. (4.27) and (4.28) by their first two elements:

$$\lim_{\xi_{th}\to\infty} \overline{\eta}_{SISO} \approx b_1 \left(1 - \frac{\alpha_1 L}{\bar{\gamma}\beta_1 + 1}\right)$$
$$\lim_{\xi_{th}\to\infty} \overline{\eta}_{OSTBC} \approx b_1 \left(1 - \frac{4\alpha_1 L}{(\bar{\gamma}\beta_1 + 2)^2}\right)$$

Finally, for positive values of the $\beta_1^2\bar{\gamma}^2$ product (as in our case), one can easily verify after some manipulations that

$$\frac{4}{(\bar{\gamma}\beta_1 + 2)^2} < \frac{1}{\bar{\gamma}\beta_1 + 1} \tag{4.29}$$

i.e., in terms of average throughput OSTBC-based schemes outperform SISO-based ones in the asymptotic case. This conclusion is not so surprising since a multi-user scenario with very low feedback channel bandwidth and high SNR is equivalent to a single-user case where, ultimately, it is the spatial diversity gain provided by space-time block coding (BER curves' slope in the high SNR regime) that makes the difference. Simulation results shown in the next section (Fig. 4.10) are aligned with this and, thus, confirm the asymptotic results obtained in this section.

4.5.5 Numerical Results and Discussion

In Fig. 4.9, we depict the ergodic *capacity* as a function of the number of active users and different feedback loads[5] ($\bar{F} = 1, ..., 0.01$). As in [Ges04], performance loss in selective-feedback system with $\bar{F} \geq 0.2$ and a moderate-to-high number of users ($K \geq 15$) can be neglected for all transmission schemes. When the average feedback load per user is further reduced ($\bar{F} = 0.05$), one can observe that the degradation experienced by the SISO-based schemes is larger than that exhibited by the OSTBC ones (approximately 24% vs. 16% respectively for a $K = 10$ case). Furthermore, for $\bar{F} = 0.01$, OSTBC approaches outperform *both* SISO schemes. As a conclusion, when the scheduling outage probability increases (due to bandwidth limitations in the feedback channel), OSTBC approaches provide additional robustness against deep fades associated to random user selection. In all cases, additional gains in terms of spatial diversity can be obtained from the introduction of AS mechanisms (albeit very moderate when feedback load is low).

As for the average system *throughput*, curves in the left-hand plot of Fig. 4.10 reveal that, for a scenario with QPSK and selective-feedback ($\bar{F} < 1$), the robustness against wrong scheduling decisions associated to OSTBC clearly exceed the multi-user diversity gains. This is true even for cases with a moderate scheduling outage probability (e.g. $\bar{F} = 0.2$). The diversity order can be further increased by introducing AS mechanisms but the improvement in terms of system throughput is rather limited due to saturation effects (as opposed to capacity enhancements). Further, throughput curves reveal that the gap between the SISO- and OSTBC-based approaches

[5]For a given feedback load, the normalized threshold ξ_{th} for each transmission scheme must be accordingly adjusted.

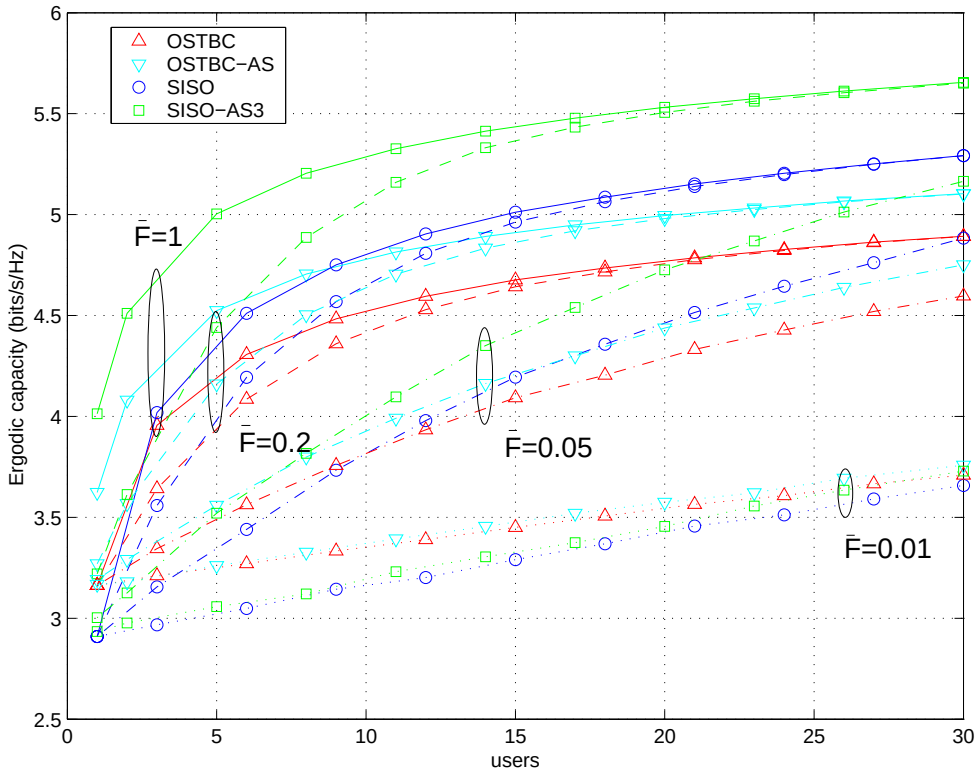


Figure 4.9: Ergodic capacity vs. number of users for the different transmission schemes and feedback loads $\bar{F}$=1 (solid), 0.2 (dashed), 0.05 (dash-dotted), 0.01 (dotted). (symbols: simulated results, curves: analytical expressions, $\bar{\gamma} = 10$ dB).

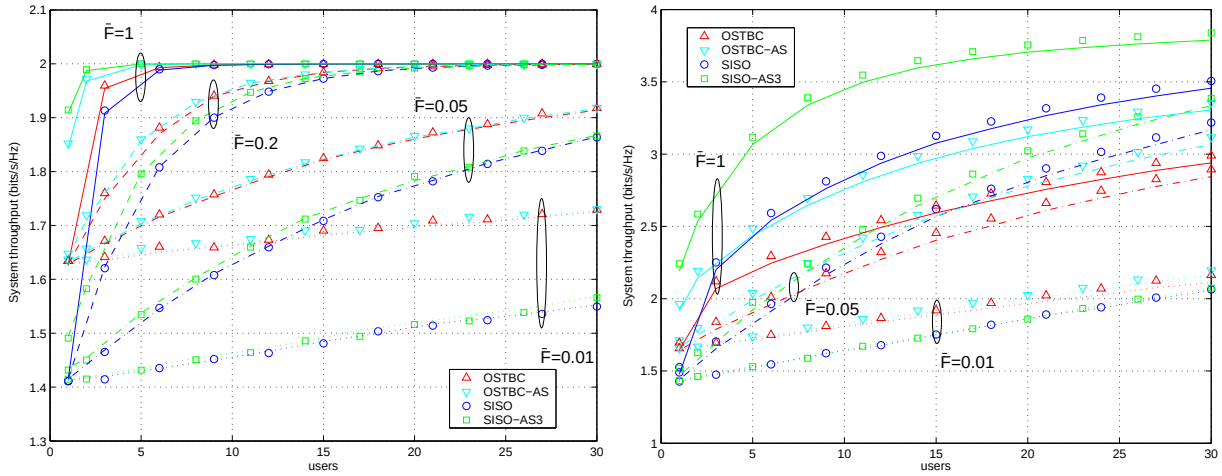


Figure 4.10: System throughput vs. number of users for the different transmission schemes and feedback loads (symbols: simulated results, curves: analytical expressions, L=10 symbols, $\bar{\gamma} = 10$ dB). Left, QPSK. Right, adaptive modulation.

becomes wider when feedback load is reduced. In other words, a OSTBC transmission scheme is far less sensitive to restrictions and constraints in the feedback channel. However, by incorporating adaptive modulation, the saturation of the system is alleviated and a greater improvement can be obtained by using AS mechanisms. This is observed in the right-hand plot of Fig. 4.10 where results corresponding to a system using adaptive modulation are shown. In particular, one can observe that MUD gains can be efficiently extracted with adaptive modulation as long as the feedback load is not substantially limited.

4.6 Analysis in a Scenario with Degraded CSI

In this section, we consider again that all the users continuously report their channel status to the BS (i.e., no SMUD schemes) but the partial CSI received at the scheduler is degraded. That is, the SNR available at the BS, $\hat{\gamma}_k$, differs from the actual SNR, γ_k.

The scheduling strategy is based on the max-normalized SNR, being the scheduling decisions based on the available SNR, $\hat{\gamma}_k$, that is,

$$k^* = \arg\max_k \left\{ \frac{\hat{\gamma}_1}{\mathbb{E}[\hat{\gamma}_1]}, ..., \frac{\hat{\gamma}_k}{\mathbb{E}[\hat{\gamma}_k]}, ..., \frac{\hat{\gamma}_K}{\mathbb{E}[\hat{\gamma}_K]} \right\}$$

Then, according to the degree of degradation in the CSI, there may be some situations where system performance is considerably deteriorated with respect to the non-degraded case. In this section, we conduct an analytical study in order to asses the impact of degraded CSI at the scheduler on SISO and OSTBC-based schemes. As shown in the previous section, antenna selection mechanisms give very moderate gains when the CSI information available at the BS is subject to restrictions. For that reason, we focus our attention on the analysis of the transmission schemes without AS. Nonetheless, we do not restrict ourselves to a specific source of imperfections. Instead, we adopt a general statistical approach to its modeling. In order to gain some insight, we present two practical examples: delayed feedback channel and imperfect channel estimation. By doing so, we analytically show that OSTBC approaches are also useful for exploiting MUD in scenarios with degraded CSI.

4.6.1 Modeling Uncertainty in the CSI

In this section, we derive a statistical model describing the degree of CSI imperfection at the BS. In particular, we consider the estimated SNR at the transmitter to be obtained from a channel vector gain, $\hat{\mathbf{h}}_k$, which differs from the actual channel response, $\mathbf{h}_k$; these two random variables being related with a Gaussian model. In other words, we assume that $\mathbf{h}_k$ conditioned on $\hat{\mathbf{h}}_k$ follows a Gaussian distribution:

$$\mathbf{h}_k|\hat{\mathbf{h}}_k \sim \mathcal{CN}(\vartheta_k \hat{\mathbf{h}}_k, \mathbf{\Sigma_k}) \tag{4.30}$$

where $\vartheta_k\hat{\mathbf{h}}_k$ and $\mathbf{\Sigma_k} = \sigma_{h_k}^2\sigma_{\epsilon_k}^2\mathbf{I}$ are the mean and covariance matrix, respectively. Notice that parameters ϑ_k and $\sigma_{\epsilon_k}^2$ are used to model different source of impairments (some examples are given in Section 4.6.5). As a case in point, $\vartheta_k = 1$ and $\sigma_{\epsilon_k}^2 = 0$ when the actual channel response can be perfectly known from the estimate, whereas $\vartheta_k = 0$ and $\sigma_{\epsilon_k}^2 = 1$ when the estimate does not give any information about the actual channel (i.e., $\mathbf{h}_k|\hat{\mathbf{h}}_k \sim \mathbf{h}_k$). Admittedly, this model might not be very accurate for some sources of imperfection but, still, it is very helpful in the analysis and design of communication schemes because of its mathematical tractability [Zho05].

Under those assumptions, it is straightforward to show from Eq. (4.30) that the actual SNR, γ_k, conditioned on its estimate, $\hat{\gamma}_k$, follows a non-central chi-square distribution with $2M$ degrees of freedom [Pro95]:

$$f_{\gamma_k|\hat{\gamma}_k}(\gamma|\hat{\gamma}_k) = \frac{M}{\bar{\gamma}_k\sigma_{\epsilon_k}^2}\left(\frac{\gamma}{\vartheta_k^2\hat{\gamma}_k}\right)^{\frac{2M-2}{4}} e^{-\frac{M(\gamma+\vartheta_k^2\hat{\gamma}_k)}{\bar{\gamma}_k\sigma_{\epsilon_k}^2}} \times I_{M-1}\left(\frac{2M\sqrt{\vartheta_k^2\gamma\hat{\gamma}_k}}{\bar{\gamma}_k\sigma_{\epsilon_k}^2}\right) \tag{4.31}$$

with $I_n(\cdot)$ standing for the nth-order modified Bessel function of the first kind. In Fig. 4.11, we plot the above pdf expression for $\vartheta_k = 1$ and different values of σ_{ϵ_k}. As it can be observed, when the value of σ_{ϵ_k} is decreased, more accurate can be the knowledge of the actual channel response as there is an increasing probability that γ_k takes a value close to $\hat{\gamma}_k$. Indeed, the pdf is equal to a delta function centered at $\gamma = \hat{\gamma}_k$ in the extreme case $\sigma_{\epsilon_k} = 0$, i.e., the actual channel response is equal to the estimated one with probability one.

4.6.2 Post-Scheduling SNR

As in the previous sections, the CDF of the post-scheduling SNR can be written as:

$$F_{\gamma^*}(y) = \sum_{k=1}^{K} \text{Prob}\left(\gamma_k \leq y|\mathcal{A}_k\right)\text{Prob}\left(\mathcal{A}_k\right) \tag{4.32}$$

In this case, the probability $\text{Prob}\left(\gamma_k \leq y|\mathcal{A}_k\right)$ can be expressed in terms of the estimated SNR available at the scheduler:

$$\text{Prob}\left(\gamma_k \leq y|\mathcal{A}_k\right) = \int_{\hat{\gamma}=0}^{\infty} F_{\gamma_k|\hat{\gamma}_k}(y|\hat{\gamma}_k) f_{\hat{\gamma}_k|\mathcal{A}_k}(\hat{\gamma}_k|\mathcal{A}_k) d\hat{\gamma}_k \tag{4.33}$$

where $F_{\gamma_k|\hat{\gamma}_k}(y|\hat{\gamma}_k)$ can be computed with the help of the conditional pdf expression obtained in Eq. (4.31):

$$F_{\gamma_k|\hat{\gamma}_k}(y|\hat{\gamma}_k) = \int_{\gamma=0}^{y} f_{\gamma_k|\hat{\gamma}_k}(\gamma|\hat{\gamma}_k) d\gamma \tag{4.34}$$

and $f_{\hat{\gamma}_i|\mathcal{A}_k}(\hat{\gamma}_i|\mathcal{A}_k)$ denotes the pdf of the *estimated* post-scheduling SNR associated to user k conditioned on event $\mathcal{A}_k$. Then, by combining Eqs. (4.33) and (4.34), and plugging them into

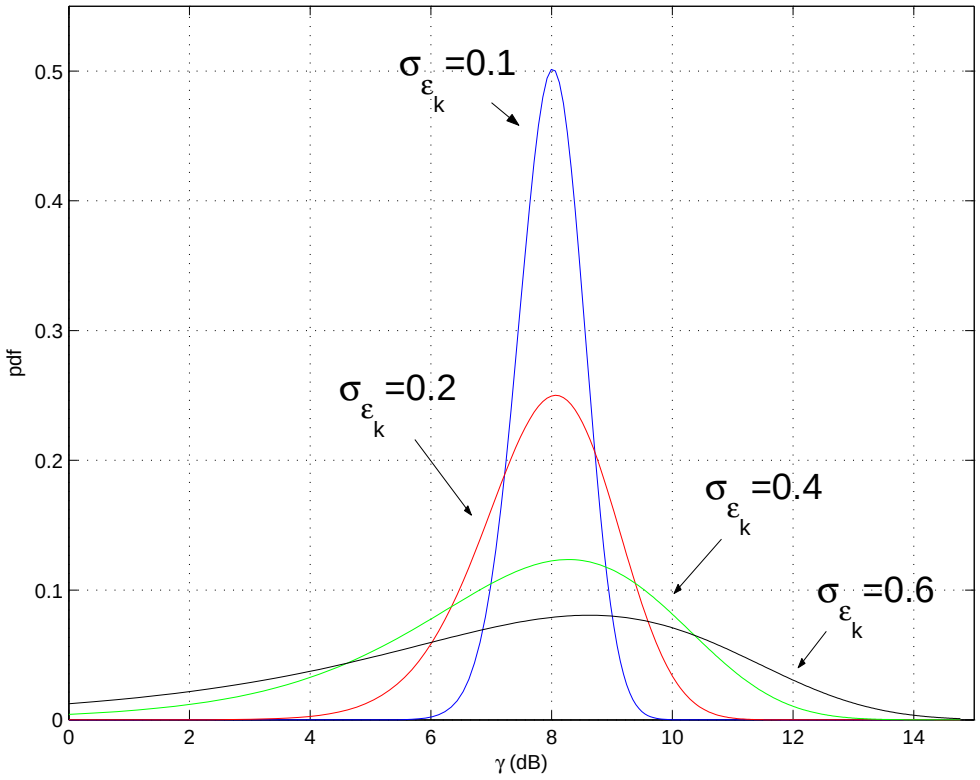


Figure 4.11: Probability density function of the actual SNR, γ_k, conditioned on its estimate, $\hat{\gamma}_k$, for different values of σ_{ϵ_k}. ($\bar{\gamma}_k$=10 dB, $\hat{\gamma}_k$=8 dB, $\vartheta_k = 1$, $M = 2$).

(4.32), we have:

$$F_{\gamma^*}(y) = \frac{1}{K}\sum_{k=1}^{K}\int_{\gamma=0}^{y}\int_{\hat{\gamma}_k=0}^{\infty} f_{\gamma_k|\hat{\gamma}_k}(\gamma|\hat{\gamma}_k) f_{\hat{\gamma}_k|\mathcal{A}_k}(\hat{\gamma}_k|\mathcal{A}_k) d\gamma d\hat{\gamma}_k \tag{4.35}$$

Before characterizing Eq. (4.35) for the SISO and OSTBC approaches, we have just to derive the conditional pdf $f_{\hat{\gamma}_k|\mathcal{A}_k}(\hat{\gamma}_k|\mathcal{A}_k)$. This can be easily done by resorting to expressions obtained in Section 4.4.1, that is:

$$f_{\hat{\gamma}_{k,SISO}|\mathcal{A}_k}(\hat{\gamma}_k|\mathcal{A}_k) = K\frac{e^{-\frac{\hat{\gamma}_k}{\mathbb{E}[\hat{\gamma}_k]}}}{\mathbb{E}[\hat{\gamma}_k]}\left(1-e^{-\frac{\hat{\gamma}_k}{\mathbb{E}[\hat{\gamma}_k]}}\right)^{K-1} \tag{4.36}$$

$$f_{\hat{\gamma}_{k,OSTBC}|\mathcal{A}_k}(\hat{\gamma}_k|\mathcal{A}_k) = \frac{4K\hat{\gamma}_k}{\mathbb{E}[\hat{\gamma}_k]^2}e^{-\frac{2\hat{\gamma}_k}{\mathbb{E}[\hat{\gamma}_k]}}\left(1-e^{-\frac{2\hat{\gamma}_k}{\mathbb{E}[\hat{\gamma}_k]}}\left(\frac{2\hat{\gamma}_k}{\mathbb{E}[\hat{\gamma}_k]}+1\right)\right)^{K-1} \tag{4.37}$$

Last, by plugging (4.36) (or (4.37)) along with (4.31) into (4.35), the corresponding post-scheduling CDF expressions can be obtained. For the SISO approach, by using the binomial expansion and identities Eq. (6.614.3), Eq. (9.220.2) and Eq. (9.215.1) in [Gra65] one can show that:

$$F_{\gamma^*_{SISO}}(y) = \sum_{k=1}^{K}\sum_{i=0}^{K-1}\binom{K-1}{i}\frac{(-1)^i}{i+1}\left(1-e^{-\frac{y(i+1)}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma^2_{\epsilon_k}}}\right)$$

On the other hand, for the OSTBC case one should resort to identities Eq. (8.406.3), Eq. (6.643.4), Eq. (8.970.1) and Eq. (3.351.1) in [Gra65] instead and, then, the CDF can be expressed in terms of the incomplete gamma function ($\Gamma_o(n,x) \triangleq \int_0^x e^{-t}t^{n-1}dt$):

$$F_{\gamma^*_{OSTBC}}(y) = \sum_{k=1}^{K}\sum_{i=0}^{K-1}\binom{K-1}{i}(-1)^i \sum_{n=0}^{i}\binom{i}{n}\sum_{s=0}^{n}\binom{n+1}{n-s}\frac{n!\left(\mathbb{E}[\hat{\gamma}_k]\right)^s \vartheta_k^{2s}\bar{\gamma}_k^{n-s}\sigma_{\epsilon_k}^{2(n-s)}}{s!(i+1)^{s+2}(\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2)^n} \times \Gamma_o\left(s+2, \frac{-2y(i+1)}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2}\right)$$

Notice that, by solving only the inner integrals in equation (4.35), the corresponding post-scheduling pdfs can be obtained:

$$f_{\gamma^*_{SISO}}(\gamma) = \sum_{k=1}^{K}\sum_{i=0}^{K-1}\binom{K-1}{i}\frac{(-1)^i}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2}e^{-\frac{\gamma(i+1)}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2}} \tag{4.38}$$

$$f_{\gamma^*_{OSTBC}}(\gamma) = 4\sum_{k=1}^{K}\sum_{i=0}^{K-1}\binom{K-1}{i}(-1)^i\sum_{n=0}^{i}\binom{i}{n}n!\gamma\bar{\gamma}_k^n\sigma_{\epsilon_k}^{2n} \times \frac{e^{\frac{-2\gamma(i+1)}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2}}}{(\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2)^{n+2}}L_n^1\left(\frac{-2\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2\gamma}{\bar{\gamma}_k\sigma_{\epsilon_k}^2\left(\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2\right)}\right) \tag{4.39}$$

where $L_n^\alpha(x)$ is the Laguerre polynomial, $L_n^\alpha(x) = \sum_{s=0}^{n}(-1)^s\binom{n+\alpha}{n-s}\frac{x^s}{s!}$.

In order to illustrate the impact of CSI uncertainty on the post-scheduling statistics, the following example is proposed. Consider a situation where the actual and estimated channel responses of an homogenous scenario are related by means of a correlation coefficient ρ. In particular, the first order statistics $\mathbb{E}[\hat{\gamma}]$ and $\bar{\gamma}$ are identical, whereas parameters ϑ and σ_ϵ are expressed in terms of ρ as follows[6]:

$$\vartheta = \rho \qquad \sigma_\epsilon^2 = 1-\rho^2$$

The cumulative densities functions $F_{\gamma^*_{SISO}}$ and $F_{\gamma^*_{OSTBC}}$ of the proposed example are plotted in Fig. 4.12 as a function of ρ ($K = 30$ active users). When the value of the correlation coefficient decreases, curves associated with both the SISO and OSTBC approaches are shifted towards the low-SNR region due to the inefficacy of the scheduler in selecting users. However, the degradation experienced by the SISO scheme is larger than that of OSTBC. In other words, the single-antenna approach is less robust to channel uncertainty (i.e., deep fades) arising from wrong scheduling decisions. In particular, the relative ordering of both curves is reversed almost for the whole SNR range in the $\rho = 0.1$ case.

4.6.3 Ergodic Capacity Analysis

In order to obtain the ergodic capacity expression of the different transmission schemes, one should introduce the corresponding pdf expressions given by Eqs. (4.38) and (4.39) in the fol-

[6]As we will show latter, the proposed example corresponds to a scenario with delayed feedback channel.

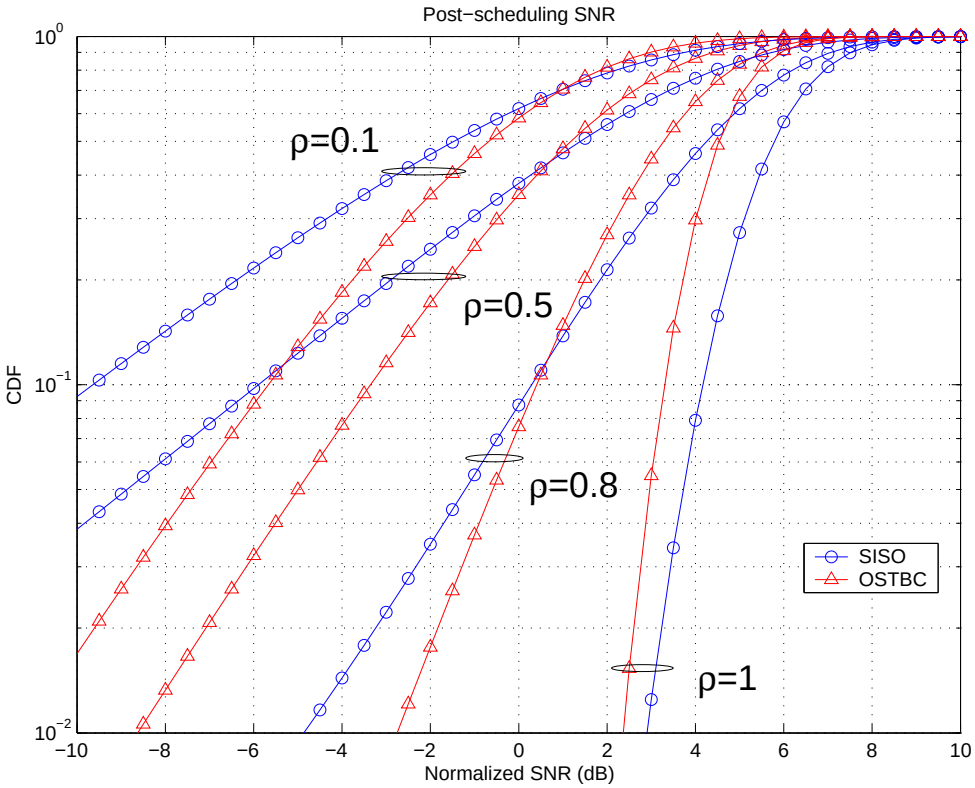


Figure 4.12: CDF of the normalized post-scheduling SNR, $\frac{\gamma}{\bar{\gamma}}$, for different values of the correlation coefficient ρ (K=30 users).

lowing equation:

$$\overline{C}\,(K) = \mathbb{E}_{\gamma}[C^*] = \int_{\gamma=0}^{\infty} \log_2(1+\gamma) f_{\gamma^*}(\gamma) d\gamma$$

As in the selective-feedback case, the ergodic capacity can be achieved when constant rate, long codewords are used with the aim of averaging the fading effect. This is because the partial CSI available at the transmitter is subject to impairments and, then, the transmitter may fail in the rate assignment when rate adaptation is encompassed.

For the SISO approach, we have

$$\overline{C}_{SISO}\,(K) = \log_2(e) \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} \frac{(-1)^i}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2 + (i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2} \times \int_{\gamma=0}^{\infty} \ln(1+\gamma) e^{-\frac{\gamma(i+1)}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2 + (i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2}} d\gamma$$

With the help of [Gra65, Eq. 4.331.2], this integral can be solved and written in closed form as:

$$\overline{C}_{SISO}\,(K) = -\log_2(e) \sum_{k=1}^{K} \sum_{i=0}^{K-1} \binom{K-1}{i} \frac{(-1)^i}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2 + (i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2} \times e^{\frac{i+1}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2 + (i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2}} E_i\left(\frac{i+1}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2 + (i+1)\bar{\gamma}\sigma_{\epsilon_k}^2}\right) \quad (4.40)$$

Analogously, for the OSTBC case we have

$$\overline{C}_{OSTBC}(K) = 4\log_2(e)\sum_{k=1}^{K}\sum_{i=0}^{K-1}\binom{K-1}{i}(-1)^i\sum_{n=0}^{i}\binom{i}{n}\sum_{s=0}^{n}\binom{n+1}{n-s}\frac{n!2^s\left(\mathbb{E}[\hat{\gamma}_k]\right)^s\vartheta_k^{2s}\bar{\gamma}_k^{n-s}\sigma_{\epsilon_k}^{2(n-s)}}{s!(\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2)^{n+s+2}} \times \Upsilon\left(0, s+2, \frac{-2(i+1)}{\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2}\right) \quad (4.41)$$

where one should recall that a closed-form expression for the integral equation $\Upsilon(a,m,\mu)$ is available in Appendix 4.A.

4.6.4 System Throughput Analysis

When the expressions related to throughput performance are derived, the post-scheduling statistics cannot be directly used. In particular, the integration intervals must be adapted to the use of adaptive modulation. By doing so, the throughput expression can be computed as:

$$\overline{\eta}(K) \approx \frac{1}{K}\sum_{k=1}^{K}\sum_{m=1}^{\text{card}(\mathcal{M})} b_m \int_{\gamma=0}^{\infty}\int_{\hat{\gamma}_k=\gamma_{th,m}}^{\gamma_{th,m+1}}(1-\alpha_m e^{-\beta_m\gamma})^L f_{\gamma_k|\hat{\gamma}_k}(\gamma|\hat{\gamma}_k) f_{\hat{\gamma}_k|\mathcal{A}_k}(\hat{\gamma}_k|\mathcal{A}_k)d\gamma d\hat{\gamma}_k \quad (4.42)$$

Notice that, it is the inner integral that depends on adaptive modulation thresholds as the adaptive modulation technique is based on the estimated SNR. Due to the use of thresholds different from 0 and ∞, the inner integral is not solvable and resorting to numerical integration is mandatory. However, we provide the reader with the closed-form expression corresponding to a system without adaptive modulation, that is, for a given modulation m:

$$\begin{aligned}\overline{\eta}_m(K) &\approx b_m\frac{1}{K}\sum_{k=1}^{K}\sum_{l=0}^{L}\binom{L}{l}(-\alpha_m)^l\int_{\gamma=0}^{\infty}\int_{\hat{\gamma}_k=0}^{\infty}e^{-\beta_m l\gamma}f_{\gamma_k|\hat{\gamma}_k}(\gamma|\hat{\gamma}_k)f_{\hat{\gamma}_k|\mathcal{A}_k}(\hat{\gamma}_k|\mathcal{A}_k)d\gamma d\hat{\gamma}_k \\ &= b_m\frac{1}{K}\sum_{k=1}^{K}\sum_{l=0}^{L}\binom{L}{l}(-\alpha_m)^l\int_{\gamma=0}^{\infty}e^{-\beta_m l\gamma}f_{\gamma^*}(\gamma)d\gamma\end{aligned}$$

This integral can be easily solved for both the SISO and OSTBC approaches by invoking [Gra65, Eqs. 3.310] and [Gra65, Eqs. 3.351.3], respectively. By doing so, the following analytical expressions can be obtained:

$$\overline{\eta}_{m,SISO}(K) \approx b_m\sum_{k=1}^{K}\sum_{i=0}^{K-1}\binom{K-1}{i}\sum_{l=0}^{L}\binom{L}{l}\frac{(-1)^i(-\alpha_m)^l}{(i+1)(1+\bar{\gamma}_k\sigma_{\epsilon_k}^2\beta_m l)+\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2\beta_m l} \quad (4.43)$$

$$\begin{aligned}\overline{\eta}_{m,OSTBC}(K) \approx\ & 4b_m\sum_{k=1}^{K}\sum_{i=0}^{K-1}\binom{K-1}{i}(-1)^i\sum_{l=0}^{L}\binom{L}{l}\sum_{n=0}^{i}\binom{i}{n}\sum_{s=0}^{n}\binom{n+1}{n-s} \\ &\times\frac{n!\left(\mathbb{E}[\hat{\gamma}_k]\right)^s\vartheta_k^{2s}\bar{\gamma}_k^{n-s}\sigma_{\epsilon_k}^{2(n-s)}2^s(s+1)(-\alpha_m)^l}{\left((i+1)(2+\bar{\gamma}_k\sigma_{\epsilon_k}^2\beta_m l)+\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2\beta_m l\right)^{s+2}(\mathbb{E}[\hat{\gamma}_k]\vartheta_k^2+(i+1)\bar{\gamma}_k\sigma_{\epsilon_k}^2)^n}\end{aligned} \quad (4.44)$$

For the case that $\vartheta_k \longrightarrow 0$ and $\sigma_{\epsilon_k} \longrightarrow 1$, it can be easily shown that the above expressions tend to those given by Eqs. (4.27) and (4.28). Therefore, the conclusions drawn in the asymptotic

study carried out in the previous section are still valid for this case. In other words, when the estimated SNR available at the scheduler is considerably degraded, it is more appropriate to adopt an OSTBC scheme due to its inherent robustness against wrong scheduling decisions.

4.6.5 Practical Examples

So far, we have derived analytical expressions for the general case of degraded CSI. In order to gain some more insight, we will now particularize those expressions to two practical cases where the Gaussian model applies:

- *Delayed feedback channel*:

Unless reciprocity between the forward and reverse links can be established, there always exists a delay between the instant when the SNR is measured at the MS and the actual transmission of data to the scheduled user. Under the assumption of a Jakes' scattering model, $\mathbf{h}_k$ and $\hat{\mathbf{h}}_k$ turn out to be samples of the same Gaussian process. In other words, $\mathbf{h}_k$ and $\hat{\mathbf{h}}_k$ follow a joint complex Gaussian distribution with correlation coefficient $\rho_k = J_o(2\pi f_{d_k} T_k)$, where f_{d_k} stands for the Doppler frequency, T_k is the delay in time units, and $J_o(\cdot)$ denotes the zero-order Bessel function of the first kind. As a consequence, the conditioned pdf can be easily obtained by applying Bayes' Theorem [Kay93, Chapter 10]:

$$\begin{aligned} f_{\mathbf{h}_k|\hat{\mathbf{h}}_k}(\mathbf{h}_k|\hat{\mathbf{h}}_k) &= \frac{f_{\mathbf{h}_k,\hat{\mathbf{h}}_k}(\mathbf{h}_k\hat{\mathbf{h}}_k)}{f_{\hat{\mathbf{h}}_k}(\hat{\mathbf{h}}_k)} \\ &= \frac{e^{-(\mathbf{h}_k-\rho_k\hat{\mathbf{h}}_k)^H\mathbf{R}_k^{-1}(\mathbf{h}_k-\rho\hat{\mathbf{h}}_k)}}{\pi^M \det(\mathbf{R}_k)} \end{aligned}$$

where $\mathbf{R}_k = \sigma_{h_k}^2(1-\rho_k^2)\mathbf{I}_M$ is the covariance matrix. Therefore, the delayed feedback channel fits into the Gaussian model since we have that:

$$\mathbf{h}_k|\hat{\mathbf{h}}_k \sim \mathcal{CN}\left(\rho_k\hat{\mathbf{h}}_k, \mathbf{R}_k\right)$$

and, hence, both the ergodic capacity and system throughput can be computed by replacing:

$$\vartheta_k = \rho_k \qquad \sigma_{\epsilon_k}^2 = 1-\rho_k^2 \qquad \mathbb{E}[\hat{\gamma}_k] = \bar{\gamma}_k \ ,$$

into Eqs. (4.40), (4.41), (4.43) and (4.44), where the last equality follows from the fact that $\mathbf{h}$ and $\hat{\mathbf{h}}$ are samples of the same Gaussian process.

- *Imperfect channel estimation*:

It is common practice to assume that the channel impulse response is perfectly known at the receiver. However, in practical situations only an estimate of the actual channel is actually

available. In the case of a linear MMSE estimator, for instance, we can model the channel estimate as [Kay93]:

$$\hat{\mathbf{h}}_k = \mathbf{h}_k + \mathbf{e}_k$$

where $\mathbf{e}_k \in \mathbb{C}^M$ is the vector corresponding to the channel estimation error for which each component is assumed to be an i.i.d circularly symmetric Gaussian random variable with variance $\sigma^2_{e_k}$ and independent from $\mathbf{h}_k$. Then, by applying the Bayes' Theorem one can find that:

$$\mathbf{h}_k|\hat{\mathbf{h}}_k \sim \mathcal{CN}\left(\frac{1}{1+\Delta_{e_k}}\hat{\mathbf{h}}_k, \sigma^2_{h_k}\frac{\Delta_{e_k}}{1+\Delta_{e_k}}\mathbf{I}_M\right)$$

where we have defined $\Delta_{e_k} = \frac{\sigma^2_{e_k}}{\sigma^2_{h_k}}$, which can be interpreted as the inverse of the SNR of the estimation process. Thus, we should take the following parameters into account in the expressions of the ergodic capacity and system throughput:

$$\vartheta_k = \tfrac{1}{1+\Delta_{e_k}} \qquad \sigma^2_{\epsilon_k} = \tfrac{\Delta_{e_k}}{1+\Delta_{e_k}} \qquad \mathbb{E}[\hat{\gamma}_k] = \bar{\gamma}_k(1+\Delta_{e_k})$$

As a final remark, it should be noted that in this second example, we are only considering the impact of noisy channel estimates on the scheduling process, whereas we disregard its impact on the detection process at the receiver. However, this second issue is out of the scope of this work since we are interested in the analytical study of the impact caused by incorrect scheduling decisions. For further details, the reader is referred to [Med00] and [Yoo06b], where some work related to SISO and MIMO channels, respectively, can be found.

4.6.6 Numerical Results and Discussion

In Fig. 4.13, we depict the ergodic capacity as a function of MS speed. As for the CSI delay, we adopt the parameters used in [Ber03] for an HSDPA scenario where the authors justify that scheduling decisions can be made every 2 ms with a time delay of $T = 4$ ms. From the curves, one concludes that in the absence of delay, the SISO approach is far more effective than its OSTBC counterpart in exploiting MUD gains. When MS speed increases, however, the degradation experienced by the SISO scheme is larger than that of OSTBC. In other words, the single-antenna approach is less robust to channel uncertainty (i.e. deep fades) arising from CSI delays. As the number of active users grows, though, the capability of generating post-scheduling SNR peaks improves faster for SISO configurations and, hence, compensates for such SNR uncertainties (i.e. SISO and OSTBC curves cross each other for higher values of the MS speed). Finally, one can also observe that beyond 50 km/h curves are driven again towards higher values of the ergodic capacity. This is because under a Jakes' scattering model assumption, the correlation depends on the zero-order Bessel function of the first kind, which is not a monotonically decreasing function. Similar conclusions can be drawn from Fig. 4.14, where the ergodic capacity is plotted as a function of the parameter Δ_e. However, in this case no extra multi-user diversity gain can

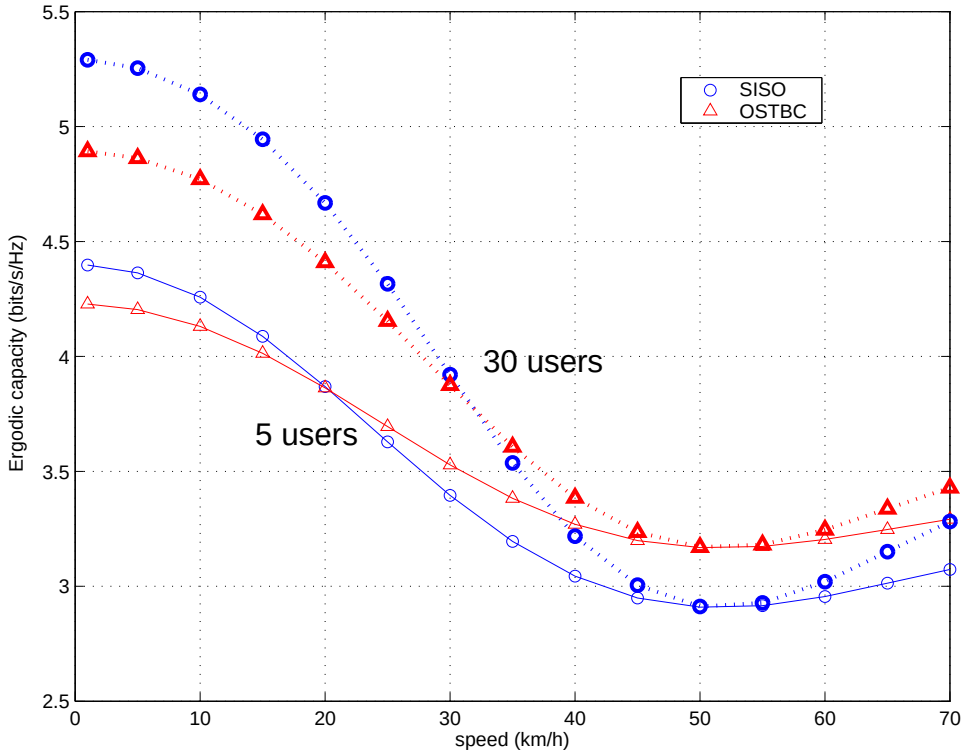


Figure 4.13: Ergodic capacity vs. MS speed for the different transmission schemes (symbols: simulated results, curves: analytical expressions, $K = 5$ and 30 users,$\bar{\gamma} = 10$ dB).

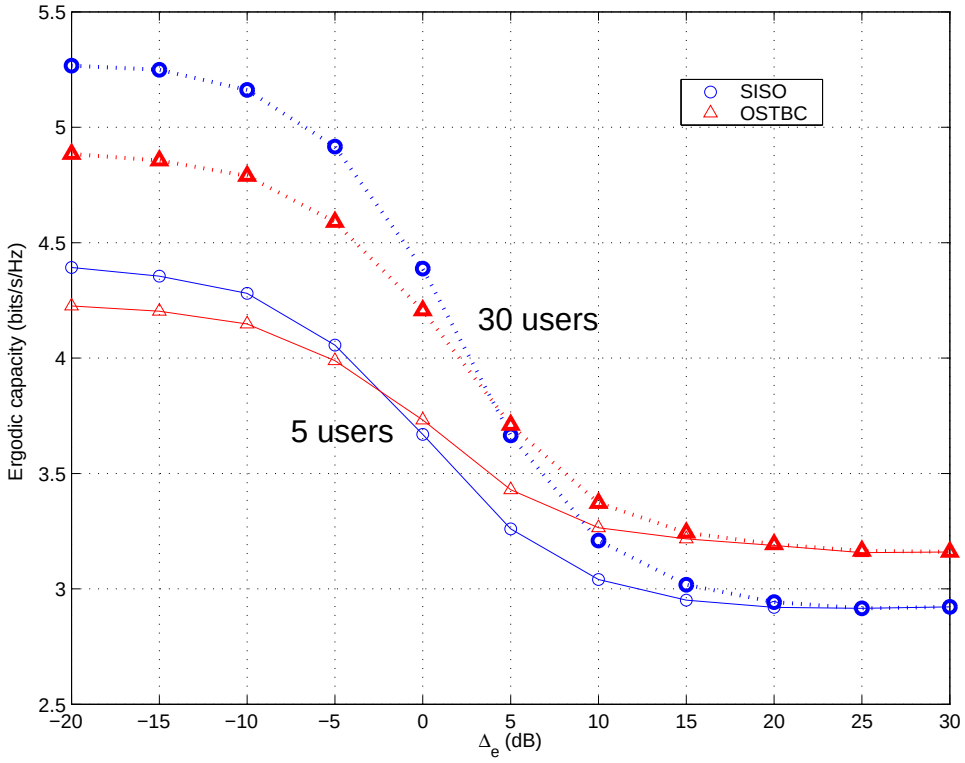


Figure 4.14: Ergodic capacity vs. Δ_e for the different transmission schemes (symbols: simulated results, curves: analytical expressions, $K = 5$ and 30 users,$\bar{\gamma} = 10$ dB).

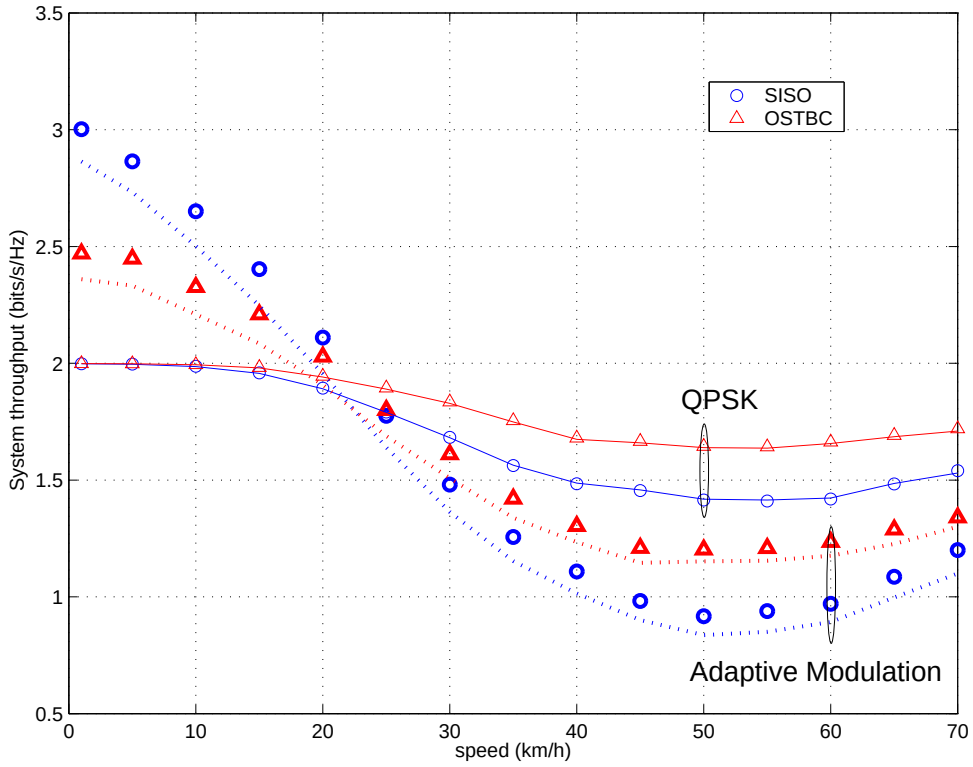


Figure 4.15: System throughput vs. MS speed for the different transmission schemes (solid line: QPSK, dotted line: adaptive modulation, symbols: simulated results, curves: analytical expressions and numerical integration, K=10 users, L=10 symbols, $\bar{\gamma} = 10$ dB).

be extracted as the degradation in the channel estimates increases. For increasing values of Δ_e, capacity curves reach a floor associated with the performance of a round-robin scheduler (i.e. no MUD gain).

Regarding results in terms of system throughput, we will only show the figure associated to the delayed feedback case (see Fig. 4.15) given the similarity between the results obtained in the proposed practical examples (delayed feedback channel and imperfect channel estimation). Analytical curves corresponding to the adaptive modulation case have been obtained by means of numerical integration, whereas Eqs. (4.43) and (4.44) have been used to evaluate the QPSK case. As shown in the previous section, better results are obtained with OSTBC when only QPSK is considered[7] and MUD gains can be exploited by increasing the degrees of freedom in the adaptive modulation scheme. In this case, however, worse results are obtained with adaptive modulation when the information available at the scheduler is degraded. This is because the BS fails to correctly adapt the modulation schemes to scenario conditions when the SNR estimates are less accurate. Therefore, the robustness of OSTBC is exploited to alleviate negative effects related to wrong scheduling decisions and inefficient rate adaptation.

[7]It is worth noting that, both schemes obtain the same performance in the lowest speed region due to the saturation effect associated with QPSK.

4.7 Assessment of Capacity vs. Robustness Trade-offs

In the previous sections, we have shown that the suppression of SNR peaks associated to the stabilizing effect of OSTBC penalizes system performance. Nonetheless, such stabilizing effect provides additional robustness against unfavorable fading conditions resulting from wrong scheduling decisions. Therefore, there exists a trade-off in terms of the degree of robustness to short-term SNR fluctuations and its impact in terms of system performance. In this section, we are devoted to illustrate such a trade-off with the help of means vs. standard deviation plots [Bar05], inspired by modern *portfolio* theory [Mar52, Mar91]. By representing in the same plot both the mean and standard deviation of the system capacity, one can easily quantify the degree of robustness and the associated gain in system performance of the different schemes. In other words, spatial vs. MUD trade-offs can be assessed from a more visual perspective.

This kind of representation is used in financial market theory with the aim of assessing the existing trade-offs in terms of the expected profit (mean) vs. the possible risk (standard deviation) when a possible investment is considered. Bartolomé introduced this methodology in multi-antenna communications systems to study the degree of fairness in resource allocation algorithms [Bar05]. In particular, he considered the MIMO Broadcast Channel with ZF transmit beamforming and compared different bit allocation techniques. By using the mean vs. standard deviation plots the trade-offs in terms of global rate vs. fairness among users can be easily showed. Then, this approach facilitates the design and comparison of different resource allocation algorithms according to the desired degree of fairness. This technique can also be found in studies about the comparison of optimum vs. ZF beamforming [Ben05] and the design of fair algorithms in a context where a orthogonal linear precoding is adopted [Bos05, Zag06].

Besides from showing the degree of robustness against channel fading, the purpose of this section is illustrating the fairness among users in terms of assigned data rates. Certainly the proposed scheduler is completely fair in the sense that all the users access the channel with the same probability $1/K$. However, we are devoted to assess the degree of fairness in terms of the quality of the links assigned to the different users by the transmission schemes (SISO or OSTBC). In other words, we study the fairness in the physical layer by analyzing the quality of the physical links (in terms of instantaneous capacity) offered to the different users. Such a view is useful for delay-limited services where the short-term fluctuations of channel capacity become relevant and one should take a closer look at the standard deviation of system capacity. In order to quantify the fairness in the physical layer, we consider an homogenous case as the objective here is to assess the differences among users resulting from using SISO or OSTBC schemes.

To start with, consider a scenario with selective-feedback scheduling. In Fig. 4.16, we depict the mean (average) vs. the standard deviation of system capacity for a varying number of users ($K = 1, ..., 30$), feedback loads ($\bar{F} = 1, ..., 0.01$), and transmission schemes (SISO/OSTBC). For

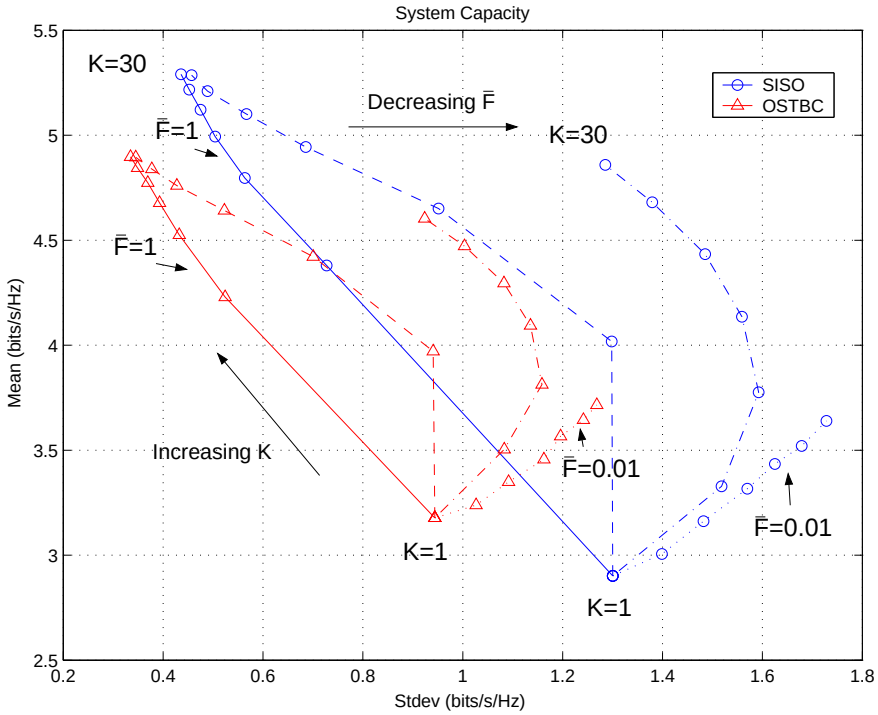


Figure 4.16: System capacity: mean (average) vs. standard deviation plot as a function of the transmission scheme (SISO/OSTBC), number of users ($K = 1, 5, 10, 15, 20, 25, 30$ users), and feedback load ($\bar{F} = 1, 0.2, 0.05, 0.01$). $\bar{\gamma}$=10dB.

the $K = 1$ and $\bar{F} = 1$ case, the average capacity is higher for OSTBC than for SISO and, simultaneously, the standard deviation is lower (i.e. higher capacity and more stable communication links). However, as soon as the number of users increases beyond $K = 1$ and for mid to high values of $\bar{F}$, SISO links outperform OSTBC ones in terms of capacity whereas OSTBC links remain more stable than SISO ones (or, alternatively, the data-rate dispersion among active users for a *short* period of time is lower). One can also observe that for decreasing values of the feedback load, both SISO and OSTBC links become less stable (to different extents) since, in those conditions, the number of random scheduling decisions increases. For high and moderate values of $\bar{F}$ this can be partially compensated by increasing the number of active users. In those conditions, the likelihood of having at least one user above ξ_{th} is higher and, hence, the reduced number of random scheduling decisions drives those curves again towards the low standard deviation region. Nonetheless, such an effect vanishes as the feedback load is further reduced (i.e $\bar{F} = 0.01$).

In Fig. 4.17, we consider that bandwidth restrictions are not imposed in the feedback channel but this is subject to delays. As in the previous case, the standard deviation of the different transmission schemes is higher when the CSI available at the scheduler is degraded, being this increase higher for the SISO case. However, the standard deviation is always a decreasing function

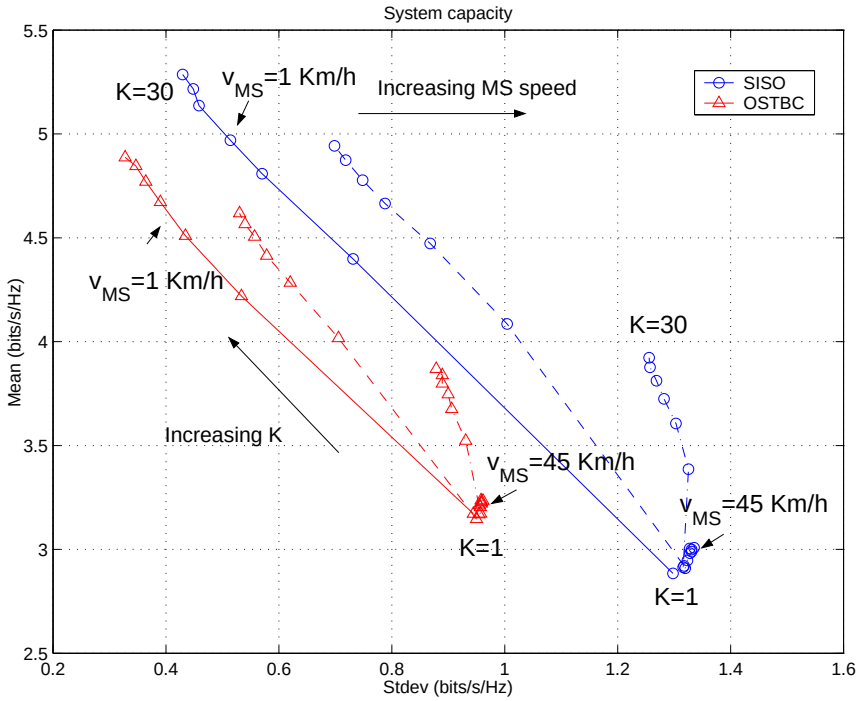


Figure 4.17: System capacity: mean (average) vs. standard deviation plot as a function of the transmission scheme (SISO/OSTBC), number of users ($K = 1, 5, 10, 15, 20, 25, 30$ users), and MS speed ($v_{MS} = 1, 15, 30, 45$ Km/h). $\bar{\gamma}$=10dB.

of K. Although the CSI is outdated, the scheduler is still selecting the maximum of a set of channel gains. Nonetheless, for $v_{MS} = 45$ Km/h corresponding points are overlapping. In this case, the available CSI is so degraded that the BS is encompassing round-robin scheduling, which does not depend on the number of users. Results corresponding to the case with imperfect channel estimation are quite similar and omitted here for brevity.

In summary, a number of non-trivial trade-offs in terms of ergodic capacity vs. robustness to short-term variations arise when considering different transmission schemes, impairments in the CSI and terminal count. As usual, design decisions at the cell level will be closely linked to the QoS requirements of the services under consideration.

4.8 Chapter Summary and Conclusions

In this chapter, a number of transmission schemes (SISO, SISO-AS, OSTBC and OSTBC-AS) aimed at jointly exploiting both multi-user and spatial diversity arising from antenna selection strategies and orthogonal space-time block coding have been presented. Different scenarios have been considered: ideal case, selective-feedback scheduling and imperfect CSI at the scheduler.

In terms of pre-scheduling SNR, we observed that the stabilizing effect (i.e. reduced number of

deep fades and SNR peaks) associated to OSTBC can be modified by the introduction of antenna selection mechanisms that partly restore the missing SNR peaks. Overall system performance, though, was assessed in terms of average system capacity and throughput, corresponding to the limiting cases of having infinite and limited AMC levels available, respectively.

In the ideal scenario, conclusions drawn from capacity curves were, as expected, totally aligned with the results obtained from SNR densities. In terms of capacity, the OSTBC-AS scheme combines antenna diversity and multi-user diversity and performs better than the MU-exploiting SISO approach in the K=1..8 users range. However, the SISO-AS configuration performs better for the whole range of users. Due to correlation effects, OSTBC-AS is equivalent to a twofold increase in the number of active users with respect to the OSTBC scheme without antenna selection, in terms of ergodic capacity. In terms of throughput different results are obtained according to the number of AMC levels. If only QPSK is considered, part or all of the advantage exhibited by the SISO and SISO-AS configurations is lost in favor of the OSTBC and OSTBC-AS configurations respectively. For instance, OSTBC outperforms SISO for the whole range of users as opposed to channel capacity where that only happened for the case with K=1 active users. These effects are associated to saturation effects, which can be avoided by introducing adaptive modulation mechanisms. With the help of 16-QAM, multi-user gains can be exploited and SISO approaches are benefited of that.

In the selective-feedback case, when the average feedback load per user is reduced one can observe that the degradation, both in terms of capacity and throughput, experienced by the SISO-based schemes is larger than that exhibited by the OSTBC ones. Additional gains can be obtained from the introduction of AS mechanisms, albeit very moderate when feedback load is low. A similar effect can be observed in those cases where the CSI available at the scheduler is not limited but is subject to imperfections. In particular, it has been shown that OSTBC-based schemes are more appropriate for scenarios where CSI available at the BS is considerably degraded, specially for a reduced number of active users.

To conclude the chapter, a novel approach have been used to assess spatial vs. multi-user trade-offs. More precisely, the robustness of the different transmission schemes against impairments in the CSI have been assessed by means of mean vs. standard deviation plots. It has been shown that, when the information available at the scheduler is limited or degraded, both SISO and OSTBC links become less stable. In particular, a lower degradation is observed for the OSTBC approaches due to their inherent robustness. By increasing the number of users, the system becomes more robust to impairments in the CSI, but such an effect vanishes as the available CSI is further reduced or degraded.

4.A Appendix: Derivation of $\Upsilon(a, m, \mu)$ and $\Phi(a, m, \mu)$

In this appendix, we derive closed-form expressions for the following two integrals:

$$\Upsilon(a, m, \mu) \triangleq \int_a^{\infty} \ln(1+t) t^{m-1} e^{-\mu t} dt \tag{4.45}$$
$$\Phi(a, m, \mu) \triangleq \int_0^{a} \ln(1+t) t^{m-1} e^{-\mu t} dt$$
$$\mu > 0; m = 1, 2, ...$$

In [Alo99], the authors solved the integral $\Upsilon(a, m, \mu)$ for the case $a = 0$:

$$\Upsilon(0, m, \mu) = (m-1)! e^{\mu} \sum_{i=1}^{m} \frac{\Gamma(-m+i, \mu)}{\mu^i} \tag{4.46}$$

where $\Gamma(\alpha, x)$ stands for the complementary incomplete gamma function ($\Gamma(\alpha, x) \triangleq \int_x^{\infty} e^{-t} t^{\alpha-1} dt$). Hence, we only need to calculate $\Upsilon(a, m, \mu)$ for finite values of a since, clearly, $\Phi(a, m, \mu) = \Upsilon(0, m, \mu) - \Upsilon(a, m, \mu)$. By conducting an integration by parts, Eq. (4.45) can be conveniently re-written as:

$$\Upsilon(a, m, \mu) = \int_a^{\infty} u dv = \lim_{t \longrightarrow \infty} (uv) - \lim_{t \longrightarrow a} (uv) - \int_a^{\infty} v du. \tag{4.47}$$

Then, we let:

$$u = \ln(1+t) \quad dv = t^{m-1} e^{-\mu t}$$
$$du = \tfrac{dt}{1+t} \qquad v = -e^{-\mu t} \sum_{i=1}^{m} \frac{(m-1)!}{(m-i)!} \frac{t^{m-i}}{\mu^i}$$

where the last equality results from [Gra65, Eq. 2.321.2]. If we use the expressions above in Eq. (4.47) and note that the first term goes to zero, the following intermediate expression results:

$$\Upsilon(a, m, \mu) = \ln(1+a) e^{-\mu a} \sum_{i=1}^{m} \frac{(m-1)!}{(m-i)!} \frac{a^{m-i}}{\mu^i} + \sum_{i=1}^{m} \frac{(m-1)!}{(m-i)!} \frac{1}{\mu^i} \int_a^{\infty} \frac{t^{m-i}}{1+t} e^{-\mu t} dt$$

By using the change of variables $x = t - a$ and the binomial expansion, the integral in the above equation can be re-written as:

$$\int_a^{\infty} \frac{t^{m-i}}{1+t} e^{-\mu t} dt = e^{-\mu a} \int_0^{\infty} \frac{(x+a)^{m-i} e^{-\mu x}}{1+a+x} dx$$
$$= e^{-\mu a} \sum_{p=0}^{m-i} \binom{m-i}{p} a^{m-i-p} \int_0^{\infty} \frac{x^p e^{-\mu x}}{1+a+x} dx$$

Next, with the help of [Gra65, Eq. 3.383.10] and after some manipulation, $\Upsilon(a,m,\mu)$ can be expressed as follows:

$$\Upsilon(a,m,\mu) = \sum_{i=1}^{m} \frac{(m-1)!}{(m-i)!} \frac{e^{-\mu a} a^{m-i}}{\mu^i} \Bigg[\ln(1+a) + \sum_{p=0}^{m-i} \binom{m-i}{p} \left(\frac{1+a}{a}\right)^p e^{\mu(1+a)} p! \Gamma\left(-p, \mu(1+a)\right) \Bigg] \tag{4.48}$$

Finally and by resorting to Eqs. (4.46) and (4.48), we can write $\Phi(a,m,\mu)$ in closed-form as well:

$$\Phi(a,m,\mu) = \sum_{i=1}^{m} \frac{(m-1)!}{\mu^i} \Bigg[e^{\mu} \Gamma(-m+i,\mu) - \frac{e^{-\mu a} a^{m-i}}{(m-i)!} \times \left(\ln(1+a) + \sum_{p=0}^{m-i} \binom{m-i}{p} \left(\frac{1+a}{a}\right)^p e^{\mu(1+a)} p! \Gamma\left(-p, \mu(1+a)\right) \right) \Bigg] \tag{4.49}$$

4.B Appendix: Proof of Eq. (4.24)

Since in the $\gamma_k/\bar{\gamma}_k \leq \xi_{th}$ region, all users remain silent, both the scheduled user and the transmit antenna subset in the BS are randomly selected. Therefore, the probability $\text{Prob}(\gamma_k \leq \gamma)$ in Equation (4.19) can be expressed as:

$$\begin{aligned} \text{Prob}(\gamma_k \leq \gamma) &= \text{Prob}(\gamma_k \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k | \gamma_k = \gamma_{min}) \text{Prob}(\gamma_k = \gamma_{min}) \\ &\quad + \text{Prob}(\gamma_k \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k | \gamma_k = \gamma_{med}) \text{Prob}(\gamma_k = \gamma_{med}) \\ &\quad + \text{Prob}(\gamma_k \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k | \gamma_k = \gamma_{max}) \text{Prob}(\gamma_k = \gamma_{max}) \\ &= \text{Prob}(\gamma_{min} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) \text{Prob}(\gamma_k = \gamma_{min}) \\ &\quad + \text{Prob}(\gamma_{med} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) \text{Prob}(\gamma_k = \gamma_{med}) \\ &\quad + \text{Prob}(\gamma_{max} \leq \gamma) \text{Prob}(\gamma_k = \gamma_{max}) \end{aligned} \tag{4.50}$$

where γ_{min}, γ_{med} and γ_{max} stand for the received SNR associated with the worst, the second worst and the best antenna subsets and, clearly,

$$\text{Prob}(\gamma_k = \gamma_{min}) = \text{Prob}(\gamma_k = \gamma_{med}) = \text{Prob}(\gamma_k = \gamma_{max}) = \frac{1}{3} \tag{4.51}$$

due to the random choice of antenna subsets. In Eq. (4.50) above, the probability that the SNR associated to the best antenna subset is below $\xi_{th}\bar{\gamma}_k$, can be readily computed from the pre-scheduling SNR distribution (see Eq. (4.11)), that is:

$$\text{Prob}(\gamma_{max} \leq \gamma) = F_{\gamma_{k,OSTBC-AS}}(\gamma) \tag{4.52}$$

However, a more elaborate analysis is needed to calculate the joint probabilities in Eq. (4.50) which cannot be factorized due to the fact that the antenna subsets are not statistically independent, i.e.:

$$\text{Prob}(\gamma_{min} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) \neq \text{Prob}(\gamma_{min} \leq \gamma)\text{Prob}(\gamma_{max} \leq \xi_{th}\bar{\gamma}_k)$$
$$\text{Prob}(\gamma_{med} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) \neq \text{Prob}(\gamma_{med} \leq \gamma)\text{Prob}(\gamma_{max} \leq \xi_{th}\bar{\gamma}_k)$$

Again, one can resort to order statistics to solve this problem. The joint distribution of the normalized channel gains (X_1, X_2, X_3) can be written as:

$$f_{X_3X_2X_1}(x_3, x_2, x_1) = \begin{cases} 3!e^{-(x_1+x_2+x_3)} & \text{for } x_1 \leq x_2 \leq x_3 \\ 0 & \text{otherwise} \end{cases}$$

For the $0 \leq \gamma \leq \frac{\xi_{th}\bar{\gamma}_k}{2}$ interval , we have:

$$\begin{aligned}
&\text{Prob}(\gamma_{min} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) = \text{Prob}(X_1 + X_2 \leq \frac{2\gamma}{\bar{\gamma}_k} = t, X_2 + X_3 \leq \frac{2\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k} = 2\xi_{th} = m) \\
&= \int_{x_3=0}^{\frac{t}{2}} \int_{x_2=0}^{x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=\frac{t}{2}}^{t} \int_{x_2=0}^{\frac{t}{2}} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 \\
&+ \int_{x_3=\frac{t}{2}}^{t} \int_{x_2=\frac{t}{2}}^{x_3} \int_{x_1=0}^{t-x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=t}^{m-t} \int_{x_2=0}^{\frac{t}{2}} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 \\
&+ \int_{x_3=t}^{m-t} \int_{x_2=\frac{t}{2}}^{t} \int_{x_1=0}^{t-x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=m-t}^{m-\frac{t}{2}} \int_{x_2=0}^{\frac{t}{2}} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 \\
&+ \int_{x_3=m-t}^{m-\frac{t}{2}} \int_{x_2=\frac{t}{2}}^{m-x_3} \int_{x_1=0}^{t-x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=m-\frac{t}{2}}^{m} \int_{x_2=0}^{m-x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3
\end{aligned} \tag{4.53}$$

$$\begin{aligned}
&\text{Prob}(\gamma_{med} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) = \text{Prob}(X_1 + X_3 \leq \frac{2\gamma}{\bar{\gamma}_k} = t, X_2 + X_3 \leq 2\xi_{th} = m) \\
&= \int_{x_3=0}^{\frac{t}{2}} \int_{x_2=0}^{x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=\frac{t}{2}}^{t} \int_{x_2=0}^{t-x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 \\
&+ \int_{x_3=\frac{t}{2}}^{t} \int_{x_2=t-x_3}^{x_3} \int_{x_1=0}^{t-x_3} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3
\end{aligned} \tag{4.54}$$

whereas for the $\frac{\xi_{th}\bar{\gamma}_k}{2} \leq \gamma \leq \xi_{th}\bar{\gamma}_k$ interval:

$$\text{Prob}(\gamma_{min} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) = \text{Prob}(X_1 + X_2 \leq \frac{2\gamma}{\bar{\gamma}_k} = t, X_2 + X_3 \leq 2\xi_{th} = m)$$
$$= \int_{x_3=0}^{\frac{t}{2}} \int_{x_2=0}^{x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=\frac{t}{2}}^{\frac{m}{2}} \int_{x_2=0}^{\frac{t}{2}} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3$$
$$+ \int_{x_3=\frac{t}{2}}^{\frac{m}{2}} \int_{x_2=\frac{t}{2}}^{x_3} \int_{x_1=0}^{t-x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=\frac{m}{2}}^{m-\frac{t}{2}} \int_{x_2=0}^{\frac{t}{2}} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3$$
$$+ \int_{x_3=\frac{m}{2}}^{m-\frac{t}{2}} \int_{x_2=\frac{t}{2}}^{m-x_3} \int_{x_1=0}^{t-x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=m-\frac{t}{2}}^{m} \int_{x_2=0}^{m-x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 \tag{4.55}$$

$$\text{Prob}(\gamma_{med} \leq \gamma, \gamma_{max} \leq \xi_{th}\bar{\gamma}_k) = \text{Prob}(X_1 + X_3 \leq \frac{2\gamma}{\bar{\gamma}_k} = t, X_2 + X_3 \leq 2\xi_{th} = m)$$
$$= \int_{x_3=0}^{\frac{t}{2}} \int_{x_2=0}^{x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=\frac{t}{2}}^{\frac{m}{2}} \int_{x_2=0}^{t-x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3$$
$$+ \int_{x_3=\frac{t}{2}}^{\frac{m}{2}} \int_{x_2=t-x_3}^{x_3} \int_{x_1=0}^{t-x_3} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 + \int_{x_3=\frac{m}{2}}^{t} \int_{x_2=0}^{t-x_3} \int_{x_1=0}^{x_2} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3$$
$$+ \int_{x_3=\frac{m}{2}}^{t} \int_{x_2=t-x_3}^{m-x_3} \int_{x_1=0}^{t-x_3} f_{X_3X_2X_1}(x_3,x_2,x_1)dx_1dx_2dx_3 \tag{4.56}$$

By solving the integrals in Eqs. (4.53)-(4.56) above and plugging the results (along with Eqs. (4.51) and (4.52)) into Eq. (4.50), we have:

$$\text{Prob}(\gamma_k \leq \gamma) = -\frac{e^{-2\xi_{th}}}{\bar{\gamma}_k}\left(4\bar{\gamma}_k e^{-\frac{\gamma}{\bar{\gamma}_k}} + 2\gamma e^{-2\frac{\gamma-\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k}} + \bar{\gamma}_k e^{-2\frac{\gamma-\xi_{th}\bar{\gamma}_k}{\bar{\gamma}_k}} - \bar{\gamma}_k e^{2\xi_{th}} + 4\gamma - 4\bar{\gamma}_k\right)$$

Finally, by substituting this last expression into Equation (4.19) and taking its derivative, Eq. (4.24) follows.

Chapter 5

Adaptive Beam Selection for Orthogonal Random Beamforming

In this chapter, we consider a scenario where an Orthogonal Random Beamforming (ORB) strategy is adopted with the aim of serving several users simultaneously. In order to improve system performance in scenarios with a low number of users, the idea of antenna selection is adapted to ORB and a beam selection algorithm is derived. Besides, the impact of feedback quantization on ORB performance is studied and a beam selection algorithm trading-off system performance vs. amount of feedback information is proposed.

5.1 Introduction

As explained in Chapter 2, orthogonal random beamforming has recently attracted significant interest. This is because these approaches merely require partial CSI at the transmitter, mostly SINR measurements for each transmit beam. Hence, those schemes have emerged as a viable alternative to more sophisticated transmit beamforming approaches, in particular in the asymptotic case of a growing number of users since, in that region, the sum-rate exhibits the same growth-rate as the capacity-achieving technique for the MIMO Gaussian Broadcast Channel do (i.e., DPC). Nonetheless, poor performance is obtained in systems with a practical number of users and, for that reason, some doubts concerning the practicality of these schemes have arisen.

In contrast with other beamforming schemes, there is not much work in the literature where the enhancement of ORB in low populated networks has been addressed. In [Kou05a], Kountoris and Gesbert proposed an approach where channel time correlation is exploited to search for the optimal set of random beamvectors. In order not to excessively depend on the properties of the channel, a different approach was proposed in [Kou05b]. More precisely, it was proposed a scheme where a low-rate feedback channel is used for selecting the best group of users, for which more efficient beamforming techniques are applied: MMSE beamformer and iterative power allocation

with full and partial CSI at the transmitter, respectively.

In this chapter, we also propose approaches aimed at improving ORB in sparse networks. However, we focus our attention on the simple philosophy of the original approach where users are served with uniform power allocation. As we will show later, activating all the available beams may not be the optimum solution for scarcely populated cells with constant transmit power at the base station. For that reason, we propose beam selection procedures where the optimum *subset of beams* is selected for transmission. In particular, we propose different beam selection techniques where the only required information are the channel gains or the SINR.

Besides, the impact of feedback quantization in an opportunistic beamforming context is studied as well. In particular, we extend the work by the authors in [Flo03] to a more general case. In that paper, it was shown that most of the MUD can be extracted in a SISO multi-user scheme when partial CSI is quantized with a very low number of bits. In this chapter we prove that similar conclusions can be drawn for the multi-antenna case as well. To do so, we derive closed-form expressions of the aggregated throughput for such communications scenario in the presence of AM. Throughput expressions are obtained both for systems with quantized and non-quantized (analog) partial CSI. Further, we propose a non-uniform quantization law (on the basis of post-scheduling SINR statistics) which provides substantial gains with respect to the uniform case to finally prove that penalties associated to quantization are small even for a very reduced number of quantization levels. It is worth noting that it was recently shown in [Dia06] that most of MUD gains can be extracted in ORB with one feedback bit. Such a result, however, is only valid in scenarios with a extremely high number of users ($K \geq 1000$ users) and, then, is out of the interests of this work based on the study of practical scenarios.

In summary, the contributions of this chapter are the following:

- It is proved that in a system with a low (practical) number of users, activating all the available beams is not the most appropriate strategy.
- Closed-form expressions are obtained for the aggregated throughput of the system with analog and quantized feedback.
- A quantization technique based on post-scheduling SINR statistics is proposed, showing considerable benefits.
- A set of optimum and sub-optimum beam selection algorithms are derived.
- It is shown that most of the MUD gains can be extracted with a few bits in the feedback channel.

This chapter is organized as follows: in Section 5.2, the corresponding signal model is presented. Closed-form expressions for the density functions (pdf and CDF) of the post-scheduling

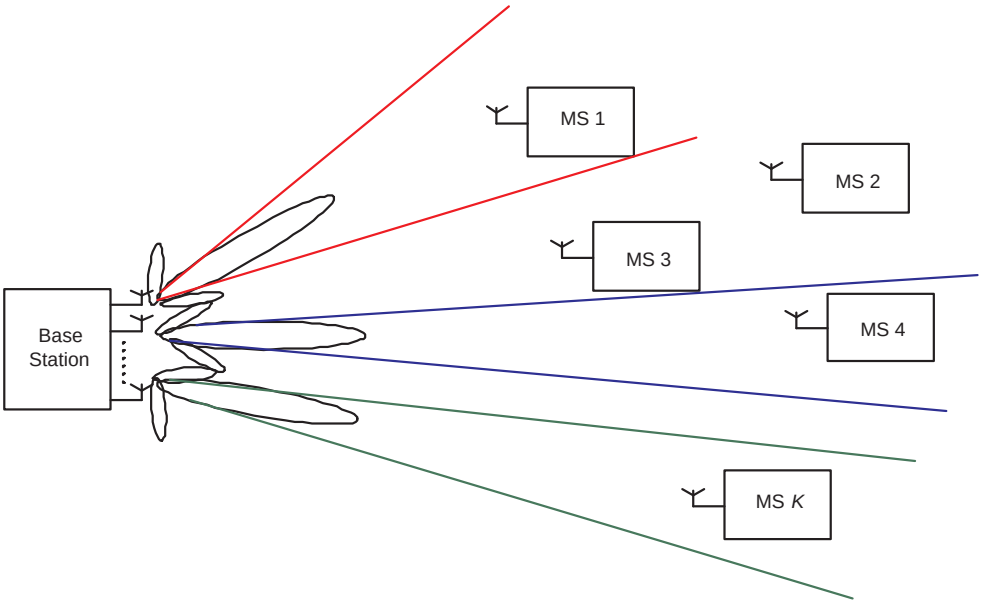


Figure 5.1: Block diagram of a multi-user communication system where ORB is considered.

SINRs are provided in Section 5.3. Next, in Sections 5.4 and 5.5 the behavior of ORB is analyzed in scenarios with a practical number of users in terms of both the sum-rate and aggregated throughput metrics, respectively. After that, the impact of feedback quantization on the aggregated throughput of the system is studied in Section 5.6. Adaptive beam selection algorithms are proposed in Section 5.7 along with an analysis in terms of system performance vs. complexity and feedback requirements trade-offs. Finally, in Section 5.8, the summary and conclusions of this chapter are presented.

The results obtained in this chapter have been published in [Vic06a] and [Vic06b], and one journal paper is still in preparation.

5.2 Signal Model

As shown in Fig. 5.1, we consider the downlink of a cellular system with one BS equipped with M antennas and K single-antenna MSs. In order to serve multiple users in the same time-slot, a linear precoding matrix is applied at the base station. In particular, we follow an orthogonal random beamforming strategy [Sha05]. More precisely, in each time-slot we construct a random matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_M]$, where $\mathbf{w}_i \in \mathbb{C}^{M\times 1}$, $i = 1, ..., M$, are random orthonormal vectors generated according to an isotropic distribution [Mar99]. Then, these vectors are used for transmitting information to the users with the highest SINRs. Unlike the approach in [Sha05], we do not necessarily transmit with all the beams vectors $\mathbf{w}_i$, that is, the transmission is made with a subset of active beams $\mathcal{B} \subset \mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_M\}$. Further details about the beam selection

procedure will be given in Section 5.7. Therefore, the received signal at the k-th MS is given by:

$$r_k = \mathbf{h}_k^T \mathbf{W}_{\mathcal{B}} \mathbf{s}_{\mathcal{B}} + n_k \tag{5.1}$$

where in the expression above the time index has been dropped for the ease of notation, $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ is the channel vector gain between the BS and the k-th MS, for which each component is assumed to be independent and identically distributed, circularly symmetric Gaussian random variable with zero mean and unit variance ($\mathbf{h}_k \sim \mathcal{CN}(0, \mathbf{I}_M)$), $\mathbf{W}_{\mathcal{B}} \in \mathbb{C}^{M\times B}$ is the precoding matrix constructed with the columns of $\mathbf{W}$ corresponding to the subset of active beams $\mathcal{B}$, $\mathbf{s}_{\mathcal{B}} \in \mathbb{C}^{B\times 1}$ is the symbol vector broadcasted from the BS, $B = \text{card}(\mathcal{B}) \leq M$ is the number of active beams (and consequently the number of simultaneously scheduled users) and $n_k \in \mathbb{C}$ denotes AWGN with zero mean and variance σ^2. The active users in the system are assumed to undergo independent Rayleigh fading processes. Further, we consider quasi-static fading, i.e, the channel response remains constant during one time-slot and changes to a new independent realization in the subsequent one.

Concerning channel state information, we assume perfect CSI knowledge for *each* user at the receive side, and the availability of a low-rate error-free feedback channel to convey partial CSI to the transmitter. Finally, the total transmit power, P_t, is constant and evenly distributed among active beams, i.e., $\mathbb{E}\{\mathbf{s}_{\mathcal{B}}^H \mathbf{s}_{\mathcal{B}}\} = P_t$. Then, we can define $\rho = \frac{P_t}{\sigma^2}$ as the average SNR of the system.

5.3 Post-scheduling SINR Statistics

According to the signal model presented in the previous section, the received signal for user k when using beamformer i for transmission can be re-written (recall Eq. (5.1)) as:

$$r_k = \mathbf{h}_k^T \mathbf{w}_i s_i + \sum_{\substack{j\in\mathcal{B} \\ j\neq i}} \mathbf{h}_k^T \mathbf{w}_j s_j + n_k \tag{5.2}$$

where s_j stands for the symbol transmitted with beam j. Notice that the last two terms in the above expression are associated with the interference-plus-noise contribution and, hence, the corresponding SINR amounts to:

$$\text{SINR}_{k,i} = \frac{|\mathbf{h}_k^T \mathbf{w}_i|^2}{B/\rho + \sum_{\substack{j\in\mathcal{B} \\ j\neq i}} |\mathbf{h}_k^T \mathbf{w}_j|^2} = \frac{z}{B/\rho + y} \tag{5.3}$$

Since we assume all users experience i.i.d Rayleigh fading and the beamformers are orthonormal to each other, z and y become independent chi-square distributed random variables, $z \sim \chi_2^2$ and

$y \sim \chi^2_{2B-2}$. Bearing this in mind, the CDF and pdf of the SINR can be expressed as [Sha05]:

$$F_{\mathsf{SINR}}(\gamma) = 1 - \frac{e^{-\frac{\gamma B}{\rho}}}{(1+\gamma)^{B-1}} \tag{5.4}$$

$$f_{\mathsf{SINR}}(\gamma) = \frac{e^{-\frac{\gamma B}{\rho}}}{(1+\gamma)^{B}} \left(\frac{B}{\rho}(1+\gamma) + B - 1 \right) \tag{5.5}$$

Notice that in a i.i.d Rayleigh fading scenario the SINR statistics depend on the *number of active beams*, B, but not on the number of *transmit antennas*, M, as long as the number of active beams is lower than M. Besides, the same SINR statistics can be obtained by transmitting B independent streams from B different transmit antennas, converting the problem into an antenna selection approach. In slow fading channels, however, generating random beams is useful for inducing channel fluctuations in order to improve MUD gains [Vis02].

The scheduling process is organized in a slot-by-slot basis following a *max*-SINR (*greedy*) rule. That is, for beam i the scheduler selects the active user k_i^* satisfying:

$$k_i^* = \arg \max_{k=1,\ldots,K} \{\mathsf{SINR}_{k,i}\}$$

where it is assumed that a different user is selected for each beam[1]. Since all the users experience i.i.d Rayleigh fading, the CDF of the *post-scheduling* SINR, $F_{\mathsf{SINR}^*}(\gamma)$, i.e. the SINR experienced by the scheduled user can be readily expressed with the help of order statistics [Dav81] in terms of Eq. (5.4) as:

$$F_{\mathsf{SINR}^*}(\gamma) = (F_{\mathsf{SINR}}(\gamma))^K = \left(1 - \frac{e^{-\frac{\gamma B}{\rho}}}{(1+\gamma)^{B-1}} \right)^K$$

Finally, by simply differentiating the above expression the corresponding pdf expression results:

$$f_{\mathsf{SINR}^*}(\gamma) = K \frac{e^{-\frac{\gamma B}{\rho}}}{(1+\gamma)^{B}} \left(\frac{B}{\rho}(1+\gamma) + B - 1 \right) \left(1 - \frac{e^{-\frac{\gamma B}{\rho}}}{(1+\gamma)^{B-1}} \right)^{K-1} \tag{5.6}$$

These expressions will be used in subsequent sections for the computation of both the sum-rate and aggregated throughput.

5.4 Behavior of the Sum-rate

Now, we are devoted to study the behavior of ORB performance as a function of the number of active beams. Our interest resides in analyzing if activating all the available beams is an appropriate strategy in scenarios with a low number of users, since in those scenarios the probability that the generated random beamforming vectors match users' channel characteristics is

[1]The probability that one user achieves the highest SINR on more than one beam is negligible when the number of users is large compared with the number of active beams ($K >> B$) [Sha05].

considerably lower. We begin our study showing the behavior of the fundamental limits of the system, that is, in terms of the sum-rate capacity.

According to the proposed scheduling policy, the sum-rate achievable when B beams are active is given by:

$$\begin{aligned} R(\mathcal{B}) &\approx \mathbb{E}_{\gamma}\left[\sum_{i\in\mathcal{B}}\log_2\left(1+\max_{1\leq k\leq K}\mathsf{SINR}_{k,i}(\mathcal{B})\right)\right] \\ &= B\int_{\gamma=0}^{\infty}\log_2\left(1+\gamma\right)f_{\mathsf{SINR}^*}(\gamma)d\gamma \end{aligned} \tag{5.7}$$

In [Sha05], Sharif and Hassibi derived a closed-form expression for the asymptotic case ($K\rightarrow\infty$) which exhibits the same sum-rate growth as DPC. For a practical scenario with a finite number of users, though, resorting to numerical integration is needed. Still, this expression is tractable when the average SNR of the system is arbitrarily high ($\rho\rightarrow\infty$) [Bos06]. In this case, the pdf of the post-scheduling SINR given by Eq. (5.6) can be re-written as follows:

$$f_{high,\mathsf{SINR}^*}(\gamma) = K\frac{B-1}{(1+\gamma)^B}\left(1-\frac{1}{(1+\gamma)^{B-1}}\right)^{K-1} \tag{5.8}$$

As a consequence, the integral in Eq. (5.7) becomes considerably simpler and a closed-form expression can be derived for the sum-rate (see Appendix 5.A for further details):

$$R_{high}\approx\frac{B}{B-1}\log_2(e)\sum_{k=1}^{K}\frac{1}{k} \tag{5.9}$$

where the term $\sum_{k=1}^{K}\frac{1}{k}$ accounts for the multi-user gain. Two main conclusions can be drawn from the expression above. First, the sum-rate tends to infinity when only one beam is used since, in this case, the sum-rate capacity grows logarithmically with the average SNR. Besides, the sum-rate decreases with the number of active beams due to the $\frac{B}{B-1}$ term. In summary, using only one active beam is the optimum strategy in the high-SNR regime.

In the low-SNR regime (i.e., $\rho\rightarrow 0$), a totally different scenario results. By neglecting the interference term in Eq. (5.3), the post-scheduling pdf can be expressed as:

$$f_{low,\mathsf{SINR}^*}(\gamma) = K\frac{e^{-\frac{\gamma B}{\rho}}}{\rho}\left(1-e^{-\frac{\gamma B}{\rho}}\right)^{K-1}$$

Notice this expression is equal to the pdf of a multi-user system with a SISO configuration (and average $SNR=\rho/B$). This is because we assume i.i.d Rayleigh channel fading and we generate the orthonormal vectors $\mathbf{w}_i$ according to an isotropic distribution. Then, by bearing in mind results obtained in Section 4.4.2 of the previous chapter, the sum-rate can be expressed as:

$$R_{low}\approx -BK\log_2 e\sum_{k=0}^{K-1}\binom{K-1}{k}\frac{(-1)^k}{k+1}e^{B\frac{(k+1)}{\rho}}E_i\left(-B\frac{(k+1)}{\rho}\right) \tag{5.10}$$

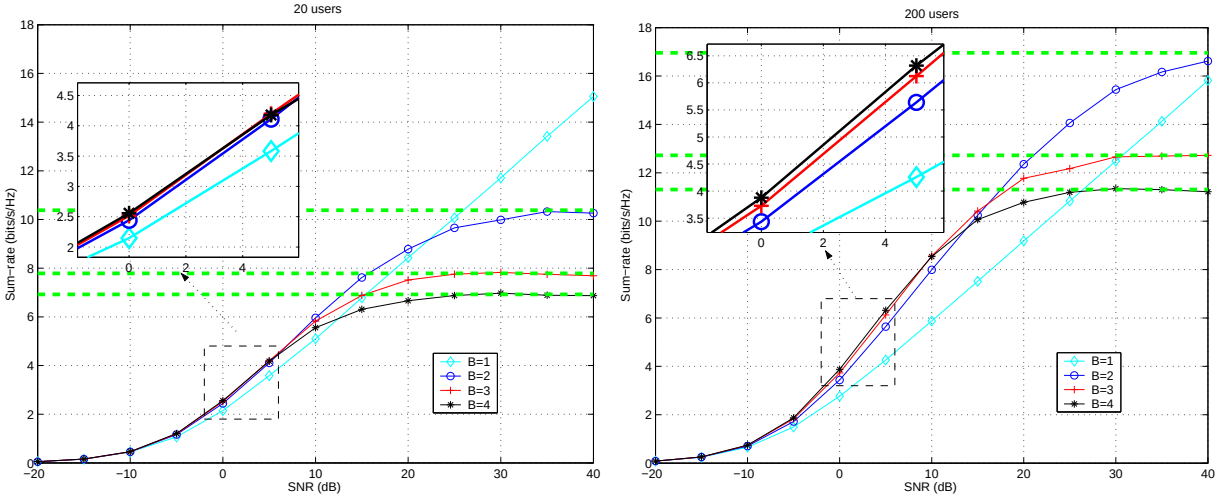


Figure 5.2: Sum-rate vs. SNR for a different number of active beams. Dotted lines correspond to the asymptotic results given by Eq. (5.9). Left, K=20 users. Right, K=200 users.

Table 5.1: Sum-rate performance in low-SNR scenarios (K=200 users, ρ=-5 dB).

	B=1	B=2	B=3	B=4
Simulation	1.485	1.759	1.836	1.893
Eq. (5.10)	1.485	1.848	2.071	2.179

with $E_i(x)$ standing for the exponential integral function ($E_i(x) \triangleq -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$, for $x < 0$).

Table 5.1 and Fig. 5.2 illustrate the accuracy of the approximate sum-rate expressions for the low- and high-SNR regimes, respectively. On the one hand, one can observe in Table 5.1 how both the simulated and approximated results reflect the same trend for a growing number of beams. On the other hand, Fig. 5.2 shows that the proposed high-SNR approximation becomes valid from SNRs equal to 25-30 dB approximately. Besides, the following conclusions can be drawn. In noise-limited scenarios a higher number of active beams turns out to be beneficial. In particular, this is more relevant when the population of users is high (see right-hand plot in Fig. 5.2). Nonetheless, when the system becomes interference-limited (i.e., the SNR is increased or the number of users is low), the use of multiple beams does not pay off. Instead, activating a lower number of beams gives better results.

5.5 Behavior of the Aggregated Throughput

In the previous section, we have analyzed system performance in terms of sum-rate. However, for practical systems with a limited number of AMC modes and realistic coding methods, data link layer throughput provides a closer idea on the actual data rates than sum-rate metrics.

5.5.1 Closed-form Expression

In this case, we can easily derive a closed-form expression for the aggregated (system) throughput. For a given modulation scheme, indexed by variable m, the aggregated throughput can be expressed as[2]

$$\begin{aligned}\eta(\mathcal{B}) &\approx \mathbb{E}_\gamma \left\{ \sum_{i\in\mathcal{B}} b_m \left(1 - PER_m(\max_{1\le k\le K} \mathsf{SINR}_{k,i})\right)\right\} \\ &= B\mathbb{E}_\gamma \left\{ b_m(1-\mathrm{SER}_m(\max_{1\le k\le K} \mathsf{SINR}_{k,i}))^L \right\} \\ &= Bb_m \int_{\gamma=0}^{\infty} (1-\mathrm{SER}_m(\gamma))^L f_{\mathsf{SINR}^*}(\gamma) d\gamma \end{aligned} \tag{5.11}$$

where L stands for the number of symbols in the burst and b_m is the number of bits per symbol. As shown in the previous chapter, the SER for M-QAM modulation schemes can be approximated by:

$$\mathrm{SER}_m(\gamma) \approx b_m 0.2 e^{-1.6\frac{\gamma}{2^{b_m}-1}} = \alpha_m e^{-\beta_m \gamma} \tag{5.12}$$

where α_m and β_m are constellation-dependent parameters[3]. Note that, by using such SER expressions we implicitly assume not only the noise component but also the overall inter-user interference to be Gaussian-distributed. Indeed, symbol constellations do not fulfill this condition but in cases where the number of interferer beams $(B-1)$ is high we can invoke the central limit theorem [Poo97]. Besides, the proposed scheduler is aimed at finding the MSs which maximize the resulting SINR or, equivalently, minimize inter-user interference. Therefore, the relative weight of the interference term in Eq. (5.2) is expected to be low, in particular when the number of users to choose from is high (or, of course in the low-SNR regime). Figure 5.3 illustrates the validity of the Gaussian approximation: even in the $B = 2$-beam case, the approximation is quite accurate for SINRs below 10 dB.

Going back to the derivation of the throughput expressions, the proposed adaptive modulation mechanism selects for each beam i the constellation size maximizing instantaneous data link layer throughput, i.e., satisfying:

$$m_i = \arg\max_{m\in\mathcal{M}} b_m(1-\alpha_m e^{-\beta_m \gamma_{k^*,i}})^L$$

where $\gamma_{k^*,i}$ stands for the SINR corresponding to the scheduled user on beam i ($\gamma_{k^*,i} = \max_{1\le k\le K} \mathsf{SINR}_{k,i}$). From the above expression, the corresponding AMC thresholds ($\gamma_{th,m}$) are obtained for a system with a number of modulation schemes given by the ordered set $\mathcal{M} = \{\text{BPSK,QPSK,16-QAM}\}$[4] as shown in Fig. 5.4. Consequently, the constellation size associated with the measured post-scheduling SINR on beam i is determined according to the following

[2] As in the previous chapters, we will restrict ourselves to uncoded transmissions.

[3] For BPSK, the approximation is also quite accurate by setting $\beta_m = 1$.

[4] In an uncoded system with ORB, the potentially low SINRs seldom support constellation sizes larger than 16-QAM. For that reason, in this chapter we also consider BPSK transmission.

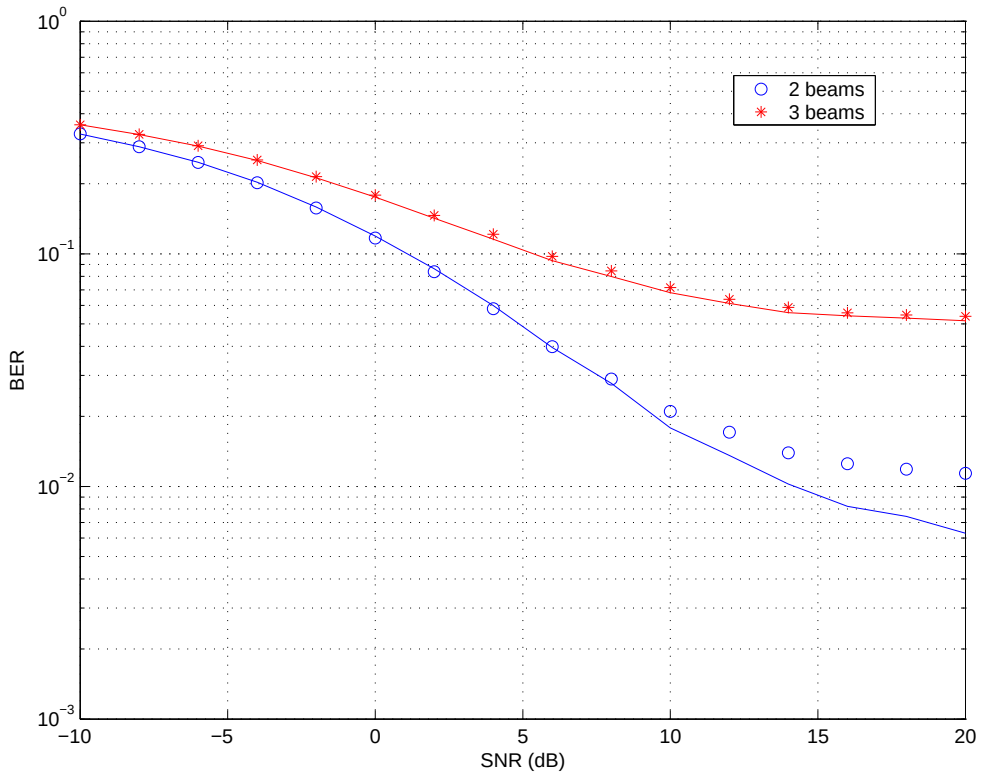


Figure 5.3: Approximated (symbols) vs. real (solid) results for the BER ($K = 5$ users, BPSK).

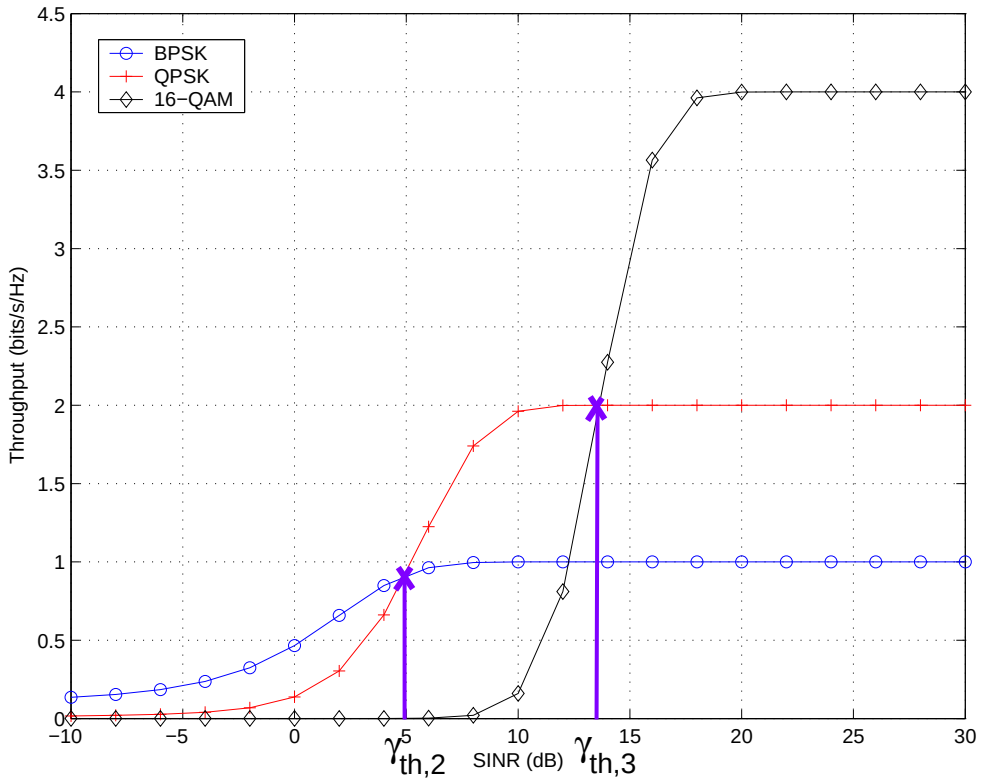


Figure 5.4: Throughput vs. SINR for the different modulation schemes (L=10 symbols).

rule:

$$m_i = m \Longleftrightarrow \gamma_{th,m} \leq \gamma_{k^*,i} < \gamma_{th,m+1}$$

with $\gamma_{th,1} = 0$ and $\gamma_{th,\mathrm{card}(\mathcal{M})+1} = \infty$. Next, by taking into consideration Eqs. (5.11) and (5.12), the aggregated throughput in the presence of adaptive modulation mechanisms can be expressed as:

$$\eta(\mathcal{B}) \approx B \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \int_{\gamma=\gamma_{th,m}}^{\gamma_{th,m+1}} (1-\alpha_m e^{-\beta_m \gamma})^L f_{\mathsf{SINR}^*}(\gamma) d\gamma$$

After some algebraic manipulation (see Appendix 5.B for details), we finally obtain the closed-form expression of the aggregated throughput as follows:

$$\begin{aligned}\eta(\mathcal{B}) \approx BK & \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \sum_{l=0}^{L} \binom{L}{l} (-\alpha_m)^l \sum_{k=0}^{K-1} \binom{K-1}{k} (-1)^k e^{\mu} \mu^c \\ & \times \left[\frac{B}{\rho\mu} \Big(\Gamma\left(1-c, (1+\gamma_{th,m})\mu\right) - \Gamma\left(1-c, (1+\gamma_{th,m+1})\mu\right) \Big) \right. \\ & \left. + (B-1) \Big(\Gamma\left(-c, (1+\gamma_{th,m})\mu\right) - \Gamma\left(-c, (1+\gamma_{th,m+1})\mu\right) \Big) \right] \end{aligned} \tag{5.13}$$

where $\mu = \beta_m l + \frac{B}{\rho}(k+1)$, $c = (k+1)(B-1)$ and $\Gamma(\alpha, x)$ stands for the complementary incomplete gamma function ($\Gamma(\alpha, x) \triangleq \int_x^\infty e^{-t} t^{\alpha-1} dt$).

5.5.2 Asymptotic Analysis

From the analytical expressions obtained in the previous subsection, it is difficult to give some insight about the behavior of the aggregated throughput. In order to alleviate that, we analyze the asymptotic SNR regime.

By assuming ($\rho \to \infty$), one can readily observe that the behavior of Eq. (5.13) depends on the number of active beams B. More specifically, the analysis can be encompassed in two different ways depending on if a single-beam strategy is carried out or several beams are activated.

Single Active Beam (SAB)

When we particularize the aggregated throughput expression to the case where only a single beam ($B = 1$) is activated (i.e., $c = 0$), Eq. (5.13) can be rewritten as:

$$\begin{aligned}\eta_{SAB} \approx & \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m K \sum_{l=0}^{L} \binom{L}{l} (-\alpha_m)^l \sum_{k=0}^{K-1} \binom{K-1}{k} (-1)^k e^{\beta_m l + \frac{k+1}{\rho}} \\ & \times \frac{1}{\rho\beta_m l + k + 1} \left(e^{-(1+\gamma_{th,m})\left(\beta_m l + \frac{k+1}{\rho}\right)} - e^{-(1+\gamma_{th,m+1})\left(\beta_m l + \frac{k+1}{\rho}\right)} \right) \end{aligned}$$

where the equality $\Gamma(1,x) = e^{-x}$ has been used. When $\rho \to \infty$, all the elements in the summations with $l \neq 0$ tend to zero. For $l = 0$, summations also tend to zero except for those elements related to $m = \text{card}(\mathcal{M})$[5]. Then, we can express the above equation as follows:

$$\begin{aligned}\eta_{high,SAB} &\approx b_{\text{card}(\mathcal{M})} K \sum_{k=0}^{K-1} \binom{K-1}{k} \frac{(-1)^k}{k+1} \\ &= b_{\text{card}(\mathcal{M})} \sum_{t=1}^{K} \binom{K}{t} (-1)^{t-1} \\ &= b_{\text{card}(\mathcal{M})}\end{aligned}$$

Differently from the sum-rate case, a saturation effect is observed in the asymptotic SNR regime. Then, using a single active beam may not be the optimum solution for the high-SNR region, specially if a low constellation size is considered for the highest available AMC level. Instead, the transmit power can be distributed in order to increase the degrees of freedom of the system by activating additional beams.

Multiple Active Beams (MAB)

For the case $B > 1$ (i.e., $c \neq 0$) and $\rho \to \infty$, one can notice the following two facts:

$$\begin{aligned}\lim_{\rho\to\infty} \mu^c &= \lim_{\rho\to\infty} \left(\beta_m l + \frac{B}{\rho}(k+1)\right)^c = 0 && \text{if } l = 0 \\ \lim_{\rho\to\infty} \frac{1}{\rho\mu} &= \lim_{\rho\to\infty} \frac{1}{\rho\beta_m l + B(k+1)} = 0 && \text{if } l \neq 0\end{aligned}$$

Then, Eq. (5.13) can be simplified as follows:

$$\begin{aligned}\eta_{high,MAB} \approx BK \sum_{m=1}^{\text{card}(\mathcal{M})} b_m \sum_{l=1}^{L} \binom{L}{l} (-\alpha_m)^l \sum_{k=0}^{K-1} \binom{K-1}{k} (-1)^k e^{\beta_m l} (\beta_m l)^c \\ \times (B-1)\Big(\Gamma\left(-c, (1+\gamma_{th,m})\beta_m l\right) - \Gamma\left(-c, (1+\gamma_{th,m+1})\beta_m l\right)\Big)\end{aligned}$$

Then, by using the following inequality:

$$\Gamma(-m,x) = \int_x^\infty \frac{e^{-t}}{t^{m+1}} dt \leq \frac{1}{x^{m+1}} \int_x^\infty e^{-t} dt = \frac{e^{-x}}{x^{m+1}}$$

an upper-bound can be obtained for the aggregated throughput:

$$\begin{aligned}\eta_{high,MAB} \lesssim B(B-1)K \sum_{m=1}^{\text{card}(\mathcal{M})} b_m \sum_{l=1}^{L} \binom{L}{l} (-\alpha_m)^l \sum_{k=0}^{K-1} \binom{K-1}{k} \frac{(-1)^k}{\beta_m l} \\ \times \left(\frac{e^{-\gamma_{th,m}\beta_m l}}{(1+\gamma_{th,m})^{(k+1)(B-1)+1}} - \frac{e^{-\gamma_{th,m+1}\beta_m l}}{(1+\gamma_{th,m+1})^{(k+1)(B-1)+1}}\right)\end{aligned}$$

[5]Notice that $\gamma_{th,\text{card}(\mathcal{M})+1} = \infty$.

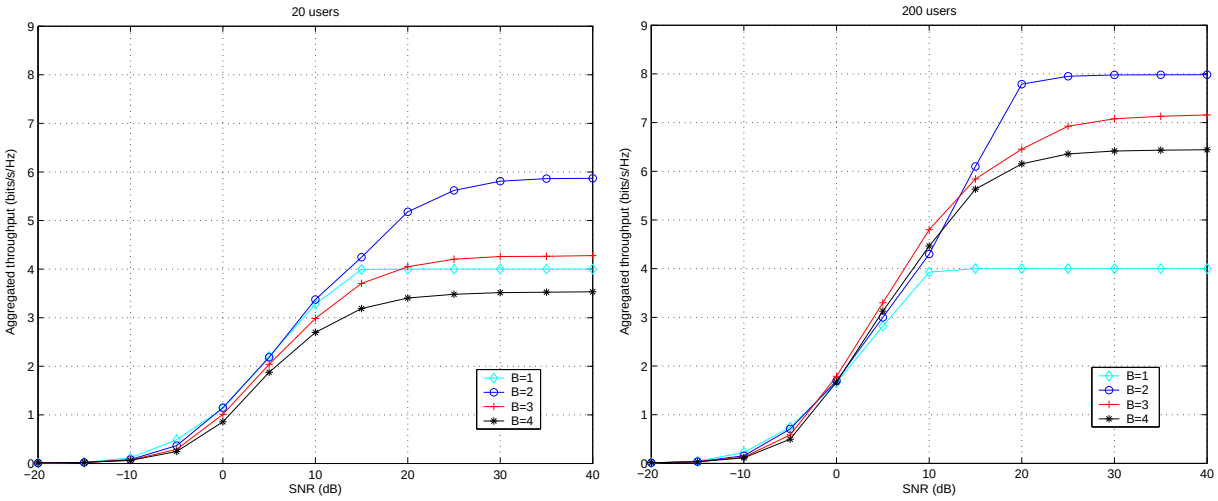


Figure 5.5: Aggregated throughput vs. SNR for a different number of active beams (L=10 symbols). Left, K=20 users. Right, K=200 users.

Notice that the above bound is valid (up to the tight approximation) due to the second term of the difference is always lower than the first one given the following inequality $\gamma_{th,m+1} > \gamma_{th,m}$. The main particularity of this bound is that it tends to 0 when $B \to \infty$. Then, the aggregated throughput also tends to 0, that is, activating more than two beams is not beneficial in the high-SNR regime[6].

In conjunction with the previous asymptotic result, one can conclude that the best strategy depends on the size of the highest available AMC level. In some situations, using only one active beam may be the best option in the asymptotic regime. However, if the constellation size of the highest AMC level is small, a more appropriate choice would be distributing the transmit power among two active beams (since setting $B > 2$ is less efficient). Simulation results shown in Fig. 5.5 are aligned with and, thus, confirm these conclusions. More precisely, it is observed that the single beam solution quickly saturates. For that reason, increasing the number of beams can be a good option in the high-SNR region as long as the number of active beams is not higher than two. Besides, in scenarios with a high population, the interference generated by the beams is reduced and, as a result, a higher gap is observed between results obtained with $B = 1$ and the remaining number of active beams.

As for the low-SNR, an asymptotic analysis cannot be carried out as the BER expression is loose approximation in this region [Chu01]. Still, one can notice from the simulation results that using several beams is not an appropriate strategy for noise-limited scenarios. This is because PER levels are very low to support several active beams, even with the lowest modulation scheme (BPSK). Spatial multiplexing capabilities can be exploited in the mid-SNR region of high-populated cells (see right-hand plot in Fig. 5.5), where the SNR is neither low enough to

[6]In this case, we are assuming that the number of active beams is $B \geq 2$.

penalize PER behavior nor high enough to enter the interference-limited region.

In summary, due to granularity and bit cap effects associated with the use of practical modulation schemes, different conclusions are drawn with respect to the sum-rate case. Nonetheless, we have observed in both cases that using all the active beams is not the best strategy in a scenario with a practical number of users. This fact will be used in Section 5.7 to develop a beam selection approach but, first, the aggregated throughput is also analyzed in a scenario with quantized feedback.

5.6 Impact of Feedback Quantization on Aggregated Throughput

In this section, we are interested in adapting the expression of the aggregated throughput to the (realistic) case where the SINRs conveyed over the feedback channel are quantized versions of the analog ones. Motivated by the work in [Flo03], this section is devoted to find out how many feedback bits are required to exploit MUD gains in an ORB context. First, we derive a closed-form expression for the aggregated throughput with quantized feedback. Next, a quantization law based on cross-layer designs is proposed. Finally, we quantify the effect of feedback quantization on system performance with the help of numerical evaluation.

5.6.1 Closed-form Expression

We start by defining $\mathcal{Q} = \{q_1, q_2, ..., q_{2^{L_q}}\}$ as the set of quantization levels. After an arbitrary user k identifies the beam i^* with the highest SINR, $\gamma_{k,i^*} = \max_{i \in \mathcal{B}} \mathsf{SINR}_{k,i}$, it is quantized according to the following rule:

$$Q(\gamma_{k,i^*}) = \gamma_{q_j} \quad \text{if } \gamma_{q_j} \leq \gamma_{k,i^*} < \gamma_{q_{j+1}}$$

where $\gamma_{q_j}, j = 1, ..., \text{card}(\mathcal{Q})$ are the different SINR thresholds associated with the quantization levels[7]. Next, an $L_t = L_q + L_b$-bit message is sent over the feedback channel, with $L_b = \lceil log_2(B) \rceil$ bits devoted to beam signalling.

Now, we define $\mathcal{A}_{k,i}$ as the event that user k is selected for transmission on beam i, and we obtain the probability of this event conditioned on the fact that $\gamma_{k,i}$ belongs to the quantization level q_j as follows [Flo03]:

$$\text{Prob}(\mathcal{A}_{k,i} | \gamma_{k,i} \in q_j) = \sum_{l=0}^{K-1} \frac{1}{l+1} \binom{K-1}{l} \text{Prob}\left(\#\text{ users different from } k \text{ in } q_j = l\right) \times \text{Prob}\left(\#\text{ users different from } k \text{ in levels lower than } q_j = K - l - 1\right)$$

[7]This quantization rule results into a conservative, but reliable, assignment of AMC modes at the transmitter.

The expression above has been derived by taking into account that user k is selected when q_j is the highest quantization level achieved by the SINR of any user. In other words, the SINR of the rest of users must lie in this level or in quantization levels lower than it. Then, each element of the summation takes into consideration that, apart from user k, l users lie in quantization level q_j and the rest of users must lie in lower quantization levels. The term $1/(l+1)$, on the other hand, is referred to the probability that user k is scheduled for transmission for each value of l, since a random user must be selected if more than one user lie in q_j.

In the proposed homogeneous scenario, $\text{Prob}(\mathcal{A}_{k,i}|\gamma_{k,i} \in q_j)$ can be easily computed as it does not depend on k and i. More specifically, by modeling the SINR of all the users with a generic random variable γ with CDF and pdf expressions given by Eqs. (5.4) and (5.5), respectively, the following expression can be obtained:

$$\text{Prob}(\mathcal{A}_{k,i}|\gamma_{k,i} \in q_j) = \sum_{l=0}^{K-1} \frac{1}{l+1} \binom{K-1}{l} \left(\text{Prob}(\gamma \in q_j)\right)^l \left(\text{Prob}\left(\gamma \in \bigcup_{p<j} q_p\right)\right)^{K-l-1}$$

After some algebraic manipulation and bearing in mind that $\text{Prob}(\gamma \in q_j) = F_{\mathsf{SINR}}(\gamma_{q_{j+1}}) - F_{\mathsf{SINR}}(\gamma_{q_j})$ and $\text{Prob}\left(\gamma \in \bigcup_{p<j} q_p\right) = F_{\mathsf{SINR}}(\gamma_{q_j})$, it can be readily shown that:

$$\text{Prob}(\mathcal{A}_{k,i}|\gamma_{k,i} \in q_j) = \frac{1}{K} \frac{\left(F_{\mathsf{SINR}}(\gamma_{q_{j+1}})\right)^K - \left(F_{\mathsf{SINR}}(\gamma_{q_j})\right)^K}{F_{\mathsf{SINR}}(\gamma_{q_{j+1}}) - F_{\mathsf{SINR}}(\gamma_{q_j})} \tag{5.14}$$

With the derived probability, one can obtain the throughput share corresponding to user k on beam i as:

$$\eta_{k,i} \approx \sum_{j=1}^{\text{card}(\mathcal{Q})} \text{Prob}(\mathcal{A}_{k,i}|\gamma_{k,i} \in q_j) b_{m_j} \int_{\gamma=\gamma_{q_j}}^{\gamma_{q_{j+1}}} (1 - \alpha_{m_j} e^{-\beta_{m_j}\gamma})^L f_{\mathsf{SINR}}(\gamma) d\gamma \tag{5.15}$$

where each integration interval corresponds to the expected throughput in each quantization level. Besides, we assume that the modulation scheme for quantization level q_j is selected according to the quantized value of $\gamma_{k,i}$, that is:

$$m_j = m \Longleftrightarrow \gamma_{th,m} \leq \gamma_{q_j} < \gamma_{th,m+1}$$

Finally, by plugging Eqs. (5.14) and (5.5) into Eq. (5.15), we can write the aggregated throughput expression as:

$$\begin{aligned}\eta(\mathcal{B},\mathcal{Q}) \approx \sum_{k=1}^{K}\sum_{i\in\mathcal{B}} \eta_{k,i} = B \sum_{j=1}^{\text{card}(\mathcal{Q})} b_{m_j} \sum_{l=0}^{L} \binom{L}{l} \\ \times (-\alpha_{m_j})^l \frac{\left(F_{\mathsf{SINR}}(\gamma_{q_{j+1}})\right)^K - \left(F_{\mathsf{SINR}}(\gamma_{q_j})\right)^K}{F_{\mathsf{SINR}}(\gamma_{q_{j+1}}) - F_{\mathsf{SINR}}(\gamma_{q_j})} \\ \times \int_{\gamma=\gamma_{q_j}}^{\gamma_{q_{j+1}}} \left(\frac{B}{\rho}(1+\gamma) + B - 1\right) \frac{e^{-\gamma\left(\beta_{m_j} l + \frac{B}{\rho}\right)}}{(1+\gamma)^B} d\gamma\end{aligned}$$

where we have added the shares of throughput obtained with all the users ($k = 1, ..., K$) and all the active beams ($i \in \mathcal{B}$). The integral term in the above expression resembles that of Eq. (5.17) in Appendix 5.B. Therefore, one can once again follow the procedure detailed in Appendix 5.B to finally obtain the aggregated throughput with quantized feedback:

$$\begin{aligned}\eta(\mathcal{B},\mathcal{Q}) \approx & B \sum_{j=1}^{\mathrm{card}(\mathcal{Q})} b_{m_j} \sum_{l=0}^{L} \binom{L}{l} (-\alpha_{m_j})^l \frac{\left(F_{\mathsf{SINR}}(\gamma_{q_{j+1}})\right)^K - \left(F_{\mathsf{SINR}}(\gamma_{q_j})\right)^K}{F_{\mathsf{SINR}}(\gamma_{q_{j+1}}) - F_{\mathsf{SINR}}(\gamma_{q_j})} e^{\mu_q} \mu_q^{c_q} \\ & \times \left[\frac{B}{\rho\mu_q} \left(\Gamma\left(1 - c_q, (1+\gamma_{q_j})\mu_q\right) - \Gamma\left(1 - c_q, (1+\gamma_{q_{j+1}})\mu_q\right) \right) \right. \\ & \left. + (B-1) \left(\Gamma\left(-c_q, (1+\gamma_{q_j})\mu_q\right) - \Gamma\left(-c_q, (1+\gamma_{q_{j+1}})\mu_q\right) \right) \right]\end{aligned}$$

where $\mu_q = \beta_{m_j} l + \frac{B}{\rho}$ and $c_q = B - 1$.

5.6.2 Quantization Laws

So far, nothing has been said concerning the quantization law (either uniform, non-uniform, etc.). We can shed some light into that issue by bearing in mind that, according to the adopted max-SINR scheduling rule, only the highest SNIRs are relevant in terms of multi-user diversity exploitation gains (see Chapter 4) and link throughput. Therefore, we should avoid uniform quantization laws (possibly in-between some limiting values) and, instead, go for a non-uniform one with smaller quantization intervals in the high-SINR region. By doing so, we minimize the clipping rate, that is, the probability that the quantized metric exceeds the highest quantization threshold), and its negative effects on system performance (further details on clipping rate effects can be found e.g. in [Jia04]). In summary and as depicted in Fig. 5.6, the quantization thresholds could be given by the inverse of either the *post-scheduling* CDF function or the *pre-scheduling*, i.e. individual, CDFs. Intuitively, a quantization law tailored to the post-scheduling SINR should give better results since this is directly related with the scheduling rule. Numerical evaluation in the next subsection confirm this extent.

In summary, the SINR thresholds related to the different quantization levels are selected as follows:

$$\gamma_{q_j} = F^{-1}_{\mathsf{SINR}^*}\left(\frac{j-1}{2^{L_q}}\right) \text{ for } j = 1, ..., \mathrm{card}(\mathcal{Q})$$

with $\gamma_{q_1} = 0$ and $\gamma_{q_{\mathrm{card}(\mathcal{Q})+1}} = \infty$. As a final remark, notice this is a *cross-layer* quantization law since quantization levels in the physical layer depend on the number of admitted users which, ultimately, is decided by access control mechanisms in the data link layer.

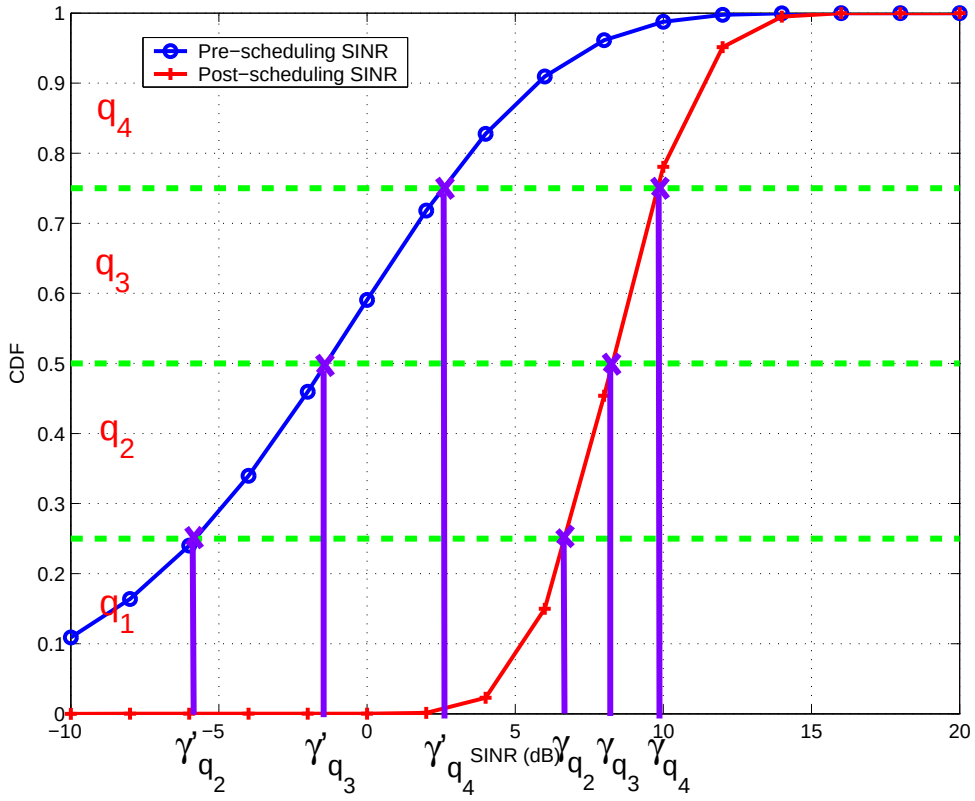


Figure 5.6: Uniform and Improved Quantization procedures for the SINR quantization (SNR=10 dB, K=20 users, B=2 active beams).

5.6.3 Numerical Results and Discussion

We consider a system with a number of active users in the range $K = 10, ..., 100$ transmitting data packets with $L = 10$ symbols in each. Since Fig. 5.5 revealed that little improvement can be obtained with four active beams, we restrict the analysis to a scenario with $M = 3$ transmit antennas.

First, we evaluate performance in terms of aggregated throughput for the quantization methodologies based on pre- and post-scheduling statistics. In Fig. 5.7, one can clearly appreciate the accuracy of the (approximate) closed-form expressions of the aggregated throughput derived in the previous subsections. Apart from that, it becomes apparent that the post-scheduling based criterion significantly outperforms its pre-scheduling counterpart for the whole range of users. For an increasing number of active users, the gap between both curves becomes wider. This is due to an increased clipping rate (the SINR of the scheduler user is potentially higher when K becomes larger) which penalizes much more the quantization law based on pre-scheduling statistics. When adopting a quantization law based on the post-scheduling CDF we can clearly see that most of the MUD can be efficiently captured with as few as $L_q = 4$ or $L_q = 2$ quantization bits. Last, one can appreciate in the right-hand plot of Fig. 5.7 how a lower performance

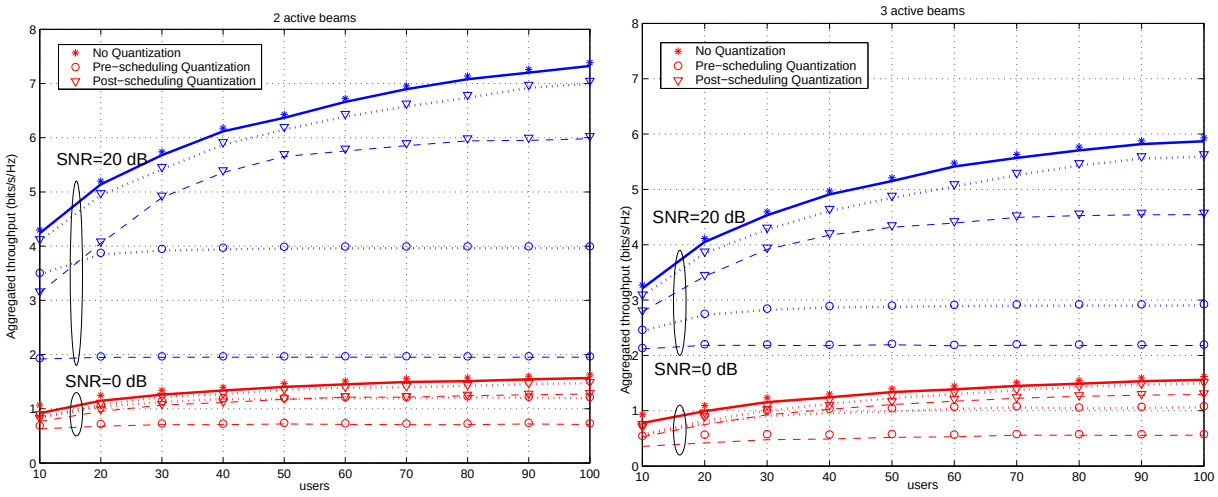


Figure 5.7: Aggregated throughput vs. number of users for the different quantization methodologies and number of quantization bits, L_q (Dotted lines: 4 quantization bits, dashed lines: 2 quantization bits, symbols: simulated results, curves: analytical expressions). Left, B=2 active beams. Right, B=3 active beams.

is obtained with three active beams. As showed in the previous section, the considered number of users and values of SNR are not appropriate to activate three beams. Further, one can also observe that in this case the gap between results attained with pre- and post-scheduling methodologies is narrower due to the lower capability of generating SINR peaks.

The curves depicted in Fig. 5.8 provide some more insights on the impact of quantization on system performance. First, the higher the SNR the larger the impact of quantization since, in this case, the range of SINR fluctuations is larger. As commented above, this is emphasized in scenarios with a high number of users. However, most of the MUD can be effectively captured for a very low number of bits. More precisely, in the worst situation in terms of quantization impact (SNR=20 dB and K=60 users) and for both the two and three active beams configurations, the proposed quantization law attains approximately 81% and 92% of the *analog* throughput by just using 2 or 3 feedback bits, respectively.

Finally, we compare results obtained with two and three active beams in a scenario where the latter one achieve better results in the analog feedback case (SNR=10 dB and K=100 users). In particular, the impact of feedback quantization is plotted as a function of the total number of bits in the feedback channel L_t. In such a case, one must bear in mind that $L_b = \lceil log_2(B) \rceil$ bits are required to indicate the selected beam. As observed in Fig. 5.9, for a number of feedback bits lower than four, MUD gains are better exploited with two active beams since only one bit is needed for beam signalling. As the number of feedback bits is increased, the three active beams configuration is benefited from an improved resolution in the SINR quantization. With L_t equal to five feedback bits, performance obtained with three active beams is even better than that obtained in the analog feedback case with two beams activated.

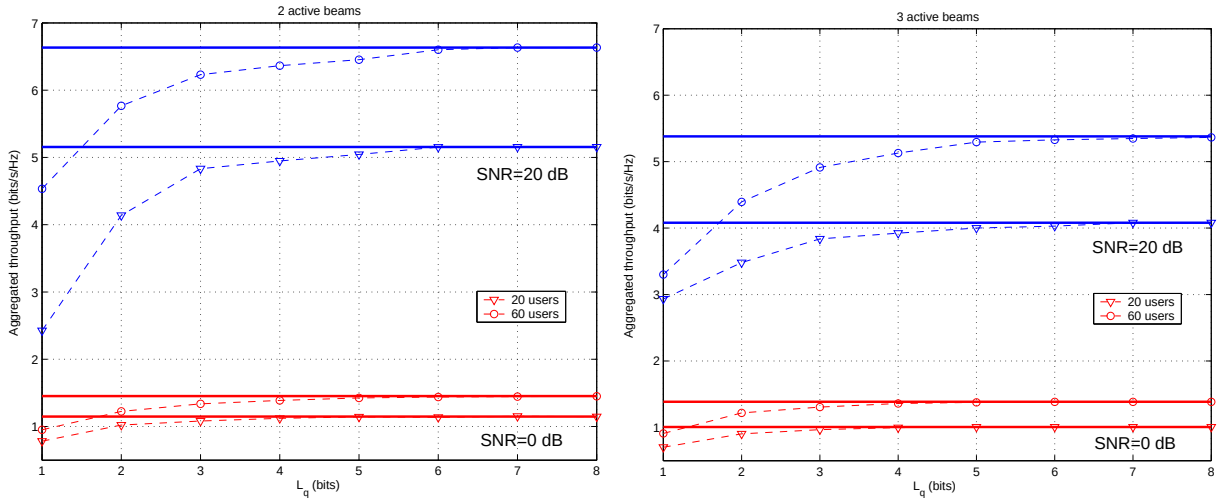


Figure 5.8: Aggregated throughput vs. number of feedback bits L_t for different number of users (Solid lines: unquantized throughput). Left, B=2 active beams. Right, B=3 active beams.

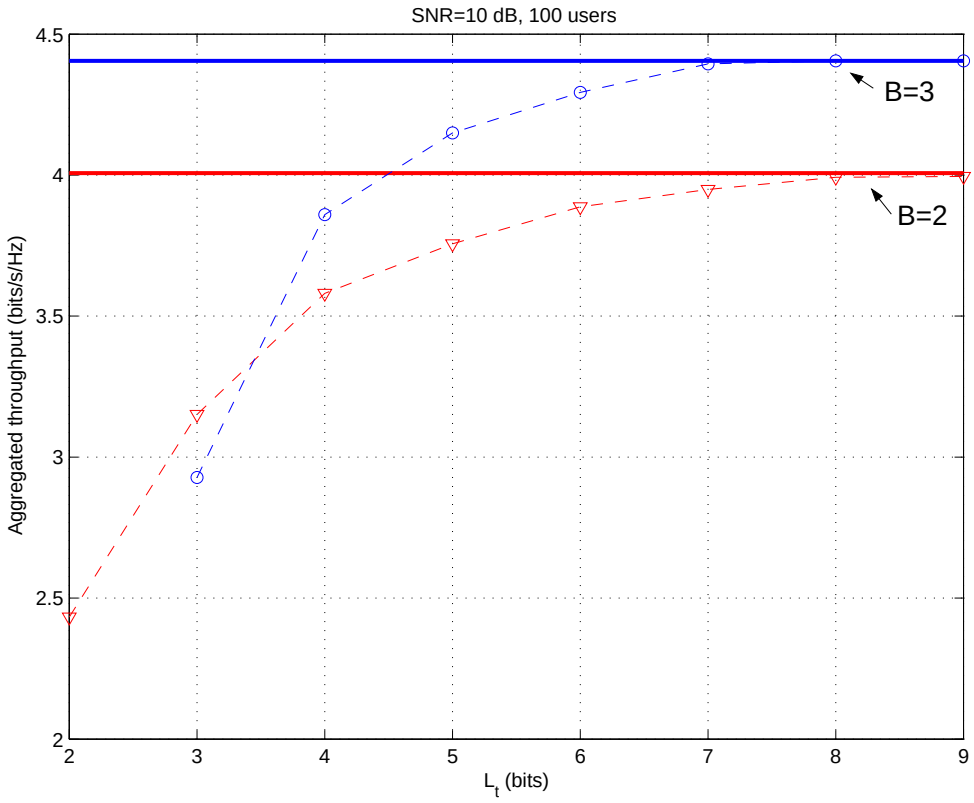


Figure 5.9: Aggregated throughput vs. number of feedback bits, L_t, for different number of active beams (Solid lines: unquantized throughput, $K = 100$ users and SNR=10 dB).

In summary, with a few number of feedback bits, most of the MUD gains provided by ORB can be exploited. In order to efficiently exploit those gains, different strategies must be considered at the base station in accordance with the number of available bits in the feedback channel, the number of users and the SNR.

5.7 Adaptive Beam Selection

In the previous sections, we have shown that using all the set of active beams may not be the optimal transmission configuration. For that reason, we now derive several beam selection algorithms capable of identifying the best subset of beams (and users) according to scenario conditions. These algorithms are quite interesting for sparse networks because the number of users in the system can be *virtually* increased. That is, by using beam selection, the number of SINR combinations is larger and, then, system performance is improved as if the number of users was increased[8].

We begin this section by proposing some beam selection algorithms with different trade-offs in terms of system performance vs. computational complexity and feedback requirements. Afterwards, numerical results are provided to analyze the proposed algorithms.

5.7.1 Beam Selection Algorithms

In a system performing ORB, the *Optimum Beam Selection* algorithm consists in conducting an exhaustive search over all the possible subsets of beams and users. This procedure is quite expensive in terms of computational complexity as the number of SINRs to be tested exponentially grows with the number of available beams. More precisely, for a fixed number of active beams B, a total of $\binom{M}{B}KB$ SINRs must be computed in order to find the best transmission configuration. Then, by considering all the possible number of active beams, the number of SINR computations is increased to a total of $\sum_{B=1}^{M}\binom{M}{B}KB = KM2^{M-1}$ operations. Further, it is worth noting that this algorithm requires all the gains $|\mathbf{h}_k^T\mathbf{w}_i|^2$ to be known. Therefore, it is necessary for each user to report M integer numbers to the base station over the feedback channel.

In order to reduce the computational complexity and the amount of required feedback information of the beam selection procedure, next we present some sub-optimum approaches:

[8] In the single beam case (B=1), for instance, the number of equivalent users it is equal to MK. This is because $\mathsf{SINR}_{k,i}$ for $k = 1, ..., K$ and $i = 1, ..., M$ are i.i.d distributed in this case.

Bottom-up Trellis Beam Selection

The first sub-optimum methodology is based on a bottom-up procedure. More precisely, the algorithm starts by selecting the best user for each beam in terms of the considered performance metric (sum-rate or aggregated throughput). After that, the selected users in the first step are combined (with their associated beams) in order to find the best combination with two active beams. The algorithm is iterated until the combination where all the beams are active ($B = M$) is reached. Basically, the objective is to reduce the computational cost by focusing on the users achieving the highest gains with only one active beam. By doing so, KM computations are still needed in the first level but only $\binom{M}{B}B$ operations are required in subsequent ones. As a result, complexity drops to $M2^{M-1} + M(K-1)$ SINR computations.

Top-down Trellis Beam Selection

In the bottom-up procedure, we are restricting the search to those users maximizing system performance when only one beam is activated. However, this subset of users may not be adequate when the number of beams increases and interference comes into play. For that reason, we propose a similar approach starting on the maximum number of active beams ($B = M$), for which the best subset of users is found. Then, the algorithm is iterated by removing one beam in each step. Again, user-beam pairs selected in the first step are kept. As a result, KM and $\binom{M}{B}B$ operations are also needed in the initializing and subsequent steps of the algorithm, respectively. Therefore, the number of required computations is the same as the bottom-up approach.

Greedy Beam Selection

Both the top-down and bottom-up procedures restrict the search to the best user subset when only one or all the beams are active, respectively. In order to extend the search to a larger set of users, we propose a greedy beam selection procedure. Specific details about the proposed greedy algorithm can be found in Table 5.2 but the basic idea consists in selecting in each step the pair user-beam leading to a higher increase of sum-rate (or aggregated throughput)[9]. The algorithm is iterated until the configuration with all the active beams is reached and, then, the best sub-set with $B = j^*$ active beams is selected. In the first iteration, MK computations are needed to find the best user-beam pair. Next, K SINRs must be computed when a new beam is added and this operation must be repeated for the remaining beams. Therefore, the

[9]In Table 5.2 the algorithm is generalized for both the sum-rate and aggregated throughput metrics. To do so, we have defined function $\Theta(x)$ as follows:

$$\Theta(x) = \begin{cases} \log_2(x) & \text{for sum-rate} \\ b_{m_i}(1 - \alpha_{m_i} e^{-\beta_{m_i} x})^L & \text{for aggregated throughput} \end{cases}$$

Table 5.2: Greedy Beam Selection Algorithm

1. Set j=1, $\mathcal{K}_1$=$\{1,\ldots,K\}$ and $\mathcal{B}_1$=$\{\mathbf{w}_1,\ldots,\mathbf{w}_M\}$.
2. Compute the best pair user-beam for the case with only one active beam as:
$$(k_1,i_1)=argmax_{(k,i)}\rho|\mathbf{h}_k^T\mathbf{w}_i|^2,\ \forall(k,i)\in\mathcal{K}_1\times\mathcal{B}_1$$
3. Compute $R_1=\Theta(1+\rho|\mathbf{h}_{k_1}^T\mathbf{w}_{i_1}|^2)$.
4. Set $j=j+1$, $\mathcal{K}_j$=$\mathcal{K}_{j-1}-\{k_{j-1}\}$ and $\mathcal{B}_j$=$\mathcal{B}_{j-1}-\{\mathbf{w}_{i_{j-1}}\}$.
5. Compute the best pair user-beam that can be added to the system as:
$$(k_j,i_j)=$$
$$=argmax_{(k,i)}\left\{\Theta\left(\frac{|\mathbf{h}_k^T\mathbf{w}_i|^2}{j/\rho+\sum_{s=1}^{j-1}|\mathbf{h}_k^T\mathbf{w}_{i_s}|^2}\right)+\sum_{p=1}^{j-1}\Theta\left(\frac{|\mathbf{h}_{k_p}^T\mathbf{w}_{i_p}|^2}{j/\rho+|\mathbf{h}_{k_p}^T\mathbf{w}_i|^2+\sum_{\substack{s=1\\s\neq p}}^{j-1}|\mathbf{h}_{k_p}^T\mathbf{w}_{i_s}|^2}\right)\right\},$$
$$\forall(k,i)\in\mathcal{K}_j\times\mathcal{B}_j$$
6. Compute
$$R_j=\sum_{p=1}^{j}\Theta\left(\frac{|\mathbf{h}_{k_p}^T\mathbf{w}_{i_p}|^2}{j/\rho+\sum_{\substack{s=1\\s\neq p}}^{j}|\mathbf{h}_{k_p}^T\mathbf{w}_{i_s}|^2}\right)$$
7. If $j<M$, go to step 4. Otherwise go to step 8.
8. Set $j^*=argmax_jR_j$.
9. The algorithm is finished and the set of selected users and beams is the following:
$$(k_1,\mathbf{w}_{i_1}),...,(k_{j^*},\mathbf{w}_{i_{j^*}})$$

algorithm performs $(M-B+1)K$ computations in each iteration with a total computational cost of $\sum_{B=1}^{M}(M-B+1)K=\frac{KM}{2}(M+1)$ SINR operations.

Enhanced Greedy Beam Selection

In the greedy beam selection scheme proposed above, the overall performance depends on the user-beam pair of the first iteration (i.e., in the absence of inter-user interference). Instead, we can defer such decision to the second iteration where some inter-user interference is already present. In other words, we initialize the algorithm by identifying the *best* user for each beam $i=1,...,M$ (i.e., in the absence of interference). Then, we run the greedy algorithm M times taking as a starting point each user-beam pair obtained in the initialization. As a result, $\frac{KM}{2}(M+1)-K(M-1)$ operations should be done each time the algorithm is executed and, given that this procedure should be repeated M times, the total computational complexity amounts to

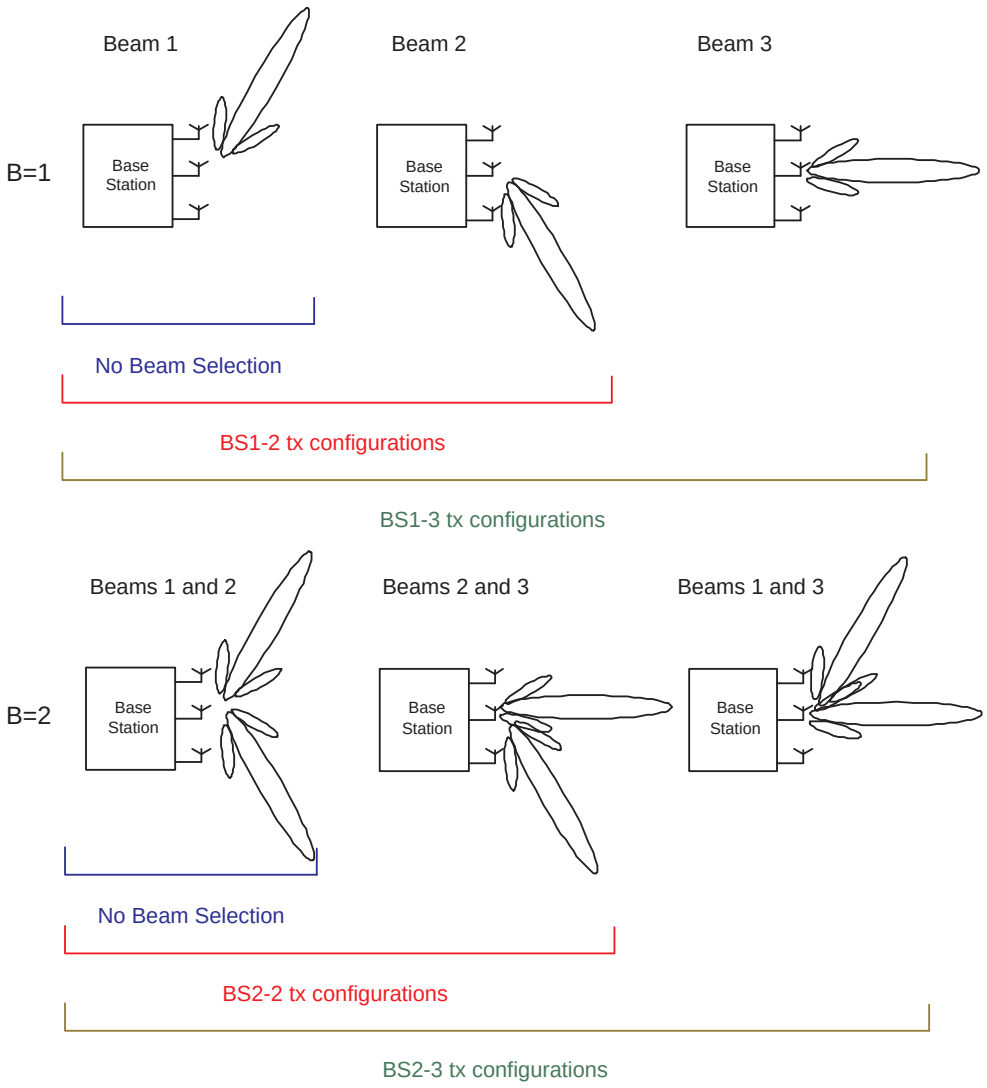


Figure 5.10: Beam selection schemes according to the restrictions imposed in the feedback channel.

$\frac{KM}{2}(M^2 - M + 2)$ SINR computations.

Restricted Beam Selection

Finally, we present a methodology where the optimum beam selection procedure is restricted to a predetermined number of active beams B. In other words, all the possible transmission configurations with B active beams are tested only. By doing so, the number of SINR computations is reduced to $\binom{M}{B}KB = \frac{KM!}{(M-B)!(B-1)!}$ operations. This strategy is very appropriate for those situations where the optimum number of active beams is known beforehand. For instance, it was shown in Section 5.4 that the optimum strategy is using a single active beam when the sum-rate metric is considered in a high-SNR scenario. In the sequel, these algorithms will be called BSX, where X will be the number of active beams.

Besides, when the number of possible transmission configurations $\binom{M}{B}$ is equal or lower than M, each user can report the highest SINRs for each configuration instead of all the gains $|\mathbf{h}_k^T \mathbf{w}_i|^2$. Then, by sending only SINRs associated to a limited number of transmission configurations, sub-optimum approaches can be derived in terms of system performance vs. feedback requirements constraints. For instance, consider a scenario with $M = 3$ available beams. In this case, three transmission configurations are available for both the BS1 and BS2 procedures (see Fig. 5.10). The system can then be simplified by considering only one (No Beam Selection) or two (BSX-2 tx) transmission configurations as shown in the figure. By doing so, the total number of bits L_t required in the feedback channel is reduced. For the BS1 case, only the quantized version of the highest SINR and the beam index must be sent to the BS. As a consequence, the number of required bits amounts to $L_t = L_q + \lceil \log_2 N_{tx_{conf}} \rceil$, where $N_{tx_{conf}}$ stands for the number of transmission configurations. Conversely, $L_t = N_{tx_{conf}}(L_q + \lceil \log_2 B \rceil)$ bits must be reported in the BS2 scheme. This is due to the necessity of reporting the quantized versions of the SINRs and the index beams for each of the different transmission configurations as more than one beam are activated.

5.7.2 Numerical Results and Discussion

As in the previous section, we restrict the analysis to a scenario with $M = 3$ transmit antennas due to the conclusions drawn from Figs. 5.2 and 5.5.

We start by showing system performance in terms of sum-rate. In Fig. 5.11, the different beam selection methodologies are compared in the low SNR regime (SNR=0 dB). As expected, the best performance is obtained with the optimum approach. Regarding the sub-optimum approaches, performance losses can be observed for both the bottom-up and greedy methodologies, whereas most of the sum-rate gains can be achieved with the top-down and enhanced trellis approaches. This is because using several active beams may be beneficial when the SNR is low and, then, incorrect decisions made in the first step of the greedy and bottom-up algorithms penalize system performance. This effect is even clearer when the number of users increases. As for the restricted beam selection procedures, it is observed that the best results are obtained with the BS2 approach.

When the SNR increases (see Fig. 5.12), the system becomes interference-limited. As a result, the impact of a wrong user selection on the denominator of the SINR (see Eq. (5.3)) is emphasized. It can also be observed that results associated to BS1 are improved, whereas performance associated with BS3 is deteriorated. This is because the optimum solution tends to use a reduced number of active beams when the SNR is increased. This effect can be clearly observed in a scenario with SNR=50 dB. Results corresponding to this scenario are not included for brevity, but it is noteworthy that the same results are obtained with the optimum, bottom-up, greedy,

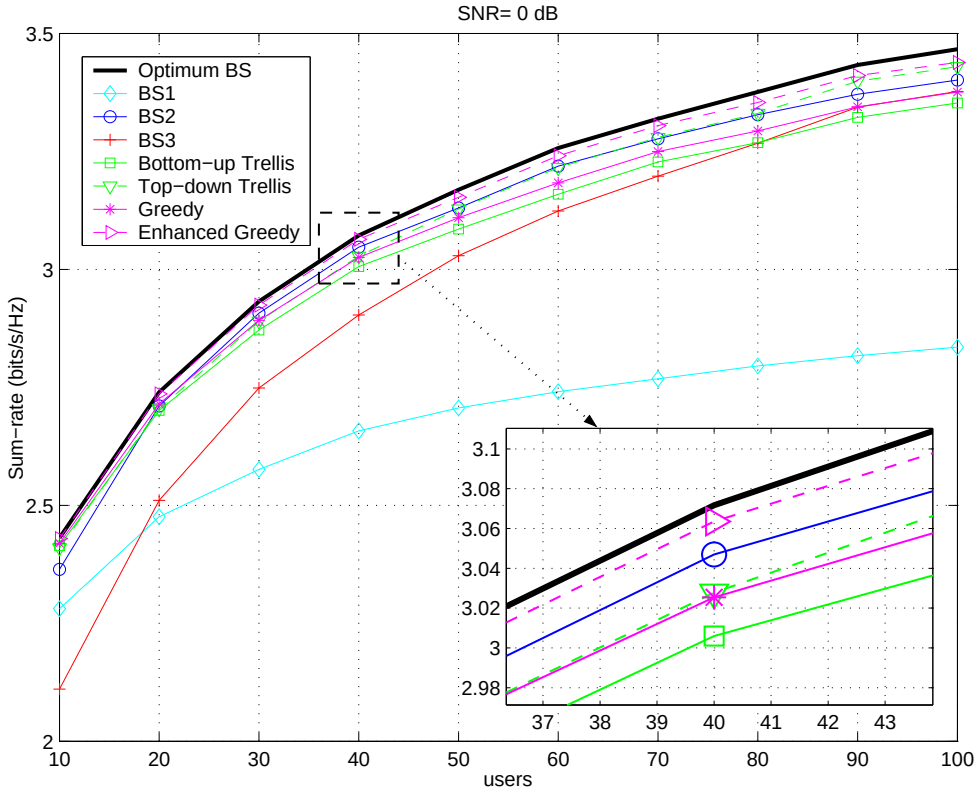


Figure 5.11: Sum-rate vs. users for the different beam selection procedures (SNR=0 dB).

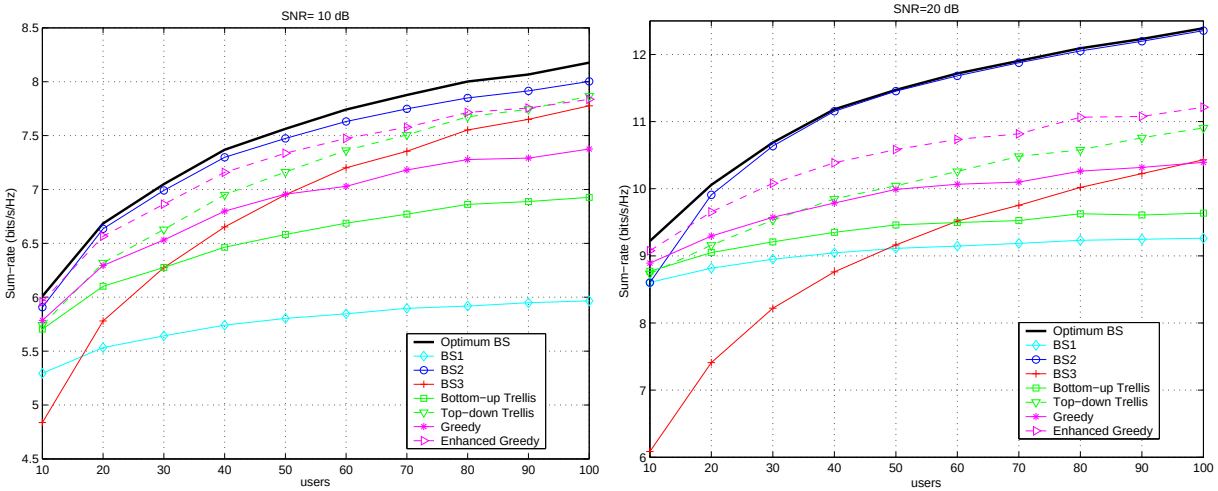


Figure 5.12: Sum-rate vs. users for the different beam selection procedures. Left, SNR=10 dB. Right, SNR=20 dB.

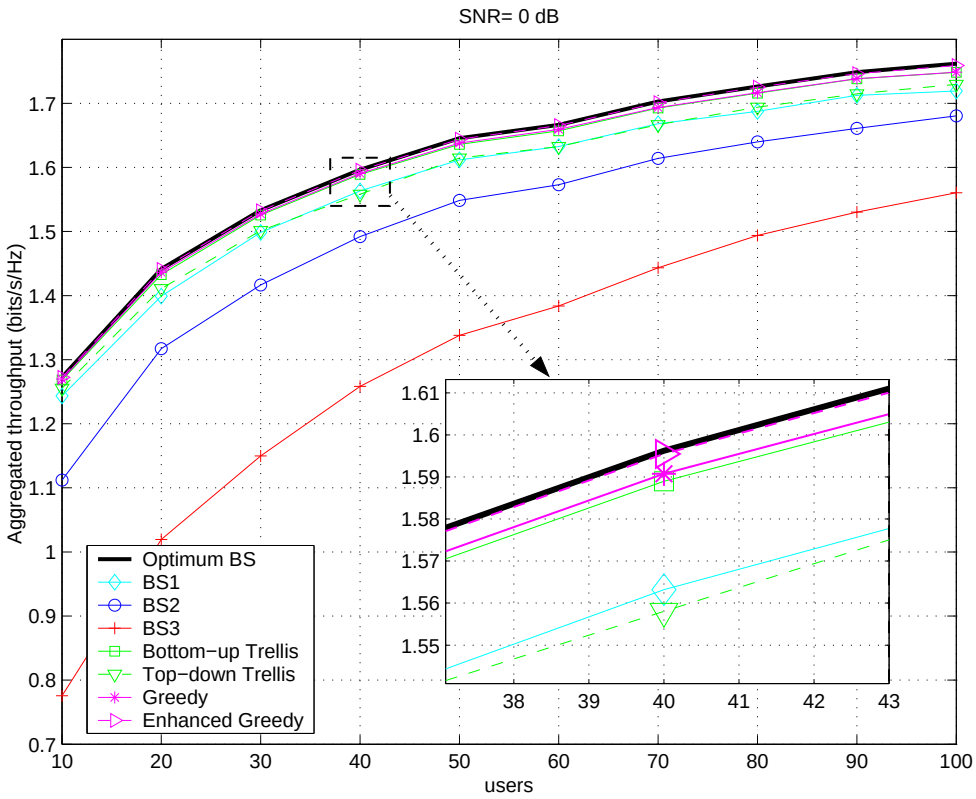


Figure 5.13: Aggregated throughput vs. users for the different beam selection procedures (SNR=0 dB).

enhanced greedy and BS1 approaches.

For the throughput case, however, a different behavior can be observed. As was discussed in Section 5.5, using a single active beam is the most appropriate option for low-SNR. Then, as confirmed in Figure 5.13, more benefits are obtained with BS1 and those recursive techniques based on starting with one active beam (bottom-up, greedy and enhanced greedy). However, throughput quickly saturates when the SNR grows and adding more active beams pays off as observed in Fig. 5.14. If the SNR is not very high and the number of users is significant, better gains are obtained with the top-down and BS3 procedures (see the high-users region of the left-hand plot in Fig. 5.14). Otherwise, BS2 seems to be a reasonable strategy due to its simplicity and good performance.

In order to provide the reader with a complementary point of view, computational complexity requirements in terms of SINR computations are shown in Table 5.3. Considerably computational savings are obtained with the sub-optimum approaches except with the enhanced greedy approach. This algorithm executes M times the greedy approach, where in each step of algorithm the set of users and beams in the search are reduced but all the elements of the set must be tested. For that reason, it is observed that the enhanced greedy approach becomes computationally efficient, with respect to the optimal approach, when the number of transmit antennas

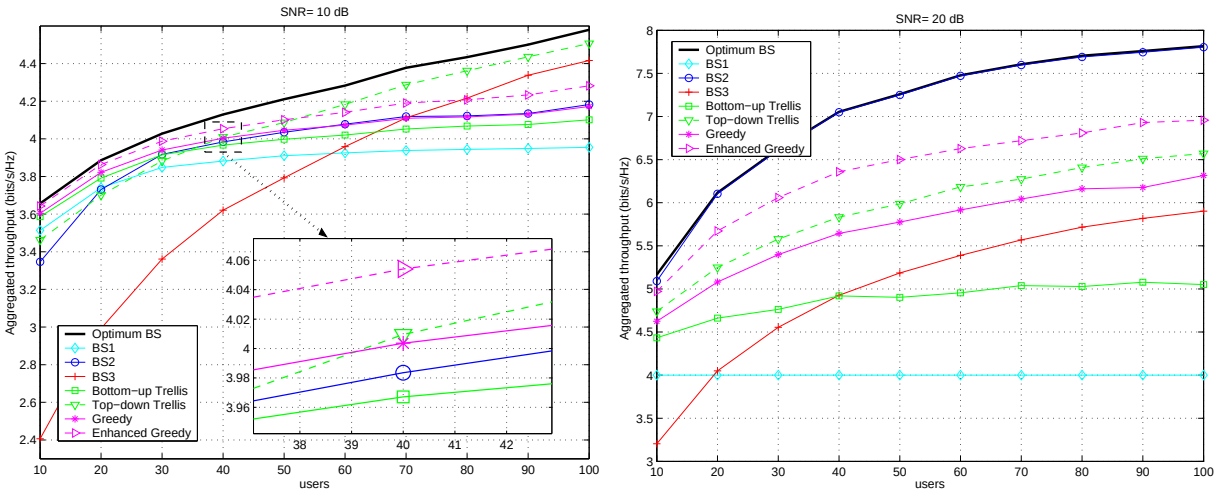


Figure 5.14: Aggregated throughput vs. users for the different beam selection procedures. Left, SNR=10 dB. Right, SNR=20 dB.

Table 5.3: Computational Complexity for the different approaches in terms of SINR computations (K=20 users).

	M=1	M=2	M=3	M=4	M=5
Optimum BS	20	80	240	640	1600
Bottom-up/Top-down	20	42	69	108	175
Greedy	20	60	120	200	300
Improved Greedy	20	80	240	560	1100
BS-1	20	40	60	80	100
BS-2	-	40	120	240	400
BS-3	-	-	60	240	600
BS-4	-	-	-	80	400
BS-5	-	-	-	-	100

is increased.

Finally, we focus on both the BS1 and BS2 procedures with the aim of analyzing their performance as a function of the total number of feedback bits L_t. These schemes provide remarkable benefits in terms of performance vs. complexity trade-offs and allow the derivation of sub-optimum approaches according to the restrictions in the feedback channel as shown in the previous subsection. In the first place, we analyze BS1 in the low-SNR region given its good response to noise-limited scenarios. In this case, only the quantized version of the highest SINR and the beam index must be sent to the BS. By recalling that the number of required bits amounts to $L_t = L_q + \lceil \log_2 N_{tx_{conf}} \rceil$, we plot in the left-hand side of Fig. 5.15 system performance for BS1 as a function of L_t. As long as the number of quantization bits are increased, more benefits are obtained. However, if some of those bits are used for incorporating beam selection

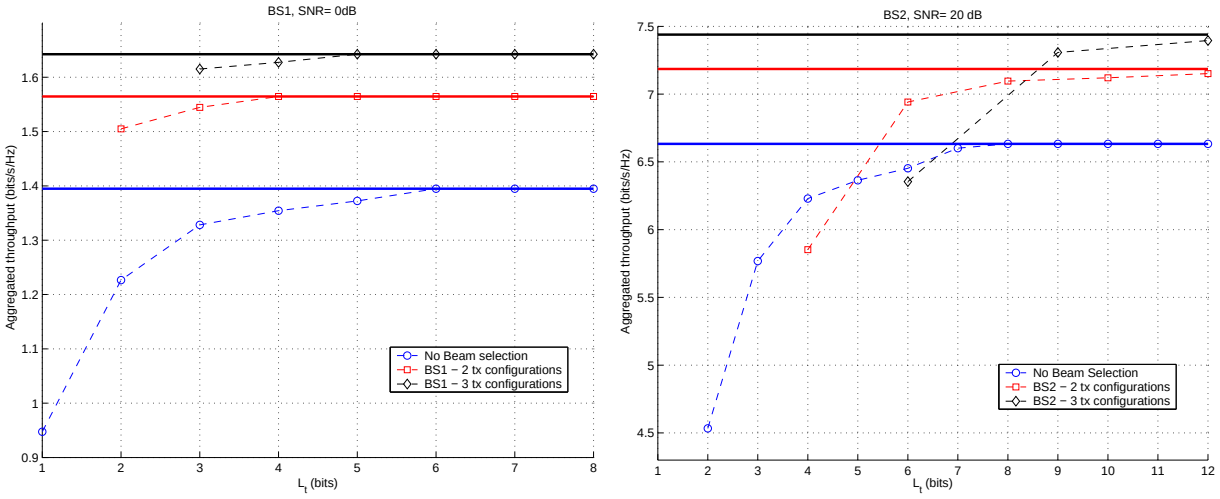


Figure 5.15: Aggregated throughput vs. number of feedback bits, L_t, for different restrictions in the BSX procedure (Solid lines: unquantized throughput, K=60 users). Left, BS1 in a scenario with SNR=0 dB. Right, BS2 in a scenario with SNR=20 dB.

capabilities, considerably gains can be achieved. As a case in point, by using $L_t = 2$ bits we can gain a 29.46% in system performance if the two bits are employed for quantization, whereas the gain is increased to the 58.34% if one of the bits are used for beam selection (BS1-2 tx). For the case that $L_t = 3$ bits are considered, BS1-3 tx can be exploited and the gain is increased to the 69.96%. As for the mid- and high-SNR regions, it was shown that BS2 is the best option for scenarios with a practical number of users. For that reason we consider this strategy in the right-hand side of Fig. 5.15, where results corresponding to a SNR equal to 20dB are shown. Similar conclusions to the BS1 case can be drawn but, as commented previously, a higher number of bits is required for exploiting beam selection capabilities.

5.8 Chapter Summary and Conclusions

In this chapter, a multi-user scenario where several users are scheduled simultaneously by means of an orthogonal random beamforming strategy has been considered. In order to improve system performance in scenarios with a low number of users, the idea of antenna selection presented in the previous chapters have been adapted to the ORB case, this resulting in a beam selection approach.

First, the behavior of both the sum-rate and aggregated throughput expressions have been analyzed as a function of the number of active beams. In terms of sum-rate performance, it has been shown that in the high-SNR region, the system becomes interference-limited and using only a single active beam is the most appropriate strategy. In noise-limited scenarios, however, a higher number of active beams turns out to be beneficial, specially when the population of

users is high. Conversely, a different behavior is observed for the aggregated throughput case due to granularity and bit cap effects. In the asymptotic SNR regime, performance obtained with a single active beam is saturated if the constellation size of the highest available AMC level is small. As a consequence, better performance is obtained by distributing the transmit power among two active beams as using a higher number of active beams is not efficient. Concerning the low-SNR region, using all the active beams is neither an appropriate strategy. This is because, PER levels are very low to support several active beams. Indeed, it has been proved that multiplexing gains can only be exploited in the mid-SNR region of high-populated cells.

Afterwards, performance of a scenario where information reported in the feedback channel is subject to quantization has been studied. More specifically, it has been shown that ORB can be an appropriate strategy in systems with considerably restrictions in the feedback channel. In particular, most of the MUD gains can be extracted with only 2 quantization bits in several situations.

Finally, motivated by the behavior of the sum-rate and aggregated throughput expressions, beam selection algorithms have been proposed. Among all of the proposed approaches, it has been proved that *restricted* beam selection algorithms give a good balance in terms of system performance and computational complexity. Further, these algorithms allow the derivation of schemes capable of trading-off system performance vs. feedback requirements. In particular, it has been shown that ORB performance can be considerably improved by requiring a few bits in the feedback channel with the help of restricted beam selection.

5.A Appendix: Proof of (5.9)

By plugging (5.8) into (5.7), one can obtain an approximated expression for the high-SNR regime by solving the following integral:

$$\begin{aligned} R_{high} &\approx B(B-1)K \int_{\gamma=0}^{\infty} \log_2(1+\gamma) \frac{1}{(1+\gamma)^B} \left(1 - \frac{1}{(1+\gamma)^{B-1}}\right)^{K-1} d\gamma \\ &= B(B-1)K \log_2 e \sum_{k=0}^{K-1} \binom{K-1}{k} (-1)^k \int_{\gamma=0}^{\infty} \frac{\ln(1+\gamma)}{(1+\gamma)^{k(B-1)+B}} d\gamma \end{aligned}$$

where the second equality follows from the application of the binomial expansion. The above integral can be computed by carrying out an integration by parts, that is:

$$\int_{\gamma=0}^{\infty} \frac{\ln(1+\gamma)}{(1+\gamma)^p} d\gamma = \int_0^{\infty} u dv = \lim_{\gamma \longrightarrow \infty}(uv) - \lim_{\gamma \longrightarrow 0}(uv) - \int_0^{\infty} v du.$$

with:

$$\begin{aligned} u &= \ln(1+\gamma) \quad & dv &= (1+\gamma)^{-p} \\ du &= \tfrac{d\gamma}{1+\gamma} \quad & v &= -\frac{(1+\gamma)^{-(p-1)}}{p-1} \end{aligned}$$

Then, by noting that the first two terms go to zero, one can easily verify the following result:

$$R_{high} \approx B(B-1) \log_2 e \sum_{k=0}^{K-1} \binom{K-1}{k} \frac{(-1)^k}{(k+1)^2 (B-1)^2}$$

which can be rewritten with the help of [Gra65, 0.155.4] as follows:

$$R_{high} \approx \frac{B}{B-1} \log_2 e \sum_{k=1}^{K} \frac{1}{k}$$

5.B Appendix: Proof of (5.13)

In order to derive a closed-from expression of the aggregated throughput, one should solve the following expression:

$$\eta(\mathcal{B}) = B \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m \int_{\gamma=\gamma_{th,m}}^{\gamma_{th,m+1}} (1 - \alpha_m e^{-\beta_m \gamma})^L f_{\mathsf{SINR}^*}(\gamma) d\gamma \tag{5.16}$$

By plugging (5.6) into (5.16) and using the binomial expansion, the following expression results:

$$\begin{aligned} \eta(\mathcal{B}) = B \sum_{m=1}^{\mathrm{card}(\mathcal{M})} b_m K \sum_{l=0}^{L} \binom{L}{l} (-\alpha_m)^l \sum_{k=0}^{K-1} \binom{K-1}{k} (-1)^k \\ \times \int_{\gamma=\gamma_{th,m}}^{\gamma_{th,m+1}} \left(\frac{B}{\rho}(1+\gamma) + B - 1\right) \frac{e^{-\gamma\left(\beta_m l + \frac{B}{\rho}(k+1)\right)}}{(1+\gamma)^{k(B-1)+B}} d\gamma \end{aligned}$$

The integral in the above equation can be re-written as:

$$\int_{\gamma=\gamma_{th,m}}^{\gamma_{th,m+1}} \left(\frac{B}{\rho}(1+\gamma)+B-1\right) \frac{e^{-\gamma\left(\beta_m l+\frac{B}{\rho}(k+1)\right)}}{(1+\gamma)^{k(B-1)+B}} d\gamma$$
$$= \frac{B}{\rho}\int_{\gamma=\gamma_{th,m}}^{\gamma_{th,m+1}} \frac{e^{-\gamma\left(\beta_m l+\frac{B}{\rho}(k+1)\right)}}{(1+\gamma)^{k(B-1)+B-1}} d\gamma + (B-1)\int_{\gamma=\gamma_{th,m}}^{\gamma_{th,m+1}} \frac{e^{-\gamma\left(\beta_m l+\frac{B}{\rho}(k+1)\right)}}{(1+\gamma)^{k(B-1)+B}} d\gamma \quad (5.17)$$

Since the two integrals in Eq. (5.17) are of the type:

$$\int_{t=u}^{v} \frac{e^{-at}}{(1+t)^n} dt = \int_{t=u}^{\infty} \frac{e^{-at}}{(1+t)^n} dt - \int_{t=v}^{\infty} \frac{e^{-at}}{(1+t)^n} dt$$
$$a>0; m=1,2,... \quad (5.18)$$

the problem is reduced to solve the following integral:

$$\int_{t=u}^{\infty} \frac{e^{-at}}{(1+t)^n} dt$$
$$a>0; m=1,2,...$$

By using the change of variables $x=(1+t)a$, the integral in the above equation can be re-written as:

$$\int_{t=u}^{\infty} \frac{e^{-at}}{(1+t)^n} dt = e^a a^{n-1} \int_{x=(1+u)a}^{\infty} e^{-x} x^{-n} dx$$

which can be easily solved by resorting to the identity [Gra65, Eq. 8.350.2]:

$$\int_{t=u}^{\infty} \frac{e^{-at}}{(1+t)^n} dt = e^a a^{n-1} \Gamma(1-n,(1+u)a) \quad (5.19)$$

where $\Gamma(\alpha,x)$ stands for the complementary incomplete gamma function ($\Gamma(\alpha,x) \triangleq \int_x^{\infty} e^{-t} t^{\alpha-1} dt$). Finally, by using Eq. (5.19) in Eq. (5.18) one can obtain the integrals in Eq. (5.17) and verify that Eq. (5.13) holds.

Chapter 6

Conclusions and Future Work

In this PhD dissertation we have explored and justified the use of antenna selection mechanisms in single- and multi-user wireless scenarios from a cross-layer perspective. First, we have considered a single-user MIMO wireless system scenario conducting an H-ARQ strategy at the data link layer. In this scenario, we have derived a cross-layer antenna selection strategy in the sense that the optimal sub-set of transmit antennas (at the physical layer) is selected with the aim of maximizing the data link layer throughput, showing considerably gains with respect to conventional approaches. After that, the study of antenna selection strategies has been extended to the multi-user case. Apart from assessing system performance behavior in terms of link throughput, the cross-layer philosophy has also been adopted in the design of the scheduling strategies. More precisely, opportunistic scheduling schemes have been taken into consideration in two different contexts. On the one hand, antenna selection has been utilized in order to jointly exploit spatial and multi-user diversity in a scenario where only one user is scheduled for transmission in each time-slot. On the other hand, an orthogonal random beamforming scheme has been adopted to serve several users simultaneously and, in this case, antenna selection principles have been adapted resulting in a beam selection scheme aimed at improving system efficiency in sparse networks.

6.1 Conclusions

After motivating this PhD thesis and giving an overview of the state of the art, in Chapter 3, a single-user MIMO wireless communication system adopting a V-BLAST scheme has been considered. In order to improve system performance with a low-rate feedback channel, a transmit antenna selection methodology has been taken into consideration. According to channel conditions, the *best* sub-set (with a varying number) of transmit antennas is selected in each time-slot. As for the selection procedure, it has been carried out in accordance with a cross-layer criterion in the sense that the selection is devoted to maximize performance at the data link

layer. This cross-layer approach has been shown to exhibit superior performance in comparison with conventional physical layer criteria since all the system characteristics in both the physical and data link layers are taken into consideration in the optimization procedure.

Besides, the proposed antenna selection has been compared to other restricted selection procedures in the literature, where a pre-determined number of antenna elements is selected out of a largest pool of transmit antennas. Although a simpler scheme results, it has been shown that setting the number of antennas in advance may have a negative impact on V-BLAST performance, specially in the low- and mid-SNR regions due to error propagation effects. In those cases, the most appropriate strategy consists in allocating (i.e. splitting) the available transmit power to a *lower* number of antennas in such a way that error propagation among V-BLAST layers is prevented.

Afterwards, antenna selection algorithms have been further enhanced by allowing adaptive modulation schemes to be included in the optimization procedure. With the proposed cross-layer design, the integration of an adaptive modulation scheme with the antenna selection strategy comes in a very natural way, since both concepts are directly taken into account in the throughput expression. Some attention has also been paid to the fact that an exhaustive search is required in the optimization process and, for that reason, a recursive (albeit optimal) approach has been derived.

Finally, the problem of switching between spatial diversity and multiplexing modes has also been addressed from a cross-layer perspective. More specifically, it has been considered a system where the transmitter is equipped with three transmission schemes (Q-OSTBC, D-STTD and V-BLAST) and a low-rate feedback channel is used to report what is the transmission mode maximizing data link layer performance. By doing so, it has been shown that significant gains can be obtained with respect to the case where such MIMO techniques are used separately.

Chapter 4 has been devoted to discuss the cross-layer interaction between spatial (physical layer) and multi-user diversity (data link layer). More specifically, the downlink of a multi-user wireless system has been studied in a context where only one user is scheduled for transmission in each time-slot. An opportunistic scheme has been considered as scheduling strategy and multi-antenna schemes have been adopted to improve link reliability. In such a context, antenna selection mechanisms have been proposed as a means of efficiently exploiting both gains with low additional requirements in the feedback channel. In addition, different impairments have been introduced in the feedback channel and, in order to improve system robustness against them, antenna selection mechanisms have also been encompassed in conjunction with OSTBC approaches.

Performance assessment has been conducted in terms of ergodic capacity and data link layer throughput metrics, showing a different behavior due to granularity and saturation effects

related with the use of a finite set of modulation schemes. In ideal scenarios where the CSI information of all the users is continuously and perfectly received at the scheduler, antenna selection mechanisms have been proved to be better exploited in combination with single antenna transmission schemes. In other words, the *best* policy consists in focusing all the transmit power on the transmit antenna with the highest gain. Conversely, when the information available at the scheduler is restricted (e.g., by means of SMUD) or subject to degradation, the use of OSTBC-based approaches pays off. This is because in those situations, the inherent robustness of OSTBC to fading alleviates the impact of wrong scheduling decisions.

The interaction of spatial and multi-user diversity has also been assessed with the help of mean vs. standard deviation plots inspired by portfolio theory used in economy. By using this representation, one can easily quantify the degree of robustness of the different transmission schemes against different impairments introduced in the feedback channel, and their impact in terms of system performance.

Finally, in Chapter 5, a multi-user system encompassing an ORB strategy in order to serve several users simultaneously has been considered. In this scenario, the problem of improving ORB performance in scenarios with a low number of users has been addressed. To do so, the antenna selection philosophy presented in previous sections has been adapted to the ORB case, this resulting in a beam selection approach. The rationale behind the proposed approach resides in the fact that using all the active beams may not be the optimal solution in some situations. More precisely, it has been shown that in terms of sum-rate performance, the system becomes interference-limited in the high-SNR regime and the best transmission option is to select a single active beam. Conversely, activating a higher number of beams can be more appropriate in the low-SNR region, specially in those scenarios with a high number of users. In terms of link throughput, however, different conclusions have been drawn. Due to saturation effects, using a single active beam may not be the best strategy in the high-SNR regime; whereas in the low-SNR region, PER levels are too degraded to support several active beams.

Sub-optimum beam selection procedures have also been proposed. It has been shown that by pre-defining a number of active beams, good levels of system performance can be obtained with low expenses in computational complexity. Further, by using this algorithm the required information to be conveyed by the users to the base station is limited to sending only the SINRs associated with the different sub-sets of beams, in contrast to other procedures where all the channel gains must be reported to the base station. Then, according to the number of available sub-sets beams, methodologies capable of trading-off system performance vs. amount of feedback information can be derived. In particular, we have proved that most of the gains can be extracted with a few bits in the feedback channel.

6.2 Future Work

The work presented in this PhD dissertation can be extended as follows:

- To take into account more realistic channels, possibly considering correlation between transmit antennas and users' fading statistics.
- To consider systems with coding schemes and more sophisticated ARQ techniques.
- To take into account random packets arrivals and finite buffer sizes at the transmitter. Then, the cross-layer design should be oriented to stabilize the size of the queues as well.
- To introduce multi-antenna capabilities at the receive side of the proposed multi-user scenarios.
- To extend the cross-layer design to upper layers of the OSI protocol stack. The final objective would be the optimization of the end-to-end communication by considering all the layers in the protocol stack.

As for the specific problems addressed in each chapter, some interesting topics to be addressed are the following:

- In Chapter 3, a possible extension would be the analytical evaluation of the proposed approach by deriving performance bounds and approximations. Also, impairments in the feedback channel should be considered in order to derive a more robust transmit antenna selection procedure.
- In Chapter 4, we have proved that in scenarios with ideal feedback channels MUD is better exploited with single antenna schemes. However, OSTBC becomes a useful technique in the case that information conveyed in the feedback channel is subject to impairments. Then, a more robust solution could encompass power allocation among transmit antennas according to the quality of feedback information. Also, receivers with multiple antenna elements should be studied in order to asses the impact of receive diversity and/or introduce spatial multiplexing capabilities.
- In Chapter 5, we have shown that when quantifying the SINR, most of the gains of ORB can still be extracted. In such a context, the study of more sophisticated quantization techniques should be addressed. In addition, fairness issues should be considered in the design of beam selection procedures. Finally, the impact of imperfect feedback channels should also be addressed.

Bibliography

[Ala98] S.M. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE J. Sel. Ar. Commun.*, 1998.

[Alo99] M.-S. Alouini, and A.J. Goldsmith, "Capacity of rayleigh fading channels under different adaptive transmission and diversity-combining techniques", *IEEE Trans. on Vehicular Technology*, vol. 48, no. 4, pp. 1165–1181, July 1999.

[Bad04] B. Badic, P. Fuxjäger, and H. Weinrichter, "Performance of a quasi-orthogonal space-time code with antenna selection", *IEE Electronics Letters*, vol. 40, no. 20, pp. 1282–1284, Sep. 2004.

[Bar05] D. Bartolomé, "Fairness analysis of wireless beamforming schedulers", *Ph.D. dissertation, Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) [Online]. Available: http://www.cttc.es/profiles/dbartolome/publications.htm*, Jan. 2005.

[Bay05] A. Bayestesh, and A. K. Khandani, "An efficient method for user selection in MIMO broadcast channels", *In Proc. CISS*, 2005.

[Ben99] M. Bengtsson, and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization", *in Proc.of Allerton Conf. on Comm. Cont and Comp.*, 1999.

[Ben01] M. Bengtsson, and B. Ottersten, "Optimal and Suboptimal Transmit Beamforming", L. C. Godara (ed.), *Handbook of Antennas in Wireless Communications*, CRC Press, 2001.

[Ben03] J. Benesty, Y.A. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system", *IEEE Trans. on Signal Processing*, vol. 51, no. 7, pp. 1722–1730, July 2003.

[Ben05] M. Bengtsson, D. Bartolome, J. L. Vicario, and C. Anton-Haro, "Beamforming and bit-loading strategies for multi-user SDMA with admission control", *in Proc. IEEE PIMRC 2005*, 2005.

[Ber92] D. Bertsekas, and R. Gallager, *Data Networks*, Prentice Hall, New Jersey, 1992.

[Ber03] L.T. Berger, T. E. Kolding, J. Ramiro-Moreno, P. Ameigeiras, L. Schumacher, and P. E. Mogensen, "Interaction of transmit diversity and proportional fair scheduling", *in Proc. IEEE VTC Spring*, 2003.

[Ber04] I. Berenguer, X. Wang, and I.J. Wassell, "Transmit antenna selection in linear receivers: geometrical approach", *IEE Electronics Letters*, vol. 40, no. 5, pp. 292–293, March 2004.

[Big91] E. Biglieri, D. Divsalar, P. McLane, and M. Simon, *Introduction to Trellis-Coded Modulation with Applications*, Macmillan, New York, 1991.

[Bos05] R. Bosisio, G. Primolevo, O. Simeone, and U. Spagnolini, "Fair scheduling and orthogonal linear precoding/decoding in broadcast MIMO systems", *in Proc. IEEE PIMRC*, 2005.

[Bos06] R. Bosisio, J. L. Vicario, C. Anton-Haro, and U. Spagnolini, "Diversity-multiplexing tradeoff in multi-user scenario with selective feedback", *in Proc. IST Mobile & Wireless Communications Summit*, 2006.

[Boy04] S. Boyd, and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[Cai03] G. Caire, and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel", *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[Cai06] G. Caire, S. Shamai, Y. Steinberg, and H. Weingarten, "On information-theoretic aspects of MIMO broadcast channels", H. Boelcskei, D. Gesbert, C. Papadias, A. J. van der Veen (eds.), *Space-Time Wireless Systems: From Array Processing to MIMO Communications*, Cambridge University Press, 2006.

[Car04] G. Carneiro, J. Ruela, and M. Ricardo, "Cross-layer design in 4G wireless terminals", *IEEE Wireless Communications*, vol. 11, no. 2, pp. 7–13, April 2004.

[Cau02] S. Cautreaux, P.F. Driessen, and L.J. Greensteins, "Data throughputs using multiple-input multiple-output (MIMO) techniques in a noise-limited cellular environment", *IEEE Trans. on Wireless Communications*, vol. 1, no. 2, pp. 226–234, April 2002.

[Cha04] C.-B. Chae, M. Katz, C. Suh, and H. Jeong, "Adaptive spatial modulation for MIMO-OFDM", *in Proc. IEEE WCNC*, 2004.

[Che04] Z. Chen, "Assymptotic performance of transmit antenna selection with maximal-ratio combining for generalized selection crtierion", *IEEE Communications Letters*, vol. 8, no. 4, pp. 247–249, Apr. 2004.

[Che06] Chiung-Jang Chen, and Li-Chun Wang, "A unified capacity analysis for wireless systems with multiuser scheduling and antenna diversity in nakagami fading channels", *IEEE Trans. on Communications*, vol. 54, no. 3, pp. 469–478, March 2006.

[Chu01] S.T. Chung, and A.J. Goldsmith, "Degrees of freedom in adaptive modulation: A unified view", *IEEE Trans. on Commun.*, Sept. 2001.

[Chu04] S.T. Chung, A. Lozano, H. Huang, A. Sutivong, and J.M. Cioffi, "Aproaching the MIMO capacity with a low-rate feedback channel in V-BLAST", *EURASIP Journal of Applied Signal Processing (JASP)*, vol. 2004, no. 5, pp. 662–675, May 2004.

[Cos83] M. Costa, "Writing on dirty paper", *IEEE Trans. Information Theory*, vol. 29, pp. 439–441, May 1983.

[Cov91] T.M. Cover, and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.

[Dav81] H.A. David, *Order statistics*, John Wiley & Sons, New York, 1981.

[Dia06] J. Diaz, O. Simeone, and Y. Bar-Ness, "Sum-rate of MIMO broadcast channels with one bit feedback", *in Proc. IEEE International Symposium on Information Theory (ISIT)*, 2006.

[Dri99] P.F. Driessen, and G.J. Foschini, "On the capacity formula for multiple-input multiple-output wireless channel: A geometric interpretation", *IEEE Trans. Commun.*, vol. 47, pp. 173–176, Feb. 1999.

[DS05] M. van Der Schaar, and N. Sai Shankar, "Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms", *IEEE Wireless Communications*, vol. 12, no. 4, pp. 50–58, Aug. 2005.

[DST01] "Double-STTD scheme for HSDPA systems with four transmit antennas: Link level simulation results", *Temporary document 21(01)-0701, 3GPP TSG RAN WG1*, June 2001, release 5.

[Fem04] G. Femenias, "Adaptive SR-ARQ combined with STBC in markov channels", *IEE Electronic Letters*, vol. 40, no. 6, March 2004.

[Flo03] F. Floren, O. Edfors, and B.-E. Molin, "The effect of feedback quantization on the throughput of a multiuser diversity scheme", *in Proc. Globecom*, 2003.

[For04] A. Forenza, A. Pandharipande, H. Kim, and R.W. Heath Jr., "Adaptive transmission scheme selection for MIMO systems", *in Proc. WWRF*, 2004.

[Fos96] G.J. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas", *Bell Labs Tech. J.*, pp. 41–59, 1996.

[Fos98] G.J. Foschini, and M.J. Gans, "On limits of wireless communications in a fading environtment when using multiple antennas", *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[Fos99] C.J. Foschini, G.D. Golden, P.W. Wolniansky, and R.A. Valenzuela, "Simplified processing for wireless communications at high spectral efficiency", *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Jan. 1999.

[Fre01] P. Frenger, S. Parkvall, and E. Dahlman, "Performance comparison of HARQ with chase combining and incremental redundancy for HSDPA", *in Proc. IEEE VTC Spring*, vol. 3, pp. 2287–2291, May 2001.

[Fre02] F. Frederiksen, and T. E. Kolding, "Performance and modeling of WCDMA/HSDPA transmission/H-ARQ schemes", *Proc. IEEE Vehicular Technology Conference Fall, Vancouver (Canada)*, vol. 1, pp. 472–476, Sept. 2002.

[GA04] M. Gharavi-Alkhansari, and A.B.Gershman, "Fast antenna subset selection in MIMO systems", *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 339–347, Feb. 2004.

[Gan02] G. Ganesan, P. Stoica, and E.G. Larsson, "Diagonally weighted orthogonal space-time block codes", *in Proc. of Asilomar Conference on Signals, Systems and Computers*, 2002.

[Ges03] D. Gesbert, and M. Alouini, "Selective multi-user diversity", *in Proc. IEEE Intern. Symposium on Signal Proc. And Info. Techn. (ISSPIT)*, 2003.

[Ges04] D. Gesbert, and M. Alouini, "How much feedback is multi-user diversity really worth?", *in Proc. IEEE ICC*, 2004.

[Gol96] G.H. Golub, and C.F. Van Loan, *Matrix Computations*, Baltimore, MD: Johns Hopkins Univ. Press, 3$^{\underline{\text{rd}}}$ ed., 1996.

[Gol97] A.J. Goldsmith, and S.G. Chua, "Variable-rate variable-power MQAM for fading channels", *IEEE Trans. on Communications*, vol. 45, pp. 1218–1230, Oct. 1997.

[Gol99] G.D. Golden, C.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky, "Detection algorithm and initial laboratory results using v-BLAST space-time communication architecture", *Electronics Letters*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[Gon04] C. Gonzalez, L. Szczecinski, and S. Aissa, "Throughput maximization of ARQ transmission protocol employing adaptive modulation and coding", *in Proc. IEEE/IEE ICT*, 2004.

[Gor00] D. Gore, R. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas in wireless systems", *in Proc. ICASSP*, 2000.

[Gor02a] D. A. Gore, and A. J. Paulraj, "MIMO antenna subset selection with space-time coding", *IEEE Trans. on Signal Proc.*, Oct. 2002.

[Gor02b] D.A. Gore, R.W. Heath Jr., and A.J. Paulraj, "Transmit selection in spatial multiplexing systems", *IEEE Communications Letters*, Nov. 2002.

[Gor02c] A. Gorokhov, "Antenna selection algorithms for MEA transmission systems", *in Proc. ICASSP*, vol. 3, pp. 2857–2860, May 2002.

[Gor03] A. Gorokhov, D.A. Gore, and A.J. Paulraj, "Receive antenna selection for mimo spatial multiplexing: Theory and algorithms", *IEEE Trans. on Signal Processing - Special Issue on MIMO Wireless Communications*, vol. 51, no. 11, pp. 2796–2807, Nov. 2003.

[Gor04] A. Gorokhov, M. Collados, D. Gore, and A. Paulraj, "Transmit/receive mimo antenna subset selection", *in Proc. IEEE ICASSP*, 2004.

[Goz03] R. Gozali, R.M. Buehrer, and B. D. Woerner, "The impact of multiuser diversity on space-time block coding", *IEEE Commun. Letters*, May 2003.

[Gra65] I.S. Gradsthteyn, and I.M. Ryzhik, *Tables of Integrals; Series and Products*, New York: Academic, 1965.

[Hea00] R.W. Heath, and A.J. Paulraj, "Switching between spatial multiplexing and transmit diversity based on constellation distance", *in Proc.of Allerton Conf. on Comm. Cont and Comp.*, 2000.

[Hea01] R.W. Heath Jr., S. Sandhu, and A.J. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers", *IEEE Communications Letters*, Apr. 2001.

[Hea05] R.W. Heath, and A.J. Paulraj, "Switching between diversity and multiplexing in mimo systems", *IEEE Trans. on Communications*, vol. 53, no. 6, pp. 962–968, June 2005.

[Hoc04] B.M. Hochwald, T.L. Marzetta, and V. Tarokh, "Multi-antenna channel-hardening and its implications for rate feedback and scheduling", *IEEE Trans. on Information Theory*, vol. 50, pp. 1893–1909, Sept. 2004.

[Hol00] J.M Holtzman, "CDMA forward link waterfilling power control", *in Proc. IEEE VTC Spring*, 2000.

[Jaf01a] S. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback", *in Proc. IEEE ICC*, 2001.

[Jaf01b] H. Jafarkhani, "A quasi-orthogonal space-time block code", *IEEE Commun. Letters*, Jan. 2001.

[Jak74] W.C. Jakes, *Microwave Mobile Communications*, John Wiley & Sons, New York, 1974.

[Jia04] J. Jiang, R.M. Buehrer, and W.H. Tranter, "Antenna diversity in multiuser data networks", *IEEE Trans. on Communications*, vol. 52, no. 3, pp. 490–497, Mar. 2004.

[Jin02] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of gaussian multiple-access and broadcast channels", *in Proc. IEEE ISIT*, 2002.

[Jin04] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of gaussian multiple-access and broadcast channels", *IEEE Trans. on Information Theory*, vol. 50, no. 5, pp. 768–783, May. 2004.

[Jön02] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming with orthogonal space-time block coding", *IEEE Trans. on Inf. Theory*, vol. 48, no. 3, pp. 611–627, March 2002.

[Kaw05] V. Kawadia, and P.R. Kumar, "A cautionary perspective on cross-layer design", *IEEE Wireless Communications*, vol. 12, no. 1, pp. 3–11, Feb. 2005.

[Kay93] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory, Vol. I*, Prentice Hall, New Jersey, 1993.

[Kie98] M. Kiessling, "Unifying analysis of ergodic mimo capacity in correlated rayleigh fading environments", *European Transactions on Telecommunications*, pp. 17–35, Jan. 1998.

[Kno95] R. Knopp, and P.A. Humblet, "Information capacity and power control in single-cell multiuser communications", *in Proc. IEEE ICC*, 1995.

[Kob04] M. Kobayashi, G. Caire, and D. Gesbert, "Antenna diversity versus multiuser diversity: Quantifying the tradeoffs", *in Proc. IEEE ISITA*, 2004.

[Kou05a] M. Kountouris, and D. Gesbert, "Memory based opportunistic multi-user beamforming", *In Proc. of IEEE International Symposium on Information Theory (ISIT)*, 2005.

[Kou05b] M. Kountouris, and D. Gesbert, "Robust multi-user opportunistic beamforming for sparse networks", *in Proc. of IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2005.

[Lar03] E. G. Larsson, and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, Cambridge, 2003.

[Lar04] E. G. Larsson, "On the combination of spatial diversity and multiuser diversity", *IEEE Commun. Letters*, Aug. 2004.

[Li01] L. Li, and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels-Part I:Ergodic capacity", *IEEE Trans. on Information Theory*, vol. 47, pp. 1083–1102, 2001.

[Lin84] S. Lin, D.J. Costello, and M. J. Miller, "Automatic-repeat-request error-control schemes", *IEEE Commun. Mag.*, vol. 22, no. 12, pp. 5–17, Dec. 1984.

[Liu04] Q. Liu, S. Zhou, and G. Giannakis, "Cross-layer combinig of adaptive modulation and coding with truncated ARQ over wireless links", *IEEE Trans. on Wireless Communications*, vol. 3, no. 5, pp. 1746–1755, Sept. 2004.

[Liu05] Q. Liu, S. Zhou, and G.B. Giannakis, "Queuing with adaptive modulation and coding over wireless links: Cross-layer analysis and design", *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1142–1153, May 2005.

[Lu05] X. Lu, G. Zhu, G. Liu, and L. Li, "A cross-layer design over mimo rayleigh fading channels", *in Proc. IEEE WCNMC*, 2005.

[Ma05] Q. Ma, and C. Tepedelenlioglu, "Comparison of multiuser diversity using STBC and transmit beamforming with outdated feedback", *in Proc. IEEE ICASSP*, 2005.

[Maa04a] A. Maaref, and S. Aissa, "Combined adaptive modulation and truncated ARQ for packet data transmission in MIMO systems", *in Proc. IEEE Globecom*, 2004.

[Maa04b] A. Maaref, and S. Aissa, "A cross-layer design for MIMO rayleigh fading channels", *in Proc. IEEE CCECE*, 2004.

[Maa05] A. Maaref, and S. Aissa, "Cross-layer design for MIMO nakagami fading channels in the presence of gaussian estimation errors", *in Proc. IEEE CCECE*, 2005.

[Mag02] J.R. Magnus, and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley Series in Probability and Statistics, John Wiley & Sons, England, 2002.

[Mar52] H. Markowitz, "Portfolio selection", *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, March 1952.

[Mar91] H. Markowitz, "Foundations of portfolio theory", *The Journal of Finance*, vol. 46, no. 2, pp. 469–477, June 1991.

[Mar99] T.L. Marzetta, and B.M. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayeligh flat fading", *IEEE Trans. Information Theory*, vol. 45, pp. 139–157, Jan. 1999.

[Mec04] C. Mecklenbräuker, and M. Rupp, "Generalized alamouti codes for trading quality of service against data rate in MIMO UMTS", *EURASIP Journal of Applied Signal Processing (JASP)*, vol. 2004, no. 5, pp. 662–675, May 2004.

[Med00] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel", *IEEE Trans. Information Theory*, vol. 46, pp. 933–946, May 2000.

[Mes03] X. Mestre, J.R. Fonollosa, and A. Pages-Zamora, "Capacity of mimo channels: asymptotic evaluation under correlated fading", *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 829–838, June 2003.

[Mil02a] A. Milani, V. Tralli, and M. Zorzi, "Improving protocol performance in BLAST-based wireless systems using channel adaptive antenna selection", *Proc. IEEE VTC Spring*, vol. 1, pp. 409–413, 2002.

[Mil02b] A. Milani, V. Tralli, and M. Zorzi, "On the use of per-antenna rate and power adaptation in V-BLAST systems for protocol performance improvement", *Proc. IEEE VTC Fall*, vol. 4, pp. 2126–2130, 2002.

[Mil06] A. Milani, V. Tralli, and M. Zorzi, "On the use of rate and power adaptation in V-BLAST systems for data protocol performance improvement", *IEEE Transactions on Wireless Communications*, vol. 5, pp. 16–22, 2006.

[Mol01a] A.F. Molisch, M.Z. Win, and J.H. Winters, "Capacity of MIMO systems with antenna selection", *in Proc. IEEE ICC*, 2001.

[Mol01b] A.F. Molisch, M.Z. Win, and J.H. Winters, "Reduced-complexity transmit/receive-diversity systems", *in Proc. IEEE VTC Spring*, Nov. 2001.

[Mol03a] A. F. Molisch, "MIMO systems with antenna selection - an overview", *in Proc. Radio and Wireless Conference, Boston (USA)*, pp. 167–170, Aug. 2003.

[Mol03b] A.F. Molisch, M.Z. Win, and J.H. Winters, "Reduced-complexity transmit/receive-diversity systems", *IEEE Trans. on Signal Processing - Special Issue on MIMO Wireless Communications*, vol. 51, no. 11, pp. 2729–2738, Nov. 2003.

[Mol04] A.F. Molisch, and M.Z. Win, "Mimo systems with antenna selection", *IEEE Microwave Magazine*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[Nak60] M. Nakagami, "The m-distribution- A general formula of intensity distribution of rapid fading", *in Statistical Methods in Radio Wave Propagation. Oxford, U.K.: Permagon*, pp. 3–36, 1960.

[Ong03a] E.N. Onggosanusi, A.G. Dabak, Yan Hui, and Gibong Jeong, "Hybrid ARQ transmission and combining for MIMO systems", *in Proc. IEEE ICC*, vol. 5, pp. 3205–3209, May 2003.

[Ong03b] N. Onggosanusi, A. M. Sayeed, and B. D. V. Veen, "Efficient signaling schemes for wideband space-time wireless channels using channel state information", *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 1–13, Jan. 2003.

[Oza94] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio", *IEEE Trans. on Vehicular Technology*, May. 1994.

[Pal03] 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. on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sept. 2003.

[Pau03] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, Cambridge, 2003.

[Pay04] M. Payaró, X. Mestre, and M. A. Lagunas, "Optimum transmit architecture of a MIMO system under modulus channel knowledge at the transmitter", *in Proc. IEEE Information Theory Workshop*, 2004.

[Pay06] M. Payaró, A. Wiesel, J. Yuan, and M. A. Lagunas, "On the capacity of linear vector gaussian channels with magnitude knowledge and phase uncertainty", *in Proc. IEEE ICASSP*, 2006.

[PI06] A. Pascual-Iserte, Daniel P. Palomar, Ana I. Pérez-Neira, and Miguel A. Lagunas, "A robust maximin approach for MIMO communications with partial channel state information based on convex optimization", *IEEE Trans. on Signal Processing*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[Poo97] H. V. Poor, and S. Verdu, "Probability of error in mmse multiuser detection", *IEEE Trans. Information Theory*, vol. 43, no. 3, pp. 858–871, May 1997.

[Pro95] J.G. Proakis, *Digital Communications*, McGraw Hill, 3$^{\underline{\text{rd}}}$ ed., 1995.

[Pro01] J.G. Proakis, *Digital Communications*, Mc Graw Hill, New York, 2001.

[Rai04] V. T. Raisinghani, and S. Iyer, "Cross-layer design optimizations in wireless protocol stacks", *Computer Communications*, vol. 27, pp. 720–724, May 2004.

[RF98] F. Rashid-Farrokhi, K.J.R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems", *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.

[Rup02] M. Rupp, and C. Mecklenbräuker, "On extended alamouti schemes for space-time coding", *in Proc. Wireless Personal Multimedia Communications (WPMC)*, 2002.

[San00] S. Sandhu, R.U. Nabar, D.A. Gore, and A. Paulraj, "Near-optimal selection of transmit antennas for a mimo channel based on shannon capacity", *in Proc. of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*, 2000.

[San04a] S. Sanayei, and A. Nosratinia, "Antenna selection in mimo systems", *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[San04b] A. Sanei, A. Ghrayeb, Y. Shayan, and T.M. Duman, "Antenna selection for space-time trellis codes in fast fading", *in Proc. IEEE PIMRC*, 2004.

[San05] S. Sanayei, and A. Nosratinia, "Exploiting multiuser diversity with only 1-bit feedback", *in Proc. IEEE WCNC*, 2005.

[Sch04] M. Schubert, and H. Boche, "Solution of the multiuser downlink beamforming problem with individual sinr constraints", *IEEE Trans. on Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[Seb00] P. Sebastian, H. Sampath, and A. Paulraj, "Adaptive modulation for multiple antenna systems", *in Conf. Rec. 34 Asilomar Conf. Signals, Syst. Comput.*, pp. 506–510, 2000.

[Sha03] S. Shakkottai, T. S. Rappapaort, and P.C. Karlsson, "Cross-layer design for wireless networkss", *IEEE Communications Magazine*, vol. 41, no. 10, pp. 74–80, Oct. 2003.

[Sha04] M. Sharif, and B. Hassibi, "Scaling laws of sum rate using time-sharing, DPC, and beamforming for MIMO broadcast channels", *in Proc. IEEE ISIT*, 2004.

[Sha05] M. Sharif, and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information", *IEEE Trans. Information Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[Shi00] D. Shiu, G.J. Foschini, M.J. Gans, and J.M. Kahn, "Fading correlation and its effect on the capacity fo multielement antenna systems", *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Nov. 2000.

[Sri05] V. Srivastava, and M. Motani, "Cross-layer design: a survey and the road ahead", *IEEE Communications Magazine*, vol. 43, no. 12, pp. 112–119, Dec. 2005.

[Sta99] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*, Prentice Hall, Upper Saddle River, NJ, 1999.

[Tar98] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space time codes for high data rate wireless communications", *IEEE Trans. on Information Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[Tar99] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Trans. on Inform. Theory*, July 1999.

[Tel95] E. Telatar, "Capacity of multi-antenna gaussian channels", *AT&T Bell Labs, Internal Tech. Memo*, June 1995.

[Tse97] D. Tse, "Optimal power allocation over parallel gaussian broadcast channels", *in Proc. IEEE ISIT*, 1997.

[Tse03] D. Tse, and P. Viswanath, "On the Capacity of the Multiple Antenna Broadcast Channel", G. J. Foschini, S. Verdú (eds.), *Multiantenna Channels: Capacity, Coding and Signal Processing*, American Mathematical Society, DIMACS, 2003.

[UTR01] "Physical layer aspects of UTRA high speed downlink packet access", *Technical Specification, 3rd Generation Partnership Project*, vol. 25.848, March 2001, release 4.

[Vic03a] J. L. Vicario, and C. Antón-Haro, "Optimización del throughput en sistemas V-BLAST usando configuración de antenas de transmisión adaptativa", *in Proc. Jornadas Telecom I+D*, 2003.

[Vic03b] J. L. Vicario, and C. Antón-Haro, "Throughput optimisation for MIMO systems via cross-layer designs", *in Proc. 9th WWRF9*, 2003.

[Vic04a] J. L. Vicario, and C. Antón-Haro, "Full vs. restricted tranmsit antenna selection schemes for mimo cross-layer designs", *in Proc. 12th WWRF12*, 2004.

[Vic04b] J. L. Vicario, and C. Antón-Haro, "Joint transmit antenna selection and adaptive modulation in cross-layer oriented designs for HSDPA systems", *in Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2004.

[Vic04c] J. L. Vicario, and C. Antón-Haro, "Performance and complexity issues in cross-layer MIMO designs for HSDPA", *in Proc. 8th WWRF8bis*, 2004.

[Vic04d] J. L. Vicario, and C. Antón-Haro, "Transmit antenna selection for rate adaptation in HSDPA systems", *in Proc. WWC*, 2004.

[Vic04e] J. L. Vicario, C. Mecklenbräuker, and C. Antón-Haro, "Reduced-complexity methods for throughput maximization in MIMO channels", *in Proc. IEEE ICC*, June 2004.

[Vic04f] J.L. Vicario, and C. Antón-Haro, "Adaptación de la velocidad de transmisión en sistemas multi-antena HSDPA mediante criterios cross-layer", *in Proc. URSI*, 2004.

[Vic05a] J.L. Vicario, and C. Antón-Haro, "Spatial vs. multi-user diversity trade-offs in SMUD systems", *in Proc. Winterschool on Coding and Information Theory*, 2005.

[Vic05b] J.L. Vicario, and C. Antón-Haro, "Adaptive swithching between spatial diversity and multiplexing: a cross-layer approach", *in Proc. IST summit*, 2005.

[Vic05c] J.L. Vicario, and C. Antón-Haro, "Joint exploitation of spatial and multi-user diversity via space-time block-coding and antenna selection", *in Proc. IEEE ICC*, 2005.

[Vic05d] J.L. Vicario, and C. Antón-Haro, "Robust exploitation of spatial and multi-user diversity in limited-feedback systems", *in Proc. IEEE ICASSP*, 2005.

[Vic06a] J. L. Vicario, R. Bosisio, U. Spagnolini, and C. Anton-Haro, "Adaptive beam selection techniques for opportunistic beamforming", *in Proc. IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2006.

[Vic06b] J. L. Vicario, R. Bosisio, U. Spagnolini, and C. Anton-Haro, "A throughput analysis for opportunistic beamforming with quantized feedback", *in Proc. IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2006.

[Vic06c] J. L. Vicario, M. A. Lagunas, and C. Antón-Haro, "A cross-layer approach to transmit antenna selection for high-speed downlink packet access", *accepted for publication in IEEE Trans. on Wireless Comm.*, 2006.

[Vic06d] J.L. Vicario, and C. Antón-Haro, "Analytical assessment of capacity vs. robustness trade-offs in systems with selective multi-user diversity", *in Proc. IEEE ICASSP*, 2006.

[Vic06e] J.L. Vicario, and C. Antón-Haro, "Analytical assessment of multi-user vs. spatial diversity trade-offs with delayed channel state information", *accepted for publication in IEEE Communications Letters*, 2006.

[Vic06f] J.L. Vicario, and C. Antón-Haro, "Cross-layer interaction between spatial and multi-user diversity in selective feedback systems: Outage capacity analysis", *in Proc. IEEE IWCMC*, 2006.

[Vic06g] J.L. Vicario, and C. Antón-Haro, "A unified approach to the analytical assessment of multi-user diversity with imperfect channel state information", *in Proc. European Wireless Conference*, 2006.

[Vic06h] J.L. Vicario, and C. Antón-Haro, "Spatial vs. multi-user diversity trade-offs for cross-layer scheduling in limited feedback systems", *accepted for publication in Eurasip Signal Processing Journal - Special issue on Signal Processing-assisted Cross-layer designs*, 2006.

[Vis01] E. Visotsky, and U. Madhow, "Space-time transmit precoding with imperfect feedback", *IEEE Trans. on Information Theory*, vol. 47, pp. 2632–2639, Sep. 2001.

[Vis02] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas", *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[Vis03a] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian MIMO broadcast channels", *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[Vis03b] S. Vishwanath, G. Kramer, S. Shamai, S. Jafar, and A. Goldsmith, "Capacity Bounds for Gaussian Vector Broadcast Channels", G. J. Foschini, S. Verdú (eds.), *Multiantenna Channels: Capacity, Coding and Signal Processing*, American Mathematical Society, DIMACS, 2003.

[Vis03c] P. Viswanath, and D. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality", *IEEE Trans. on Information Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[Wei04] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian MIMO broadcast channel", *in Proc. CISS, Princeton, USA*, 2004.

[Win83] J. H. Winters, "Switched diversity with feedback for dpsk mobile radio systems", *IEEE Trans. on Vehicular Technology*, vol. 32, no. 1, pp. 134–150, Feb. 1983.

[Win98] J. H. Winters, "Smart antennas for wireless systems", *IEEE Personal Communications*, vol. 5, no. 1, pp. 23–27, Feb. 1998.

[Win99a] M.Z. Win, G. Chrisikos, and J.H. Winters, "Error probability for M-ary modulation using hybrid selection/maximal-ratio-combining in rayleigh fading", *in Proc. Military Comm. Conf.*, 1999.

[Win99b] M.Z. Win, and J.H. Winters, "Analysis of hybrid selection/maximal-ratio-combining in rayleigh fading", *IEEE Trans. on Communications*, vol. 47, pp. 1773–1776, Dec. 1999.

[Win01] M.Z. Win, and J.H. Winters, "Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio-combining in rayleigh fading", *IEEE Trans. on Communications*, vol. 49, pp. 1926–1934, Nov. 2001.

[Xyl99] G. Xylomenos, and G. C. Polyzos, "Internet protocol performance over networks with wireless links", *IEEE Network*, vol. 13, no. 4, pp. 55–63, July/August 1999.

[Yan94] J. Yang, and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems", *IEEE Trans. on Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[Yan04] L. Yang, M. S-. Alouini, and D. Gesbert, "Further results on selective multi-user diversity", *in Proc. ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, 2004.

[Yoo06a] T. Yoo, and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming", *To appear in IEEE Journal of Selected Areas in Commun., available for download at http://wsl.stanford.edu/ yoots/*, 2006.

[Yoo06b] T. Yoo, and A. J. Goldsmith, "Capacity and optimal power allocation for fading MIMO channels with channel estimation error", *To appear in IEEE Trans. on Information Theory, available for download at http://wsl.stanford.edu/∼yoots/*, 2006.

[Yoo06c] T. Yoo, R. Lavery, A. Goldsmith, and D. Goodman, "Throughput optimization using adaptive techniques", *Submitted to IEEE Communication Letters, available for download at http://wsl.stanford.edu/Andrea_publications.html*, 2006.

[Zag06] A. Zagami, "Beamforming and bit-loading strategies for multi-user MIMO systems", *Master Thesis, Universitat Politecnica de Catalunya and Politecnico di Torino*, June 2006.

[Zen04] X. N. Zeng, and A. Ghrayeb, "Performance bounds for space-time block codes with receive antenna selection", *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 2130–2137, Sept. 2004.

[Zhe03] L. Zheng, and D. Tse, "Diversity and multiplexing: a fundamental trade-off in multiple-antenna channels", *IEEE Trans. on Inform. Theory*, May 2003.

[Zho05] S. Zhou, and G. B. Giannakis, "MIMO Communications with Partial Channel State Information", A. B. Gershman, N. D. Sidiropoulos (eds.), *Space-Time Processing for MIMO Communications (to appear)*, John Wiley & Sons, 2005.

[Zhu03] H. Zhuang, L. Dai, S. Zhou, and Y. Yao, "Low complexity per-antenna rate and power control approach for closed-loop V-BLAST", *IEEE Trans. on Wireless Communications*, vol. 51, no. 11, pp. 1783–1787, Nov. 2003.