PREDICTION OF SINGLE AND MULTI-USER DOWNLINK CHANNEL CAPACITIES
BASED ON MULTIPLE ANTENNA PROPAGATION MEASUREMENTS

BY
ELIANA YEPEZ

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

OTTAWA-CARLETON INSTITUTE FOR ELECTRICAL AND COMPUTER ENGINEERING
FACULTY OF ENGINEERING
DEPARTMENT OF SYSTEMS AND COMPUTER ENGINEERING

Carleton University
Ottawa, On

© Copyright by Eliana Yépez 2007
All Rights Reserved

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-27119-3

Our file Notre référence

ISBN: 978-0-494-27119-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

A la huella ejemplar que mi padre dejó en cada uno de nosotros.

¡Salve, Luis E.!

Abstract

In this thesis we develop methods to analyze the capacities of multiple-antenna channels based on a limited number of radio propagation measurements. The results can be applied to the deployment of high-speed fixed Internet wireless services. We design and build a multiple antenna broadband channel sounder to measure the radio channel. In our measurements, the four antenna base station (*BS*) located at the rooftop of a building, transmits to four non-line-of-sight users where each user, also equipped with four antennas, is located inside a different building. Our objective is to compare the downlink channel capacity of an orthogonal space division multiplexing (*OSDM*) scheme, where the *BS* transmits simultaneously to all users, with that of a time division multiplexing (*TDM*) scheme, where the *BS* transmits to each user in a different time slot. We select these schemes because they exploit the channel differently. At each channel use, the spatial sub-channels in *OSDM* are orthogonal by means of transmit-receive spatial processing but they may be unbalanced due to largely separated users. The spatial sub-channels of each user in *TDM* are balanced but they may not be orthogonal due to nearly co-located antennas of each user. Initially, we analyze the data to predict the capacities of the single-user channels. We find that a significant amount of shadowing on the measured channels, which causes variations in the mean power for each receive sub-area, can severely affect their capacities. The importance of this analysis is the connection we make between channels with shadowing and their capacities. We extend this analysis to multi-user channels and find that they are mainly affected by variations in the mean receive powers for each user due to different transmit-user separations. We show that, compared to *TDM*, the capacity of *OSDM* is more sensitive to these power variations.

Acknowledgements

My most special thanks go to my supervisor, David Falconer, for giving me the great opportunity to work on this research that involved both theory and practice. I am grateful to him for his patience and constant support, which gave me the freedom to explore and develop my own perspectives. His great vision and insight will be an inspiring role model in my future career. I also would like to thank my co-supervisor, Robert Bultitude. I have benefited tremendously from his knowledge and experience during all the stages of my work. His detailed questions and criticism motivated me to look for tools to analyze the data, which resulted in the method we propose in this thesis. My gratitude also extends to Prof. A.K.Md. Ehsanes Saleh from the School of Math. & Stat., for many useful discussions on the subject of multivariate statistics and especially for introducing me to bootstrap methods. This research would not have been possible without the financial support of Prof. Samy Mahmoud who made available funding for building the equipment. Also, I acknowledge the NSERC financial support given to my supervisor. I feel fortunate for the great working experience I had in the Communication Research Centre (CRC) during the design and implementation of the MIMO sounder. I thank the support and technical expertise of the team at CRC, especially to Bob Hahn for the excellent job he did in building the equipment. I also thank the enthusiastic participation of Pamela Quan, Rajpreet Dijon, and Rami Youssef during the channel measurements. My Ph.D. experience was made enjoyable by the friendship of Gama, Reza and Farhad and by the unique support of my sister Nena and my brothers Juan Pablo and Carlos. I thank to the graduate secretary Blazenka Power for her diligent work and her great spirit of cooperation. I am deeply indebted to my husband Kamran for his advice, insight and constant support during all this journey. He was also a real guru in fixing all computer and network related problems that I had with my 'mobile office'. I also wish to give thanks to the moral support of all my family, Tia and Aiza. Last but far from least, I am especially grateful to my mother Auri who inspired me to pursue my graduate work.

Notation and Acronyms

ABBREVIATIONS	
<i>BS</i>	Base station
<i>cdf</i>	Cumulative distribution function
<i>iid</i>	Independent identically distributed
MIMO	Multiple input multiple output
<i>NLOS</i>	Non-line-of-sight
<i>pdf</i>	Probability density function
<i>Prob</i>	Probability
<i>OSDM</i>	Orthogonal spatial division multiplexing
<i>SNR</i>	Signal to noise ratio
<i>SVD</i>	Singular value decomposition
<i>TDM</i>	Time division multiplexing
NOTATIONAL CONVENTIONS	
$\approx$	Approximately equal to
$\sim$	Approximately equal to (in distribution)
$\sim$	Equal to (in distribution)
x, X	Scalars
$\mathbf{x}$	Vector
$\mathbf{X}$	Matrix
$\mathbf{I}_n$	Identity matrix of size $(n \times n)$
$\mathbf{X} \otimes \mathbf{Y}$	Kronecker product between $\mathbf{X}$ and $\mathbf{Y}$
$\mathbf{X}^T$	Transpose of $\mathbf{X}$
$\mathbf{X}^\dagger$	Hermitian transpose of $\mathbf{X}$
$Tr(\mathbf{X})$	Trace of $\mathbf{X}$

$\det(\mathbf{X})$	Determinant of $\mathbf{X}$
$ x $	Absolute value of x
$\ \mathbf{X}\ _F$	Frobenius norm of $\mathbf{X}$
$\ \mathbf{x}\ _2$	Vector norm of $\mathbf{x}$
$\text{vec}(\mathbf{X})$	Forms a column vector by stacking the transpose of each row of $\mathbf{X}$
$\langle \cdot \rangle$	Forms a short-wide matrix by stacking each of the elements in $\cdot$ one beside the other in a horizontal fashion i.e., $\langle \mathbf{A}, \dots, \mathbf{M} \rangle = [\mathbf{A} \ \dots \ \mathbf{M}]$
$E\{\cdot\}$	Expectation
$\log(\cdot)$	Logarithm base 2
$\ln(\cdot)$	Natural logarithm
$\mathbb{C}^{m \times n}$	set of $m \times n$ complex entries
$\Re\{c\}$	Real part of the complex variable c
$\Im\{c\}$	Imaginary part of complex variable c
$\mathcal{LN}(\mu, \sigma)$	Lognormal distribution with mean μ in dB and standard deviation σ in dB
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
$\mathcal{CN}(\mu, \sigma^2)$	Complex proper normal distribution with mean μ and variance σ^2
$\mathcal{N}_m(\mu, \Sigma)$	m-variate normal distribution with mean vector μ and covariance matrix Σ
$\mathcal{CN}_m(\mu, \Sigma)$	Complex m-variate proper normal distribution with mean vector μ and covariance matrix Σ
$[\theta_{lo}, \theta_{up}]$	Estimated 90% bootstrap confidence interval
P_{out}	Outage probability
C_{erg}	Ergodic capacity
$C_{out,x}$	Outage capacity with $P_{out} = x$

Contents

Abstract	iv
Acknowledgements	v
Notation and Acronyms	vi
1 Introduction	1
2 Sounder and impulse response estimates	15
2.1 MIMO channel sounder	16
2.1.1 Transmitter	16
2.1.2 Receiver	18
2.2 Operational characteristics	21
2.3 Link budget	21
2.4 Impulse response estimates	22
2.5 Impulse response estimates after threshold	23
3 Channel measurements	27
3.1 Description of the measurement scenarios	27
3.2 Description of the impulse response and its transfer function	30
3.3 Measured impulse response estimates and their transfer functions	37
4 Estimation of spatial correlations	40
4.1 Estimation of the independent frequency sub-bands	42
4.1.1 Mathematical framework for the independence test	42
4.1.2 Application of the test	43
4.1.3 Test results	45

4.2	Validation of the Kronecker model	48
4.2.1	Mathematical framework for the multiple homogeneity test	51
4.2.2	Application of the test	52
4.2.3	Test results	54
4.3	Estimation of spatial correlations	55
4.3.1	Mathematical framework for the principal components test	55
4.3.2	Application of the test	58
4.3.3	Test results	60
4.4	Discussion	65
5	MIMO single-user channel capacities	68
5.1	Estimated path loss differences within a building	69
5.2	Estimated shadowing losses for each building	71
5.3	The MIMO single-user channel model and its ergodic and outage rates . . .	76
5.3.1	Ergodic rates	79
5.3.2	Outage rates	80
5.4	Estimated rates of the measured single user channels	84
5.4.1	Ergodic rates	87
5.4.2	Outage rates	88
5.5	Discussion	95
6	MIMO downlink channel capacities	97
6.1	Estimated mean receive power for each building	99
6.2	Downlink fixed channels with Gaussian noise	101
6.2.1	<i>OSDM</i> schemes for the downlink	103
6.2.1.1	Step 1: transmit processing	104
6.2.1.2	Step 2: transmit-receive processing	108
6.2.2	<i>TDM</i> scheme for the downlink	109
6.2.3	Example: two user rate region of a downlink fixed channel	109
6.3	Downlink fading channels with Gaussian noise	111
6.3.1	Extension of the <i>OSDM</i> schemes to fading channels	112
6.3.1.1	Ergodic rates	113
6.3.1.2	Outage rates	115
6.3.2	<i>TDM</i> scheme for fading channels	117

6.4	Finding a set of balanced rates for all users	118
6.5	Estimated downlink rates of the measured channels: <i>OSDM</i> vs. <i>TDM</i> . . .	120
6.5.1	Ergodic and outage rates	121
6.5.2	Transmit power, time allocation and <i>SNR</i> for each user	126
6.6	Discussion	129
7	Conclusions	135
A	Background on radio channels	141
B	Complex proper normal distribution	152
C	Empirical distributions of the test statistics	154
D	Multivariate normality validation	158
E	Bootstrap methods	168
F	Estimation of the power ratio	175
G	Independence of transmit vectors	180
H	Principal components test - complex case	182
I	Optimal power allocation algorithms	187
J	Accuracy of the capacity estimates	191
K	Example: path loss and shadowing	198

List of Tables

2.1	MIMO 4×4 system specifications	21
2.2	Link budget	22
3.1	Distance between transmit and receive areas	28
4.1	Estimation of frequency separation between two independent frequencies . .	46
5.1	Maximum distance separation between measured sub-areas in a building . .	70
D.1	Number of rejections of the skewness and kurtosis tests for normality assessment	164
D.2	Number of rejections of the Shapiro-Wilk W test for normality assessment .	167
D.3	Number of rejections of all the tests for normality assessment	167
G.1	Number of rejections of independence test for transmit vectors	181
K.1	TDM vs. sum capacity of the cooperative multi-user channel (4×2) coop .	204
K.2	$OSDM$ vs. sum capacity of the cooperative multi-user channel (4×2) coop	204

List of Figures

1.1	Pictorial of a downlink multi-user channel	2
1.2	Two user rate region for a fixed channel	4
1.3	Two user rate region for a fading channel	6
2.1	MIMO 4×4 transmitter	17
2.2	Transmit antenna array	17
2.3	MIMO 4×4 receiver	18
2.4	Receive antenna array	19
2.5	Acquisition time at the receiver	20
2.6	Probability of detection vs. false alarm	25
2.7	Enlarged view of P_D vs. P_{FA}	25
2.8	Estimate of impulse response with noise	26
2.9	Estimate of impulse response after application of the CFAR algorithm	26
3.1	Measured receive areas	31
3.2	Transmit location: Mackenzie Building	32
3.3	Receive area: Dunton Tower	33
3.4	Receive area: Library	34
3.5	Receive area: Loeb Building	35
3.6	Receive area: Southam Hall	36
3.7	A measured impulse response	38
3.8	A measured transfer function	38
3.9	Back to back measurement	39
4.1	A measured transfer function	46
4.2	Estimated independent frequency sub-bands of a transfer function	47

4.3	Test results for validity of Kronecker model	55
4.4	Finding the values of h and q that minimize $\theta_{h,q}$	61
4.5	Finding the values of h and q that minimize $L\theta_{h,q}$	62
4.6	Estimated 90% confidence interval for $\tilde{C}_q(f)$	64
4.7	Effect of spatial correlations on the <i>cdf</i> of the channel capacity	67
5.1	Estimated path loss within LI	71
5.2	Confidence interval for σ	74
5.3	Point estimates of the ergodic rates of measured channels	86
5.4	Ergodic and outage rates of <i>iid</i> Rayleigh channels	88
5.5	Point estimates of the outage rates of the measured channels	90
5.6	Effect of shadowing on the <i>cdf</i> of the channel capacity	93
5.7	Ergodic and outage rates of uncorrelated channels with shadowing	94
5.8	Rates of uncorrelated channels with different degrees of shadowing	96
6.1	Confidence interval for r_{dB}	100
6.2	Two-user capacity region of a downlink fixed channel with <i>OSDM</i> and <i>TDM</i>	110
6.3	Downlink fading channel	111
6.4	Point sum rate estimates of <i>OSDM</i>	122
6.5	Point balanced rate estimates of <i>OSDM</i> and <i>TDM</i>	124
6.6	Point estimates of transmit power, time allocation and <i>SNR</i>	128
6.7	Ergodic and outage sum rates of cooperative multi-user channels	131
6.8	<i>cdfs</i> of the sum capacities of cooperative multi-user channels	132
C.1	Empirical distribution of $\tilde{\lambda} = -2 \ln \lambda$ with $q = 2, p = 1, N = 17$	155
C.2	Empirical distribution of $\tilde{\lambda} = -2 \ln \lambda$ with $q = 4, p_i = 4, N = 17$	155
C.3	Empirical distribution of $-2 \ln \lambda_5$	156
C.4	Empirical distributions of the sample skewness and kurtosis	157
C.5	Empirical distributions of W	157
D.1	Skewness test results	165
D.2	Kurtosis test results	165
D.3	Shapiro-Wilk test results	166
E.1	Histogram of the bootstrap sample mean	171

F.1	Estimated 90% confidence interval for the power ratio	179
G.1	Test results for independent transmit vectors	181
J.1	Confidence interval for C_{erg} and point estimate ($F = 11$)	192
J.2	Confidence interval for $C_{out,0,06}$ and point estimate ($F = 11$)	194
J.3	Confidence interval for C_{erg} and point estimate ($F = 1000$)	196
J.4	Confidence interval for $C_{out,0,06}$ and point estimate ($F = 1000$)	197
K.1	Path loss, shadowing and the <i>cdf</i> of the channel capacity of <i>TDM</i>	201
K.2	Time allocation per user vs. overall <i>SNR</i> of <i>TDM</i>	206
K.3	<i>SNR</i> per user vs. overall <i>SNR</i> of <i>TDM</i>	207
K.4	Path loss, shadowing and the <i>cdf</i> of the channel capacity of <i>OSDM</i>	208
K.5	Transmit power allocation per user vs. overall <i>SNR</i> of <i>OSDM</i>	209
K.6	<i>SNR</i> per user vs. overall <i>SNR</i> of <i>OSDM</i>	210

Chapter 1

Introduction

Recently there has been great interest in using high speed wireless systems to deliver Internet access from the backbone network (which may be connected with optic fiber) to the last mile end users, as an alternative to cable and digital subscriber line (*DSL*) connections. The Internet access alternative using wireless channels is attractive since it is easy to deploy and the installation cost can be relatively low [Eklund *et al.* (2002)]. On the other hand wireless channels are affected by shadowing due to the obstruction of different interfering objects as well as path loss¹ which degrades their performance in terms of capacities² when compared to wired channels.

Also due to the growing demand in the number of users who want to obtain wireless Internet access as well as the number of Internet applications that require high data rates, it is important that wireless channels offer high capacities. We could increase the capacity of wireless channels by increasing the transmit power (for which channel capacity only grows

¹In this work path loss refers to the difference between the mean transmit and receive powers that is due mainly to the distance separation between the transmitter and the receiver. i.e., it does not include shadowing.

²We will use the terms 'capacities', 'achievable rates' and 'rates' interchangeably. We will explicitly mention whether or not there is channel knowledge at the transmitter when necessary.

Figure 1.1: Downlink multi-user channel: base station, located at the rooftop of a building and equipped with four antennas, transmits to four non-line-of-sight users located inside different buildings where each user terminal is also equipped with four antennas.

logarithmically) or by transmitting over a wider frequency spectrum, however these resources are limited by regulation bodies.

An efficient solution to increase the capacity of single-user wireless channels in rich scattering environments, that does not require an increase in transmit power or frequency spectrum, was proposed by [Winters (1987), Telatar (1995), Foschini (1996)]. Assuming channel knowledge at the receiver, the authors found that by increasing the number of antenna elements at the transmitter and at the receiver, the channel capacity will grow linearly (at high signal-to-noise-ratio (SNR)) with the minimum value between the number of transmit antennas and the number of receive antennas.

For the case of multiple-antenna multi-user wireless channels (see an example in Figure 1.1), assuming they are *iid* Rayleigh distributed, it has also been observed in [Vishwanath

(2003)] that there is a linear growth in the sum capacity of the uplink channel³ [Telatar (1995)] with channel knowledge at the receiver and sum capacity of the downlink channel⁴ [Yu and Cioffi (2001), Viswanath and Tse (2003), Vishwanath *et al.* (2003)] with channel knowledge at the transmitter and the receiver. For these channels sum capacity grows linearly with the minimum value between the number of users K and the number of antennas at the base station N_T . The authors in [Vishwanath (2003)] found the proportionality constant of sum capacity growth for the case when $K < N_T$.

We note that in multi-user channels a significant fraction of the sum capacity can be achieved as long as there are as many users as number of antennas at the base station, and each user occupies one sub-channel of the system. This behavior has been observed in various studies. In [Mazzarese and Krzymień (2003)] the authors find that for the case of two-user fading channels there is only a marginal loss in terms of their ergodic and outage capacities⁵ if the number of users is limited to the number of transmit antennas as opposed to a larger number of users. In [Rhee *et al.* (2004)] the authors find that for the case of uplink MIMO (multiple-input-multiple-output) channels, beamforming at the transmitter (one sub-channel per user) is optimal as long as the number of users is large enough to occupy all the sub-channels of the system which are limited by the number of receive antennas in this case. In [Blum *et al.* (2001)] the authors consider the capacity of MIMO cellular systems in which the performance of the system is often improved by limiting the number of sub-channels

³The uplink channel (from user terminals to base station) sum capacity is the maximum total throughput rate (sum of all user rates) that can be achieved reliably given per user transmit power constraints P_i $i \in [1, \dots, K]$.

⁴The downlink channel (from base station to user terminals) sum capacity is the maximum total throughput rate that can be achieved reliably given a sum transmit power constraint $\sum_{i=1}^K P_i \leq P$ where P_i is the power of the i^{th} user.

⁵Considering that the random channel realizations occur in time then we give the following definitions. The ergodic capacity assumes that there are no decoding delay constraints and it corresponds to the maximum rate that can be achieved reliably by averaging the capacity over all channel realizations. The outage capacity, with an associated outage probability $P_{out} \leq \epsilon$, assumes that there are decoding delay constraints and it corresponds to the maximum rate that can be achieved on the channel with a probability of outage $\leq \epsilon$

Figure 1.2: Fixed channel: downlink two-user rate region (R_1, R_2) achieved by different user power allocations $\{P_1, P_2\}$ such that the total transmitted power is $P_1 + P_2 \leq P$

per user, since transmitting through all the sub-channels causes co-channel interference for other users degrading the overall performance. Finally, in [Yu and Rhee (2006)] the authors find that the optimal number of users for the maximum sum capacity of MIMO downlink fading channels whose elements are *iid* Rayleigh, is given by N_T^2 , the square of the number of transmit antennas, however as pointed out in the earlier studies there is only a marginal loss if the number of users is limited by the number of transmit antennas.

Figure 1.2 shows an example of a downlink two-user rate region (i.e., all user rate tuples (R_1, R_2) that can be achieved reliably) of a non-fading channel. Each point in the two-user rate region boundary can be achieved by different user power allocations i.e., $\{P_1, P_2\}$, subject to a constraint in the total transmitted power $P_1 + P_2 \leq P$. In general, different optimization criteria will lead to different points on the boundary of the region.

There are two observations that we want to point out regarding the performance of multi-user channels. We observe that the sum capacity optimization criterion is adequate only when the statistics of all user channels are similar so that we deliver balanced rates to all users. Otherwise this optimization criterion will deliver high rates to users with relatively high mean receive powers at the expense of delivering negligible rates to users with relatively low mean receive powers (see the sum rate point in the two-user rate region of Figure 1.2). In practice, due to different separation distances between users and the transmitter, different user surrounding environments, etc. the user channel statistics may be considerably different. As a result, in order to deliver fair rates to all users, a more relevant measure of performance is the weighted sum rate (see the weighted sum rate point in the two-user rate region of Figure 1.2 where μ_1, μ_2 are the user rate rewards introduced by [Tse and Hanly (1998)]).

For the case of a fading channel i.e., when the channel varies at each realization, the performance of a multi-user system is measured in terms of its ergodic and its outage rate regions. To illustrate these concepts we show in Figure 1.3 the rate regions based on a two-user down-link fading channel example. Figure 1.3 (a) shows how the two-user rate region varies at each channel realization. This figure depicts the instantaneous two-user rate region at each channel realization $n \in [1, \dots, N]$ i.e., the instantaneous user rate tuples $(R_1(n), R_2(n))$, that are achieved with different user power allocations $\{P_1(n), P_2(n)\}$ subject to a constraint on the average transmit power $P = P_1 + P_2$, where $P_i = E\{P_i(n)\}$, $i \in [1, 2]$.

Figure 1.3 (b) shows the long-term rate regions under no decoding delay constraints (ergodic rate region) as well as under decoding delay constraints (outage rate region). More specifically this figure depicts the ergodic two-user rate region i.e., user rate tuples (R_1, R_2) , that can be obtained reliably by averaging over all channel realizations. It also shows the outage rate region, with an associated user outage probability $P_{out,1} = P_{out,2} \leq \epsilon$ i.e., user rate

Figure 1.3: Fading channel: downlink two-user rate region achieved by different average user power allocations $\{P_1, P_2\}$ such that the average transmitted power is $P_1 + P_2 \leq P$

tuples (R_1, R_2) , that can be obtained reliably such that the probability of outage for user one is $\leq \epsilon$ and such that the probability of outage for user two is $\leq \epsilon$.

An algorithm that finds the maximum weighted sum rate of the downlink fading channels under a long-term input power constraint is given in [Mohseni *et al.* (2006)]. However the implementation of this algorithm is rather involved when compared to the implementation of other sub-optimal schemes such as orthogonal space division multiplexing schemes.

From these observations one can see that, when studying the performance of multi-user channels, it is important to use an optimization criterion that guarantees fair (balanced) rates to all users and it is also important to consider the performance of simple multi-user schemes that although sub-optimal, are easy to implement.

Note that one of the conditions for the linear growth in channel capacities described above is

that the spatial variations in the complex gains on the links between different transmit and receive antenna elements, which here we will refer as the elements of the MIMO channel, are uncorrelated⁶. Under realistic propagation environments this condition may especially affect the capacity of single-user MIMO channels whose elements may be correlated e.g., when the separation between the antenna elements is not sufficient, decreasing their channel capacities [Chuah *et al.* (2002)].

Also, since real world measured MIMO wireless channels may be severely affected by physical link channel factors such as shadowing and path loss, the *iid* Rayleigh channel distribution assumption may no longer be valid. This should be especially considered for multi-user channels with largely spaced users.

Even though the linear growth in channel capacity mentioned above occurs regardless of the distribution of the elements of the MIMO channel [Chuah *et al.* (2002)], it is important to know whether or not these factors can significantly affect the performance of the measured channels when compared to that of the theoretical *iid* Rayleigh channels, since this will give us practical limits on the rates that can be achieved on real world measured channels.

Objective

In this thesis we are interested to study the achievable rates of real world channels for fixed broadband wireless Internet access such as the ones specified by IEEE 802.16a [IEEE P802.16a

⁶In our work we exclude the keyhole channels presented in [Chizhik *et al.* (2002)], since this type of channels do not typically occur in real world propagation scenarios [Almers *et al.* (2003), Bonek *et al.* (2003)]. In fact we confirm in Appendix(D) that the elements of the measured channels are normal or log-normal distributed, which eliminates the possibility of key-hole measured channels whose elements are double Rayleigh distributed [Gesbert *et al.* (2002)]

(2002)]. We focus our study on downlink channels where the transmitter is in non-line-of-sight (*NLOS*) with respect to the receiver, the transmitter is located outdoors and the receiver is located indoors (see Figure 1.1). These types of channels are of interest to wireless Internet providers who want to deploy their services in places where the transmitter is located on rooftops of relatively low-rise buildings.

Our target is to compare the relative performance of measured multi-user channels assuming the following sub-optimal schemes for transmission on MIMO downlink channels:

- (1) Orthogonal Space Division Multiplexing (*OSDM*): the base station (*BS*) uses all the system bandwidth to simultaneously transmit to all users located in different receive areas. This is achieved by projecting the user signals onto orthogonal sub-channels so that users are separated according to their spatial signatures.
- (2) Time Division Multiplexing (*TDM*): the base station (*BS*) sends information to each user in a different time slot so that users are separated in time but information to a given user flows through several orthogonal spatial sub-channels.

Approach

To analyze these types of channels, we design and build a 4×4 MIMO broadband channel sounder. We use this equipment to measure *NLOS* outdoor to indoor channels on the Carleton University campus.

We assume that the measured channels are fixed in time and variable in space. The random variations of the channel occur when we place the receive unit in different locations of a given receive area. We take measurements in four different receive areas (buildings). In our analysis, we assume that a user moves through the different measured locations of a receive area. We perform our analysis in two stages:

- (1) Single-user MIMO channels: (performance of one user in one measured receive area) the objective is to estimate the single-user channel performance for each measured receive area and to compare it with the performance of the theoretical MIMO single-user *iid* Rayleigh channels. The main importance of this analysis is the connection that we are able to establish between the physical link factors of the measured single-user channels and the effect on their capacities. We use the insight gained through this analysis in the next stage, where we extend this connection to explain the effect of the physical link factors of the multi-user channels on their capacities.
- (2) Multi-user MIMO channels: (performance of four users, where each user is in one of the four measured receive areas) the objective is to compare the relative multi-user channel performances of the *OSDM* schemes, given in [Spencer *et al.* (2004)], with that of the *TDM* scheme. We assume that in the *OSDM* schemes the transmitter and the receiver know the channel while in the *TDM* scheme the receiver knows the channel and the transmitter has a limited channel knowledge so that it can adjust the average allocated transmission time for each user.

For the case of MIMO single-user channels, we evaluate the ergodic and outage rates using the closed form channel capacity expressions given in [Telatar (1995), Foschini (1996)] assuming that only the receiver knows the channel, where we use equal power allocation as a reasonable strategy. When both the receiver and the transmitter know the channel we allocate the power using the optimal water-filling solution described in [Telatar (1995)] for the case of ergodic capacity and the optimal power allocation solution described in [Biglieri *et al.* (2001)] for the case of outage capacity. We identify and quantify the physical link channel factors that affect the performance of the single-user MIMO channels.

For the case of MIMO multi-user channels, we evaluate the ergodic and outage rates by extending the algorithms given in [Spencer *et al.* (2004)] for fixed frequency-flat channels

to the case of fading frequency-selective channels, where we use the same optimal solutions of the MIMO single-user channels. This procedure will lead to a region of achievable user rate tuples. Since our objective is to find the user rate vector that has the most balanced rates for all users, we propose an optimization criterion to find this user rate vector. We also identify and quantify the physical link channel factors that affect the channel performance of each of the multi-user schemes: *OSDM* and *TDM*.

Contributions

The following are the contributions of our work. For each of the contributions first we describe related work and then we explain the differences with our work.

- (1) *Use of bootstrap methods for estimation of channel parameters from a small number of samples:* in 1979 the author in [Efron (1979a), Efron (1979b)] first introduces bootstrap methods to estimate the standard error of virtually any estimator. Since then important developments have been achieved in this field and these methods have turned into very powerful statistical tools that are especially useful to estimate the accuracy of a parameter of interest when the number of samples is small. Bootstrap methods are widely used in areas such as radar signal processing, sonar signal processing, biomedical engineering, seismology, etc. [Zoubir and Boashash (1998)].

In this work we extend the use of bootstrap methods to analyze different parameters of measured radio channels based on a limited number of measurements. Some examples where we use these methods are: to approximate the confidence interval for the test statistic of principal components so that we can evaluate the related hypothesis; to estimate the confidence intervals for physical channel parameters such as: shadowing and path loss; to establish a comparison in terms of ergodic and outage rates of the

measured single-user channels and the theoretical *iid* Rayleigh channels.

- (2) *Estimation of spatial correlations in the MIMO channel elements:* The authors in [Ivrlač and Nossek (2003), Ivrlač and Nossek (2005)] proposed a measure to find the amount of correlation in the elements of the MIMO channel. This correlation measure is directly related to what they call a diversity measure, where the latter is based on the definition of the coefficient of variation of a random variable [Efron (1993)].

In our work we estimate the spatial correlations of the elements of the MIMO channel by analyzing their effect on the conditioning of the channel covariance matrix. By using some of the tests of multivariate statistics, we are able to perform this analysis, which is described in the following steps:

- (a) Validation of multivariate normal distribution of the measured data.
- (b) Validation of Kronecker model to represent spatial correlations in the elements of the measured MIMO channel.
- (c) Estimate of the number of most significant eigenvalues of the channel covariance matrix.

The estimate in the last step is similar to the estimate of the diversity measure of the previous study. The difference in our work is that we estimate the diversity measure using one of the likelihood ratio tests of principal components, while the authors in [Ivrlač and Nossek (2003)] propose their own estimator. We think that the main advantage of using our approach is that the statistical properties of the test statistic we use are well established. We know for example its asymptotic distribution which we use to build a confidence interval and test the hypothesis that a given number of eigenvalues of the covariance matrix is non significant, by checking whether the null value lies inside this interval. We also know that this test statistic is limited for use

with multivariate normal distributed data as well as for the cases where the eigenvalues of the covariance matrix have multiplicity one. In order to evaluate the performance of these two estimators a further study is needed to compare their statistical properties such as bias and variance; however this is out of the scope of this work.

- (3) *Effects of shadowing and path loss on the cdf of channel capacity*: Based on indoor channel measurements performed in different locations of an office environment the authors in [Bonek *et al.* (2003), Özcelik *et al.* (2003b), Herdin *et al.* (2002)] found that the ergodic rates of single-user channels under equal power allocation are mostly determined by different mean received powers experienced for the different measured receive locations and not by the spatial correlations in the elements of the measured MIMO channels⁷.

In our work, after finding that the elements of the MIMO measured channels can be considered uncorrelated (from the analysis results of the previous contribution), we also found that the rates of our measured channels are mainly affected by differences in the mean receive powers corresponding to the different measured locations. However our work differentiates from the previous work in the following aspects:

- (a) Our analysis allows us to exactly identify whether the differences in mean receive powers are caused by shadowing or caused by the different distances between transmit and receive locations that we call path loss differences. It is important to differentiate these two factors, since they are caused by different physical propagation mechanisms. Also, as we will later see, these parameters affect differently the *cdf* of the channel capacity.
- (b) We study the effect of these factors not only in single-user channels but also in

⁷A similar conclusion was also obtained by [Nguyen *et al.* (2004)] when estimating the ergodic capacity of indoor to indoor as well as outdoor to indoor channel propagation measurements

multi-user channels. In our analysis, we found that shadowing affects especially the outage performance of the single-user channels, while the channel capacity of multi-user channels is not only affected by shadowing but also by path loss differences among the largely spaced users.

- (c) We do not limit our analysis to the ergodic performance case. Instead, we quantify each of these parameters and relate their effect on the *cdf* of the channel capacity. The results of our analysis are important since a prior estimate of these parameters can be used to predict upper bounds on the capacities of the measured channels.
- (4) *Comparison of TDM and OSDM schemes:* a theoretical comparison of the sum capacity of the downlink between *TDM* and the optimal dirty paper coding (*DPC*) strategy [Caire and Shamai (2003)] was performed by [Jindal and Goldsmith (2005)]. In this study the authors assumed symmetrical user channels (same statistics for all user channels), channel knowledge at both the transmitter and the receiver and a short-term input power constraint. They found that for Rayleigh channels (at high *SNR*, for a large number of transmit antennas N_T , or for a large number of users K) the ergodic sum capacity of the downlink is $\leq \max\left(\min\left(\frac{N_T}{N_R}, K\right), 1\right) \times \text{TDM sum capacity}$, where N_R is the number of antennas at each user. We observe that the benefit of *DPC* over *TDM* is minimum when $N_T = N_R$.

The main differences of our work with the previous study are: we compare sub-optimal *OSDM* and *TDM* schemes based on real world measured channels. Our measured channels are in fact non symmetrical, mainly because of the largely spaced users. We are interested to evaluate the multi-user channel performance not only in terms of ergodic rates but also in terms of outage rates. Also, our objective is to deliver fair rates for all users instead of obtaining the maximum sum rate. As we will see later in

this thesis, the asymmetry among the different user channels particularly degrades the performance of the *OSDM* schemes.

Some of these contributions were presented in [Yépez *et al.* (2006b), Yépez *et al.* (2006a)].

Thesis overview

The next two chapters of this thesis are descriptive in nature, and the following three chapters of this thesis constitute the analytical contributions of our work. Below is the description of each chapter. Chapter 2 gives a description of the 4×4 MIMO channel sounder that we built as well as a description of the procedure we follow in order to obtain the estimates of channel impulse response. Chapter 3 presents a description of the measurement scenarios and the mathematical description for the measured transfer functions. Chapter 4 presents the statistical analysis we perform in order to determine whether the elements of the measured channels are correlated. In this chapter we also estimate the independent frequency sub-bands in order to analyze the full bandwidth of the measured channels. We use these independent frequency sub-bands estimates in our subsequent analysis. Chapter 5 compares the estimated performances of the measured single-user channels with the corresponding performance of *iid* Rayleigh channels. We also estimate the confidence interval for the shadowing factor, which is shown to be the main physical link channel factor that affects the performance of the measured single-user channels. Chapter 6 compares the relative performance of the measured multi-user channels using different sub-optimal schemes. We also estimate the confidence interval for the mean receive power per user which together with shadowing affect the performance of the measured multi-user channels. Chapter 7 gives a summary of the results, list of contributions and conclusions of our work as well as suggested ideas for future work. Finally, we note that the review of earlier literature is indicated in relevant sections throughout Chapters 4, 5, 6.

Chapter 2

Channel sounder and impulse response estimates

This chapter presents a description of the MIMO measurement system used to record data suitable for the estimation of radio channel impulse responses. We remark that at the time the equipment was developed there was no availability of MIMO channel sounders on the market and therefore to build this equipment was a necessary first step in order to perform the channel measurements for later analysis, which is the subject of this thesis.

Section 2.1 describes the transmitter and receiver, it also describes how the real time transmission takes place. Section 2.2 gives a summary of the main characteristics of the MIMO equipment. Section 2.3 presents a link budget calculation for the operational area of this equipment. Section 2.4 gives a brief explanation of the offline data processing that we conduct to obtain the impulse response estimates as well as an algorithm we use to remove the noise from the measured impulse response estimates.

2.1 MIMO channel sounder

The project of building a MIMO channel sounder was a joint effort between The Communication Research Centre (CRC) and Carleton University. The different phases of design, implementation and testing of the equipment were all completed as part of this thesis, and the outcome is a 4×4 MIMO channel sounder. Further characteristics of the MIMO equipment are presented below.

2.1.1 Transmitter

This module consists of four physical transmitters, and each transmit antenna radiates a maximum power of 30 *dBm*. A block diagram of the transmitter is shown in Figure 2.1. The pseudo-noise generator (pn-generator) consists of four parallel output channels; each channel output generates the same pseudo-noise sequence (pn-sequence) of length 4095 chips at a rate of 40 *Mchps*. The output of each channel is then mixed with a 2.44 *GHz* carrier signal coming from the phase lock oscillator (PLO). The rubidium standard generates a stable 10 *MHz* frequency reference signal that is fed to the clock of the pn-generator and the PLO. Finally, the output of each mixer is bandpass filtered, amplified and sent through its respective antenna.

Array antenna configuration: all the antennas at the transmitter are vertically polarized quarter-wavelength monopoles¹. The antenna array at the transmitter can be configured either in a linear shape or in a square shape. In our measurements we used the linear configuration as shown in Figure 2.2. We also used the maximum separation distance between transmit antennas of 20λ . The transmit antenna array can be easily placed on a building rooftop.

¹The radiation pattern of a quarter-wavelength monopole antenna is the same as the radiation pattern of a dipole antenna

Figure 2.1: MIMO 4×4 transmitter

Figure 2.2: Transmit antenna array

Figure 2.3: MIMO 4×4 receiver

2.1.2 Receiver

Figure 2.3 shows the block diagram of the receiver. The signal is received by a four-element antenna array that is controlled by a rf-switch. The rf-switch selects one element at a time letting its signal pass to the receiver. At the receiver, this signal is bandpass filtered and amplified by the low noise amplifier (LNA). The output signal is then split, in power, in two identical components. One of the components is inphase demodulated and the other component is quadrature demodulated. The outputs of each demodulator are finally lowpass filtered and sampled by the data acquisition card (DAQ) at a sampling rate of 160 MS/s , each. The rubidium standard creates a stable 10 MHz frequency reference signal that is fed to the clock of the DAQ card and to the PLO.

Array antenna configuration: all the antennas at the receiver are vertically polarized quarter-wavelength monopoles. Similar to the transmitter, there are two antenna configurations that can be used in the measurements: a linear shape and a square shape. We

Figure 2.4: Receive antenna array

have used in our measurements the square configuration, see Figure 2.4. The distance separation between the antennas is fixed to one λ . The receive antenna array can easily be accommodated in small spaces such as offices or small rooms.

Acquisition time: for recording the data at the receiver we assume the channel stays constant during the time interval of 1 ms . This is a reasonable assumption since it is less than 100 ms , which is the coherence time of a fixed wireless outdoor to outdoor channel operating at 2.5 GHz with a Doppler spread of 1 Hz reported in [Baum *et al.* (2000)]².

The required time for recording the pn-sequence from each of the four antennas is $102\text{ }\mu\text{s}$. The switching time from one antenna to the other is $\approx 10\text{ }\mu\text{s}$; however we increase the switching time to $102\text{ }\mu\text{s}$ to give enough time to the rf-switch to set to the new position. Therefore the total acquisition time is $204\text{ }\mu\text{s} \times 4 \approx 0.82\text{ ms}$. Figure 2.5 illustrates how the receiver does this acquisition.

²Also, for the case of fixed outdoor to indoor wireless channels operating at 5.2 GHz and where measurements were performed in an office environment the work in [Wyne *et al.* (2004)] reports typical Doppler frequencies of 1 Hz

Figure 2.5: Acquisition time at the receiver

Mutual coupling: it occurs when the antenna elements are closely placed to each other, so that the field generated by one antenna element affects the current distribution of the other antenna elements [Clerckx *et al.* (2003)]. This fact not only distorts the radiation pattern of the other antennas but it also causes a mismatch in their input impedances.

The authors in [Clerckx *et al.* (2003)] showed that there are two mutual coupling effects on the channel capacity of a MIMO system. Their study is based on dipole antennas with different receive antenna separations. They found that when the antenna separation is between $[0.1 - 0.4] \lambda$ the correlation coefficient at the receiver as defined in [Oestges and Erceg (2003)] decreases but it also decreases the effective gain of the antennas. They also found that when the antenna separation is between $[0.4 - 0.8] \lambda$ the correlation coefficient at the receiver increases but it also increases the effective gain of the antennas. Finally they showed that, at $SNR = 10 \text{ dB}$, the 50% outage rates (using equal power allocation) of the MIMO system with receive antenna separations of $[0.4 - 0.8] \lambda$ were higher than the ones of the MIMO system with antenna separations of $[0.1 - 0.4] \lambda$, so that the increase in effective antenna element gains was more beneficial than the decrease in the correlation of the links between transmit and receive antenna elements.

In our measured channels mutual coupling should not be of a concern since the minimum antenna separation of the MIMO system that we used is one λ , at the receiver. There is

clear indication in [Clerckx *et al.* (2003), Özdemir *et al.* (2004)] that this antenna separation is sufficient to mitigate the effects of mutual coupling.

2.2 Operational characteristics

Table 2.1 summarizes the main characteristics of the MIMO equipment.

Centre frequency	2440 <i>MHz</i>
Bandwidth	80 <i>MHz</i>
Output power	4 W (36 <i>dBm</i>)
Noise power (with noise figure=7 <i>dB</i>)	−86 <i>dBm</i>
System gain	93 <i>dB</i>
Excitation	m-sequence (maximum length), Kasami
Sequence Length	4095 (2047, 1023, 511)
Resolution	25 <i>ns</i>
Modulation	BPSK
Transmit rate	40 <i>Mchps</i> from each transmit antenna
Data acquisition	2 quadrature baseband channels, each having 8 bit resolution, sampling at a rate of 160 <i>Mchps</i>
Number of (Tx, Rx) antennas	(4, 4) omnidirectional
Tx. antenna separation	variable up to 20 λ
Rx. antenna separation	1 λ
Antenna polarization	vertical
Frequency reference	Rubidium standard

Table 2.1: MIMO 4 $\times$ 4 system specifications

2.3 Link budget

Table 2.2 shows the link budget for our system. This calculation is useful when one is interested in predicting the coverage area. We note in the table that *MPSR* denotes the multipath power sensitivity ratio which corresponds to the level at which multipath can be detected and the processing gain was calculated assuming a pn-sequence length of 1023

chips. We see in the table that the transmitter can cover distances up to 375 m based on a path loss value of 124 dB .

The path loss value, which is based on the COST 231 model, was taken from an internal technical report by [Hendrantoro (2002)]. It assumes outdoor to indoor propagation conditions, and it is calculated based on an operating frequency of 2.44 GHz , a base station antenna height of 30 m and a user terminal antenna height of 10 m . It also assumes a street width $w = 30$ m , a roof height $h_{roof} = 12$ m , an orientation of the street on which the user terminal is located with respect to the direction of propagation $\gamma = 45^\circ$ and a building separation distance $b = 50$ m .

Later in the thesis we will estimate mean receive powers from the measured data. We note that these estimates are all relative to some unknown reference since they do not include the amplification gain of the receiver.

Effective isotropic radiated power	36 dBm
Processing Gain	30 dB
Receive antenna gain	2 dB
Path loss (according to COST 231 for outdoor to indoor)	124 dB at 375 m
MPSR required at the receiver	30 dB
Minimum received power	-86 dBm

Table 2.2: Link budget

2.4 Impulse response estimates

We explain below the two steps we follow in order to obtain the impulse response estimates from the measured data.

- (1) Online process: it refers to the real time process that is effected at the time we take the measurements. The two sides of the link operate as follows.

- Transmitters: each transmit antenna sends a shifted version of the same m-sequence (maximum length) of 4095 chips. More specifically, each transmitter sends the same m-sequence, which is delayed by 1023 chips with respect to the one sent by the previous transmitter.
 - Receivers: each receive antenna records the signal received over the channels from all four transmit antennas. In particular each receive antenna acquires four samples per chip so that the total signal length is 16380 samples.
- (2) Offline process: after the measurements have been recorded, we process the data to generate the impulse response estimates as follows:
- Correlate the signal acquired by the i^{th} receive antenna with the reference m-sequence³, to obtain the impulse response estimate corresponding to the radio channel between the i^{th} receive antenna and the four transmit antennas.
 - Separate this impulse response estimate into the individual impulse response estimates, one, between each receive-transmit antenna pair (we identify the starting point of the impulse response estimate corresponding to the i^{th} receive and first transmit antenna elements). From this point we divide the impulse response estimate into four individual impulse response estimates of length ≈ 1023 , which correspond to the receive-transmit antenna pairs (i, j) i.e., h_{ij} where $j \in [1 \cdots 4]$.
 - Remove the noise from the impulse response estimate presented in Section 2.5.

2.5 Impulse response estimates after threshold

In order to remove the noise from the impulse response estimate we follow the modified constant false alarm rate (CFAR) algorithm presented by [Sousa *et al.* (1994)].

³The transmitted m-sequence up-sampled by a factor of 4.

The CFAR algorithm consists of finding the threshold value

$$\xi = \eta \sigma_N \quad (2.1)$$

where a valid sample of the impulse response estimate is considered when its magnitude is above the threshold ξ . In (2.1) η is a constant and σ_N is the standard deviation of the noise of the impulse response estimate that we need to evaluate.

We assume that the noise of each of the impulse response estimates h_{ij} is Gaussian, so that the magnitude of the noise samples is Rayleigh distributed. The estimate σ_N is based on the median σ_m of the samples of the impulse response estimate. The median is preferred over the mean since the latter will overestimate σ_N due to the large values of valid samples.

The standard deviation of the noise (σ_N) in terms of the median (σ_m) is found to be:

$$\sigma_N = 0.85 \sigma_m \quad (2.2)$$

[Sousa *et al.* (1994)]; η is found from the false alarm probability equation

$$P_F = \exp(-\eta^2/2) \quad (2.3)$$

where P_{FA} is set to the desired value.

A modification of the CFAR algorithm is given in [Sousa *et al.* (1994)]. The authors consider a valid sample only when at least two consecutive samples are above the threshold. The modified probability of false alarm is given by

$$P_{FA} = 2 (P_F)^2 \quad (2.4)$$

Figure 2.6: Probability of detection vs. false alarm for different values of SNR

Figure 2.7: Enlarged view of P_D vs. P_{FA}

In our data processing we followed the modified CFAR algorithm using (2.4). We found a threshold value of 9.62 dB above the median level σ_m . This threshold value corresponds to a modified false alarm probability of $P_{FA} = 6.1e^{-6}$, which indicates we would have one false alarm every *ten* impulse response estimates.

We are also interested to know what is the probability of detecting a valid sample. This probability is called the probability of detection and it is given by:

$$P_D = Q(\sqrt{SNR_k}, \eta) \quad (2.5)$$

[Sousa *et al.* (1994)], where Q is the Marcum Q -function, which represents the area under the right tail of a Ricean distribution [Kay (1998)]; and $SNR_k = |h(k)|^2 / \sigma_N^2$, is the signal to noise ratio of the k^{th} sample of the impulse response estimate. The probability of false alarm P_{FA} and the probability of detection P_D , both depend on η , as can be observed in Figure 2.6; when η increases, P_{FA} and P_D decrease; when η decreases P_{FA} and P_D increase.

Figure 2.8: Estimate of impulse response with noise

Figure 2.9: Estimate of impulse response after application of the CFAR algorithm

The curves in the figure also depend on the SNR . For example it can be observed that if we set $P_{FA} = 0.1$, then $P_D = 0.4$ when $SNR = 10\text{ dB}$. If we want to increase $P_D \approx 0.99$ then the SNR should be increased to $\approx 15\text{ dB}$.

Figure 2.7 gives a closer look at the P_{FA} values in the range $[1e^{-6} \dots 1e^{-5}]$, which are the ones we are interested in analyzing. Since we have set $P_{FA} = 6.1e^{-6}$, we get $P_D = 0.4$ when $SNR = 10\text{ dB}$; and $P_D = .98$ when $SNR = 15\text{ dB}$.

For typical measured impulse response estimates we obtained 25 dB difference between the noise level σ_N^2 and the largest sample. We then have that on average we will detect a valid sample 98% of the time, whenever the relative sample power is 15 dB above the noise level. Figures 2.8 and 2.9 show the impulse response estimate before and after application of the CFAR algorithm.

Chapter 3

Channel measurements

In this chapter we describe the scenarios where we performed the channel measurements as well as the impulse response and its transfer function that is the subject of this study. This chapter is organized in two sections. Section 3.1 gives a brief description of the different sites where measurements were carried out and how the experiments were performed and Section 3.2 describes the measured channel impulse response estimates and their corresponding transfer functions.

3.1 Description of the measurement scenarios

Figure 3.1 shows a map of Carleton University campus where we took the MIMO channel measurements¹. The measurements were performed during late fall when there were not many leaves on the trees².

The transmitter was positioned on the rooftop of Mackenzie building *ME*, and the receiver

¹All maps and blue prints of the buildings were obtained with the kind help of Carleton University library staff.

²We performed these measurements with the participation of Carleton University students: Rajpreet Dijon, Pamela Quan and Rami Youssef during the fall of 2002.

<i>Receive area</i>	<i>Label</i>	<i>d (m)</i>
Dunton Tower	DU	206
Library	LI	268
Loeb	LO	351
Southam Hall	SA	372

Table 3.1: Distance in m between transmit and receive area

was positioned sequentially in four different receive areas or buildings, namely: Dunton Tower (DU), Library (LI), Loeb (LO) and Southam Hall (SA). The endpoints of the arrows in Figure 3.1 indicate these areas. The reader can refer to Table 3.1 to see the corresponding distance between the transmitter and each receive area.

At the transmitter we used a linear array of four antenna elements with the distance of 20λ between each pair of antennas, where λ is the wavelength corresponding to the operating frequency, 2.44 GHz , of our channel sounder. The total length of the transmit linear array was $\approx 8\text{ m}$. Figure 3.2 shows the positioning of this antenna array on the rooftop of Mackenzie building. The total radiated power was 36 dBm .

At the receiver we used a square array configuration where each antenna element was separated by one λ . We show in Figures 3.3 to 3.6 the four measured receive areas indicated by $[DU, LI, LO, SA]$ respectively. We divided each receive area into three sub-areas and in each sub-area we performed five to six channel measurements in different locations; the separation distance between two consecutive locations was at about 2λ , which gives a reasonable probability of low correlation among most samples [Chizhik *et al.* (2003)]. During each channel measurement the receiver was fixed in a given location. Finally, the total number of measured locations for each receive area was seventeen.

The receive area where measurements were taken in DU was the classroom (602) located on

the sixth floor. The maximum separation among receive locations in this area was 3.5 *m*. This area was further subdivided into three sub-areas and in each sub-area measurements were performed over a maximum separation among receive locations of 0.5 *m*. The separation distance between sub-areas was about 1 *m*. Figure 3.3 depicts the scenario for this receive area.

The receive area in *LI* was one of the corridors on the fifth floor of the library. The maximum separation among receive locations in this area was 45 *m*. This area was further divided into three sub-areas and within each sub-area measurements were taken over a maximum separation among receive locations of 0.62 *m*. The separation distance between sub-areas was about 20 *m*. Figure 3.4 depicts the scenario for this receive area.

The receive area in *LO* where the measurements were done was the seminar room (*C869*) located on the eighth floor. The maximum separation among receive locations in this area was 3 *m*. This area was further divided into three sub-areas and in each sub-area measurements were performed over a maximum separation among receive locations of 0.5 *m*. The separation distance between sub-areas was about 1 *m*. Figure 3.5 depicts the scenario for this receive area.

Finally, the receive area in *SA* was one of the halls located on the fifth floor. The maximum separation among receive locations in this area was 3.5 *m*. This area was further divided into three sub-areas and within each sub-area measurements were taken over a maximum separation among receive locations of 0.5 *m*. The separation distance between sub-areas was 1 *m*. Figure 3.6 depicts the scenario for this receive area.

We note that the relative floor height of *SA* was two floors below the other receive areas i.e.,

the relative floor heights of DU , LI , LO were all the same.

3.2 Mathematical description of the impulse response and its transfer function

We use the theoretical framework, originally developed by [Bello (1963)] for time variant channels, for analyzing our channel measurements with the assumption that the same theory can be applied for location varying channels that are not time variant. We use these location variations to characterize the channel performance in a given receive area. More specifically, we perform our measurements by positioning the receive unit in different locations of a given receive area and analyze the performance of each receive area in terms of ergodic and outage capacities.

Since capacity expressions and in general theoretical frameworks for channel modelling are developed for the low pass transfer function of the channel, we perform our analysis in the frequency domain and recall the fact that zero mean uncorrelated scattering (US) channels in time domain become zero mean wide sense stationary (WSS) channels in frequency domain [Bello (1963)].

It is shown in Equation (A.4) of Appendix A that the complex valued lowpass impulse response of a radio channel is given by

$$\tilde{h}(\tau, t) = \sum_{i=1}^L \alpha_i(t) \delta(\tau - \tau_i(t)) \quad (3.1)$$

Figure 3.1: The start point of the arrows indicate the transmit location Mackenzie building (*ME*). The receive areas are denoted by the cloud signs at the end of each of the arrows, from left to right: Dunton Tower (*DU*), Library (*LI*), Southam Hall (*SA*) and Loeb (*LO*).

Figure 3.2: Transmit area: *ME* building; roof on the 5th floor. The cloud sign on the upper right block denotes the location of the linear antenna array with four antennas of length of ≈ 8 m. The arrows point to the direction where the receive areas were located.

Figure 3.3: Receive area *DU*. Sixth floor, classroom (602). The small cloud signs represent the sub-areas where the measurements were taken. The arrow points in the direction of the transmitter.

Figure 3.4: Receive area *LI*. Along one of the corridors in the fifth floor. The small cloud signs represent the sub-areas where the measurements were performed. The arrow points in the direction of the transmitter.

LOEB BUILDING

LEVEL 6

APRIL 25TH, 2003

FILE: 15(LA)-LV6.DWG

Carleton
UNIVERSITY
Construction Services
Department Of Physical Plant
Prepared by: Raymond G. Dufresne (CAI)

Figure 3.5: Receive area *LO*. Eighth floor, seminar room (*C869*). The small cloud signs represent the sub-areas where the measurements were performed. The arrow points in the direction of the transmitter.

Figure 3.6: Receive area SA ; Along one of the halls in the fifth floor. The small cloud signs represent the sub-areas where the measurements were performed. The arrow points in the direction of the transmitter.

where the impulse response coefficients are given by

$$\alpha_i(t) = a_i(t) e^{-j\phi_i(t)} \quad (3.2)$$

and where $a_i(t)$ represents the attenuation of the i^{th} propagation path at time t , $\tau_i(t)$ is the associated delay, $\phi_i(t)$ is the associated phase and L is the number of radio paths between transmitter and receiver.

Since our interest is the frequency representation of the lowpass impulse response we proceed to take the Fourier transform of the lowpass impulse response with respect to the delay parameter, which is

$$\tilde{H}(s, t) = \int \tilde{h}(\tau, t) e^{-2\pi s\tau} d\tau \quad (3.3)$$

In the next section we describe the measured lowpass impulse response estimates and their transfer functions, which are the subject of this study.

3.3 Measured impulse response estimates and their transfer functions

Our measurements correspond to the discrete representation of the lowpass impulse response defined in Equation (A.12) of Appendix A. In order to describe each of the elements of the MIMO channel we present the following definitions. Let $h_{i,j}(n, l)$ be the discrete representation of the lowpass impulse response and let $H_{i,j}(f, l)$ be its transfer function where the subscripts i and j refer to the i^{th} receive antenna element and j^{th} transmit antenna element; n is the delay index; f represents the frequency index and l denotes the sample index (see Appendix F for more details on notation and definitions).

Figure 3.7: A measured impulse response $h_{ij}(n)$, [$25 \mu s$ corresponding to 4095 samples]

Figure 3.8: A measured transfer function $H_{ij}(f)$, [$160 MHz$ corresponding to (4095) frequency components]

Before continuing with the description of a measured channel impulse response and its transfer function, we recall that the measurements correspond to a 4×4 MIMO system, where the length of each impulse response (sampled at $160 MS/s$) is 4095 samples; there are also seventeen independent channel measurements for each of the four buildings, namely: *DU*, *LI*, *LO*, and *SA*.

Figure 3.7 shows an example of a measured impulse response from transmitter j to receiver i , after suppressing the noise as explained in Section 2.5. The length of each impulse response is 4095 samples, and the impulse response duration is $25 \mu s$.

Figure 3.8 shows the corresponding discrete transfer function mentioned above where there are 4095 frequency components spaced by $39 kHz$.

Figure 3.9: (a) A transfer function from a back to back measurement (4095 frequencies); (b) An enlarged view of the 3 dB bandwidth (1000 frequencies).

Finally, we are interested to analyze the range of frequencies that correspond to the 40 MHz system bandwidth. As an example of how we select these range of frequencies, we show in Figure 3.9 a measured transfer function from a back to back measurement. Figure 3.9 (a) depicts a transfer function corresponding to one of transmitters and the receiver. We observe in this figure that the captured range of frequencies is 160 MHz. Figure 3.9 (b) depicts the selected range of frequencies corresponding to the 3 dB bandwidth of the frequency spectrum i.e., 40 MHz that are captured in 1000 frequency components. It can also be observed in this figure that the DC component of the frequency spectrum has been removed, since we are only interested in characterizing the random component of the lowpass impulse response.

Chapter 4

Estimation of spatial correlations

It has been shown in [Winters (1987), Telatar (1995), Foschini (1996)] that at high SNR the channel capacities of $m \times m$ MIMO single-user channels increase linearly as a function of m under *iid* Rayleigh fading conditions. However, in real world propagation environments the elements of the MIMO channels¹ may be correlated e.g., when the separation between the antenna elements is not sufficient, decreasing these capacities [Chuah *et al.* (2002)].

In this chapter we propose a method to estimate these spatial correlations by analyzing the eigenvalues of the channel covariance matrix. Assuming that the elements of the MIMO channel are joint circularly symmetric complex normally distributed in space, then we note that MIMO channels with highly correlated elements would correspond to an ill conditioned channel covariance matrix, while MIMO channels with uncorrelated elements would correspond to a well conditioned channel covariance matrix. Our objective is to predict the decrease in MIMO channel capacity through the estimation of the number of non-significant eigenvalues of the channel covariance matrix that we use to indicate the degree of correlation in the elements of the MIMO channel.

¹In this chapter we focus our study on the case of MIMO single-user channels and we will often refer to them as MIMO channels

We summarize in the following three steps the method we propose to estimate the spatial correlations in the elements of the measured MIMO channels:

- (i) Estimate the independent frequency sub-bands of the transfer functions of the measured MIMO channel so that we can use them to analyze all the channel bandwidth (Section 4.1). We use the estimated independent frequency sub-bands in the next steps.
- (ii) Validate the Kronecker model assumption where the spatial correlations in the elements of the measured MIMO channels can be characterized in terms of the transmit and the receive covariance matrices (Section 4.2).
- (iii) Estimate the number of non-significant eigenvalues of the transmit and the receive channel covariance matrices (Section 4.3).

The material of this chapter is organized as follows. Sections 4.1 to 4.3 consist of a mathematical description of the parameter we want to estimate or validate from the measured channel as well as the mathematical description of the statistical tool we use for its validation or estimation; a description of how we apply the statistic to our measured data, and the results of these evaluations. Section 4.4 illustrates with some examples the decrease in capacity of some theoretical correlated channels as well as the capacity loss that would be expected in our measured channels according to their estimated degree of correlation.

The reader is referred to Appendix B for definitions of complex proper normal and circularly symmetric normal distributions used in this chapter.

4.1 Estimation of the independent frequency sub-bands

In order to perform the analysis of the measured broadband channels, it is necessary to estimate the independent frequency sub-bands from each of the measured transfer functions (consisting of 1000 frequency components each). We can then analyze the measured broadband channels based on their estimated independent frequency sub-bands by using simultaneous tests as described in [Looney (1995)].

In the following we explain the test statistic we use to estimate the independent frequency sub-bands of the measured broadband channels in order to proceed with the subsequent data analysis.

4.1.1 Mathematical framework for the independence test

We apply the test statistic for independence of sets of variates described in [Krishnaiah *et al.* (1976)] which will help us determine whether two frequency components from a given transfer function are independent or not. The following explains how to use this test statistic.

Let $\mathbf{z}_1, \dots, \mathbf{z}_i, \dots, \mathbf{z}_q$ be the p – variate (size $p \times 1$) complex proper normally distributed vectors with mean vectors μ_i and covariance matrices $\Sigma_{ij} = E\{(\mathbf{z}_i - \mu_i)(\mathbf{z}_j - \mu_j)^\dagger\}$ where $i, j \in [1, \dots, q]$. We are interested to test the hypothesis

$$\mathcal{H}_0 : \Sigma_{ij} = \mathbf{0} \text{ for } i \neq j = 1, \dots, q \quad (4.1)$$

Let $\mathbf{Z}_i = [\mathbf{z}_i(1) \cdots, \mathbf{z}_i(l) \cdots, \mathbf{z}_i(N)]$ where $\mathbf{z}_i(l)$, $l \in [1 \cdots N]$, is the l^{th} observation of the random vector $\mathbf{z}_i$ and $i \in [1, \cdots, q]$. We then define:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1q} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2q} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{A}_{q1} & \mathbf{A}_{q2} & \cdots & \mathbf{A}_{qq} \end{bmatrix} \quad (4.2)$$

where $\mathbf{A}_{ij} = (\mathbf{Z}_i - \bar{\mathbf{Z}}_i)(\mathbf{Z}_j - \bar{\mathbf{Z}}_j)^\dagger$, $i, j \in [1, \cdots, q]$ and $\bar{\mathbf{Z}}_i = N^{-1} \sum_{l=1}^N \mathbf{Z}_i(l)$. The likelihood ratio statistic for testing $\mathcal{H}_0$ is defined as

$$\lambda = \frac{\det(\mathbf{A})}{\prod_{j=1}^q \det(\mathbf{A}_{jj})} \quad (4.3)$$

[Krishnaiah *et al.* (1976)]. Let $\tilde{\lambda} = -2\ln(\lambda)$, then the test rejects $\mathcal{H}_0$ if $\tilde{\lambda} > c_\alpha$, where c_α corresponds to the upper $100\alpha\%$ point of the distribution of $\tilde{\lambda}$, and α is the level of significance.

Tabulated percentage points of $\tilde{\lambda}$, based on (4.3), are given in [Krishnaiah *et al.* (1976)], however they are not given for our parameters of interest. We find the empirical distribution of this test statistic with the parameters of interest using 100000 simulation points to get a good approximation of the distribution of $\tilde{\lambda}$ which is shown in Figure C.1 of Appendix C. We obtain the percentage points c_α from this distribution and reject $\mathcal{H}_0$ when $\tilde{\lambda}$, evaluated with our data, is greater than c_α .

4.1.2 Application of the test

Below we first describe two assumptions we make in order to estimate the independent frequency sub-bands of a measured transfer function and then we describe the procedure we

follow in order to apply this test to our data.

We presuppose that the delay elements of the lowpass impulse response, or equivalently the frequency components of the transfer function, are multivariate complex proper normally distributed²; otherwise the results of this analysis can still be viewed as an approximation to the true estimate. Also, in order to estimate the independent frequency sub-bands we assume that the width of these sub-bands is the same through all the measured bandwidth.

The following describes the procedure we use in order to estimate the independent frequency sub-bands of the broadband measured channels. Since our objective is to divide the measured bandwidth of the channel into independent frequency sub-bands of the same width, we estimate the minimum frequency separation (sub-band width) for which a pair of frequencies of a given transfer function are independent. Also, since there are sixteen transfer functions in the 4×4 MIMO measured channel our approach is to select the maximum minimum frequency separation, over all the transfer functions, that guarantees simultaneous independence in all the transfer functions of the MIMO channel.

In order to estimate the minimum separation distance for which a pair of frequency components are independent for all transfer functions simultaneously, we perform the test over the sixteen transfer functions with an adjusted level of significance that is determined as follows. Let's denote the nominal level of significance by α^* ; since we need to perform sixteen simultaneous tests, it is appropriate to use Bonferroni's adjusted level of significance $\alpha = \alpha^*/s$ [Looney (1995)]³, where $s = 16$ is the number of simultaneous tests.

²Using the fact that if $\mathbf{x}$ is distributed according to $\mathcal{N}_m(\mu, \Sigma)$, then any linear combination $\mathbf{y} = \mathbf{C}\mathbf{x}$ is distributed according to $\mathcal{N}_m(\mathbf{C}\mu, \mathbf{C}\Sigma\mathbf{C}^\dagger)$ [Anderson (1984)]

³In this work the adjusted level of significance α will be referred simply as the level of significance

Finally we choose 0.05 as the value for the nominal level of significance for most of the tests we use in our analysis, since this value is commonly employed for hypothesis testing. However, for testing the null hypotheses related to multivariate normality and independent frequency sub-bands we select 0.1 as the nominal significance level. This is because we do not have previous knowledge about the nature of the measured data, so we want to decrease the probability of not rejecting a false null hypothesis.

4.1.3 Test results

As indicated in Section 3.3 we first select the 3 *dB* bandwidth for each of the measured transfer functions (we recall that there are sixteen transfer functions in each MIMO channel measurement). Figure 4.1 illustrates this procedure for the case of one measured transfer function. Figure 4.1 (a) depicts a measured transfer function with a total of 4095 frequency components, and Figure 4.1 (b) depicts the 3 *dB* frequency bandwidth with 1000 equally-spaced frequency components.

We then select the independent frequency components. To do so we use the empirical distribution of the statistic $\tilde{\lambda}$, shown in Figure C.1 of Appendix C, to obtain the 90 percentage point of $c_\alpha = 0.6773$ which is calculated based on the following parameters: $p = 1$, $q = 2$, $N = 17$, $s = 16$, $\alpha = \alpha^*/s$. We evaluate this test with a nominal level of significance $\alpha^* = 0.1$ which means that there is a 10% chance of incorrectly rejecting the true hypothesis given in (4.1). Any pair of frequency components that are separated by the minimum frequency separation Δ_f in *MHz*, specified in Table 4.1.3, are considered independent. For convenience we choose the largest minimum frequency separation among all buildings as a common minimum frequency separation to facilitate the process of the succeeding statistical analysis. As a result we have that for all the measured buildings a frequency separation of $\Delta_f = 3.3$ *MHz* between the frequency components of a transfer function guarantees their independence.

Figure 4.1: (a) A measured transfer function; (b) An expanded view of the 3 *dB* bandwidth of that transfer function

Building	Δ_f (MHz)
LI	1.1
LO	2.3
SA	2.7
DU	3.3

Table 4.1: Minimum separation in *MHz* that guarantees independence between a pair of frequencies

Figure 4.2: Measured transfer function (thin line) and estimated independent frequency sub-bands (thick line with discrete steps)

Note that if there is any DC offset caused for example by the measurement equipment it will affect not only the zero frequency but also neighboring frequencies around zero (DC-spread) due to the replacement of noisy samples of the lowpass impulse response with zeroes, which is done in the noise threshold routine. In our analysis we consider frequencies starting at $\pm 1.71 \text{ MHz}$ which is enough distance to avoid any DC-spread in the frequency components.

We finally obtain that there are at least eleven independent frequency sub-bands of width 3.3 MHz . Our statistical analysis will be based on these eleven independent frequency sub-bands. Figure 4.2 shows an example of a measured transfer function and the estimated eleven independent frequency sub-bands.

In the following two sections all scalar, vector and matrix quantities are understood to be at a given arbitrarily chosen frequency.

4.2 Validation of the Kronecker model

When the entries of the MIMO channel matrix are complex proper normally distributed a complete characterization of the MIMO channel matrix is given in terms of its mean vector and its covariance matrix. Our analysis is based on the covariance matrix, which characterizes the random component of the channel only. The significance of our study is supported by the results of Appendix F where we estimated that at least 90% of the power resides in the random component of the channel. In the following we present the mathematical descriptions of the General and the Kronecker models for representing spatial correlations among the elements of the MIMO channel.

General model: Assuming that the elements of the MIMO channel matrix $\mathbf{H}$ are joint circularly symmetric complex normally distributed in space then any correlation in the elements of $\mathbf{H}$ is characterized by $\mathbf{R}$ which is the channel covariance matrix⁴. The MIMO channel can then be written as:

$$\text{vec}(\mathbf{H}) = \mathbf{R}^{1/2} \mathbf{h}_w \quad (4.4)$$

where $\mathbf{h}_w = \text{vec}(\mathbf{H}_w)$, $\mathbf{H}_w$ is a $N_R \times N_T$ matrix whose elements are $\mathcal{CN}(0, 1)$ ⁵ and $\mathbf{R} = E\{\text{vec}(\mathbf{H}) \text{vec}(\mathbf{H})^\dagger\}$.

Kronecker model: The Kronecker model allows one to represent the spatial correlations in the elements of MIMO channels in terms of transmit and receive covariance matrices.

⁴ $\mathbf{R}$ may also characterize the effect of using different antenna polarizations by having different variances [Nabar *et al.* (2004)], however this is not a concern in our measured channels since we used only one antenna polarization

⁵In general, in this work, we also write $\mathbf{x}$ is $\mathcal{F}(\cdot)$ to indicate that $\mathbf{x}$ is distributed according to $\mathcal{F}(\cdot)$

This model assumes that the transmit covariance matrices of each of the transmit vectors associated with each of the receive antennas are all the same, because the angular spectra at the transmitter is the same for signals arriving from any of the receive antennas. This also applies for the case of the receive covariance matrices of each of the receive vectors [Paulraj *et al.* (2003)]. Our objective is to find whether or not the Kronecker model is a suitable representation of the measured channels.

Let the transmit vector $\mathbf{h}_{ti}$, $i \in [1, \dots, N_R]$, denote the channel elements involved when the transmit antenna array radiates to receive antenna i i.e.,

$$\mathbf{h}_{ti} = [h_{i,1}, \dots, h_{i,N_T}]^T \quad (4.5)$$

Then the spatial variations at the transmit end caused by the i^{th} receive antenna, $i \in [1, \dots, N_R]$, can be represented by

$$\mathbf{h}_{ti} = \mathbf{R}_{T_i}^{1/2} \mathbf{h}_t \quad (4.6)$$

where $\mathbf{h}_t$ is a $N_T \times 1$ vector with elements distributed according to $\mathcal{CN}(0, 1)$ and $\mathbf{R}_{T_i}$ represents the i^{th} transmit covariance matrix given by

$$\mathbf{R}_{T_i} = E\{\mathbf{h}_{ti} \mathbf{h}_{ti}^\dagger\} \quad (4.7)$$

Similarly we define the j^{th} receive vector $\mathbf{h}_{rj}$, where $j \in [1, \dots, N_T]$, which represents the channel elements involved in the radiation from j^{th} transmit antenna to all the receive antennas

$$\mathbf{h}_{rj} = [h_{1,j}, \dots, h_{N_R,j}]^T \quad (4.8)$$

Then the spatial variations at the receive end caused by the j^{th} transmit antenna can be represented by

$$\mathbf{h}_{rj} = \mathbf{R}_{R_j}^{1/2} \mathbf{h}_r \quad (4.9)$$

where $\mathbf{h}_r$ is a $N_R \times 1$ vector with elements distributed according to $\mathcal{CN}(0, 1)$, and $\mathbf{R}_{R_j}$ represents the j^{th} receive covariance matrix given by

$$\mathbf{R}_{R_j} = E\{\mathbf{h}_{rj} \mathbf{h}_{rj}^\dagger\} \quad (4.10)$$

Assuming that the transmit random vectors $\mathbf{h}_{ti}$ $i \in [1, \dots, N_R]$ are independent then if all the transmit covariance matrices are equal i.e., $\mathbf{R}_T = \mathbf{R}_{T_1} = \dots = \mathbf{R}_{T_{N_R}}$ (or equivalently if all the receive covariance matrices are equal) then the MIMO channel can be represented in terms of the transmit and receive covariance matrices as shown below

$$\mathbf{H} = \mathbf{R}_T^{1/2} \mathbf{H}_w \mathbf{R}_R^{1/2} \quad (4.11)$$

[Chuah *et al.* (2002), Jafar (2003)], where $\mathbf{H}_w$ is the MIMO channel matrix with *iid* Rayleigh elements.

In the literature, the Kronecker model is usually validated by estimating the upper triangular correlation coefficients of the transmit and receive covariance matrices. These coefficients are in turn used to generate correlated simulated channels. Finally, one would determine the suitability of Kronecker depending of how well the ergodic capacity of the simulated channels predicts that of the measured channels [Kermoal *et al.* (2002), Chizhik *et al.* (2003), Özcelik *et al.* (2003a)].

In the approach that we propose below, one does not need to estimate a number of correlation coefficients of the covariance matrices (which grows proportionally with the number of

transmit and receive antennas). Instead we determine the suitability of Kronecker by using one of the tests of multivariate statistics to estimate only one parameter that characterizes a desired property of the structure of the covariance matrices.

4.2.1 Mathematical framework for the multiple homogeneity test

Following we give a mathematical description of the test we use in order to evaluate the suitability of the Kronecker channel model to represent our measured channel.

Let $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q$ be the p -variate (size $p \times 1$) independently distributed complex proper normal vectors, with mean vectors $\mu_1, \dots, \mu_q$ and covariance matrices $\Sigma_{11}, \dots, \Sigma_{ii}, \dots, \Sigma_{qq}$, where $\Sigma_{ii} = E\{(\mathbf{z}_i - \mu_i)(\mathbf{z}_i - \mu_i)^\dagger\}$. We are interested to test the hypothesis

$$\mathcal{H}_5 : \Sigma_{11} = \dots = \Sigma_{qq} \quad (4.12)$$

Let $\mathbf{Z}_i = [\mathbf{z}_i(1), \dots, \mathbf{z}_i(l), \dots, \mathbf{z}_i(N_i)]$, where $\mathbf{z}_i(l)$ is the l^{th} observation of the random vector $\mathbf{z}_i$ and $i \in [1, \dots, q]$ is the index for the independent random vector we want to test. We define: $\mathbf{A}_{ii} = (\mathbf{Z}_i - \bar{\mathbf{Z}}_i)(\mathbf{Z}_i - \bar{\mathbf{Z}}_i)^\dagger$ and $\bar{\mathbf{Z}}_i = N_i^{-1} \sum_{l=1}^{N_i} \mathbf{z}_i(l)$.

The likelihood ratio statistic for testing $\mathcal{H}_5$ is defined as

$$\lambda_5 = \frac{\prod_{i=1}^q \det(\mathbf{A}_{ii}/n_i)}{\det(\sum_{i=1}^q \mathbf{A}_{ii}/n_q)^{n_q}} \quad (4.13)$$

[Krishnaiah *et al.* (1976)]; where $n_i = N_i - 1$ and $n_q = \sum_{i=1}^q n_i$.

Let $\tilde{\lambda}_5 = -2 \ln(\lambda_5)$, then the test rejects $\mathcal{H}_5$ if $\tilde{\lambda}_5 > c_\alpha$, where c_α corresponds to the upper 100 α % point of the distribution of $\tilde{\lambda}_5$ and α is the level of significance.

Tabulated percentage points of this statistic are given in [Krishnaiah *et al.* (1976)] but they are not given for the parameters of interest. We find the empirical distribution of $-2\ln(\lambda_5)$ with the parameters of interest using 1000000 simulation points as shown in Figure C.3 of Appendix C. From this distribution we obtain the percentage points c_α and evaluate $\tilde{\lambda}_5 = -2\ln(\lambda_5)$ with our data; if the result of this evaluation is above c_α we reject $\mathcal{H}_5$.

4.2.2 Application of the test

Let's first introduce the definition of p-value often used in statistics when analyzing tests of significance.

P-value: Let the one sided test statistic associated with the null hypothesis be given by T and the statistic of the observed sample be given by T_{obs} then assuming the null hypothesis is true,

$$\text{p-value} = \text{Prob}\{T > T_{obs}\} \quad (4.14)$$

[Kalbfleisch (1985)]⁶. This is the probability of obtaining a value as extreme as the one observed from the sample assuming the null hypothesis is true. The lower the p-value the stronger the evidence against the null hypothesis.

We will include the mentioned p-value in our analysis of tests of significance and for convenience we will express the p-value in terms of the level of significance α . When (p-value $< \alpha$) we will reject the null hypothesis and when (p-value $> \alpha$) we will not have evidence to reject the null hypothesis.

We now continue with the validation of the conditions that need to be satisfied in order to perform the multiple homogeneity test. The first condition is that the data must be

⁶For a two sided test: p-value = $\text{Prob}\{T > T_{obs}\} + \text{Prob}\{T < -T_{obs}\}$

multivariate normally distributed. We analyzed this condition with goodness of fit tests in Appendix D. The results of these tests indicate that the data from the receive areas of $[DU, LI]$ cannot be considered multivariate normal. We will confirm later in Chapter 5 that the reason for rejection of multivariate normality is the significant amount of shadowing caused by different obstructing buildings in the transmit-receive paths of the different sub-areas of LI . We will also find in Chapter 5 that to a lesser extent DU is also affected by shadowing of near buildings. As a result, we do not reject the multivariate normality condition only for the receive areas of $[LO, SA]$.

The second condition specifies that the q random vectors: $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_q$ must be p -variate independent. We evaluated this condition for the receive areas of $[LO, SA]$ in Appendix G and could not find any evidence for rejecting $\mathcal{H}_5$. Therefore we apply the multiple homogeneity test only to the receive areas where the two test conditions could not be rejected i.e., $[LO, SA]$.

We also note that when analyzing the spatial correlations of MIMO measured channels, it is important that there are no significant distance variations over the different measured locations so that the propagation scenario remains the same [Herdin *et al.* (2005), Bonek (2006), Almers *et al.* (2006)].

In our case the measured channels of $[LO, SA]$ were taken in an seminar room and in a hall respectively. We performed the channel measurements over a maximum separation distance among receive locations of 3.5 m (or $\approx 30\lambda$). We consider that there is no significant change on the statistics of the measured channels over this separation distance⁷.

⁷We note that this separation distance is similar to that used in [Chizhik *et al.* (2003)], where the authors consider that the statistics of the measured outdoor to outdoor channels are the same

4.2.3 Test results

We now test the equality of all the transmit sample covariance matrices corresponding to each one of the transmit vectors specified above with a sample size $N = 17$. The nominal level of significance that we use with this independence test is $\alpha^* = [0.05]$. Since we apply this test over each of the eleven independent frequency sub-bands simultaneously we adjust the level of significance to $\alpha = [\alpha^*/11]$. From the empirical distribution given in Figure C.3 of Appendix C we obtain the following percentage point $c_\alpha = [86.3450]$, which is the upper limit of the non rejection region of the hypothesis $\mathcal{H}_5$.

Figure 4.3 shows the evaluated test points using our data. The thick horizontal line represents the percentage point c_α mentioned above; any evaluated point above this limit will contribute to the rejection of $\mathcal{H}_5$. We see that the evaluated points for all frequency sub-bands of the two receive areas are under c_α . We also see in the figure the maximum p-values, over all the independent frequency sub-bands, of each receive area. We observe that the p-values of both receive areas are larger than α so that there is no evidence for rejecting $\mathcal{H}_5$. The above results indicate that the data of the receive areas of $[LO, SA]$ is consistent with the Kronecker model assumption used to represent the spatial correlations in their channel elements.

Our results are in agreement with those found in a similar study given in [Kermoal *et al.* (2002)] where the authors also show Kronecker to be a suitable model for the 4×4 outdoor to indoor measured channels. It has also been shown in [Özcelik *et al.* (2003a), Bonek *et al.* (2003), Özcelik *et al.* (2005), Wyne *et al.* (2005), Almers *et al.* (2006)] that as we increase the number of transmit and receive antenna elements, e.g. 8×8 , Kronecker can no longer represent the spatial correlations of the measured channels accurately since it underestimates the channel capacity.

Figure 4.3: Multiple homogeneity test statistic vs. frequency index. The horizontal line at [86.3450] is the upper limit of the one sided test with nominal level of significance $\alpha^* = [0.05]$.

4.3 Estimation of spatial correlations

We use the fact that the spatial correlations in the elements of the channels corresponding to the measured receive areas $[LO, SA]$ can be represented using the Kronecker model given in (4.11), so that we can focus our analysis on the study of the transmit and receive covariance matrices of the MIMO channel. Our target is to estimate the number of non-significant eigenvalues of these covariance matrices to indicate the degree of correlation in the elements of the measured MIMO channel. Towards this end we use one of the tests of principal components that we describe below.

4.3.1 Mathematical framework for the principal components test

Principal components analysis explains the variability of a random vector $\mathbf{x}$ through the eigen structure of its covariance matrix $\mathbf{\Sigma}$. More specifically, principal components analysis characterizes the variability of $\mathbf{x}$ by a new variable with a smaller number of components with large variance [Muirhead (1982)]. To do this, it uses the fact that the total variability

of a random vector $\mathbf{x}$ can be measured in terms of $Tr(\Sigma)$ or $\det(\Sigma)$. In our work we are particularly interested in one of the tests of principal components that allows one to estimate the number of non-significant eigenvalues of the covariance matrix. We will use this test in our analysis to measure the degree of correlation in the variables of the random vector $\mathbf{x}$.

In the following, we first make an extension to the mathematical derivation of the principal component test, which assumes real normal random variables to the case of complex normal random variables and find an analytical expression for its distribution. We then build a confidence interval for the parameter of interest that will allow us to estimate the number of non-significant eigenvalues of a given covariance matrix of a random vector with circularly symmetric complex normal entries.

We extend the results given in [Anderson (1984)] for the case of real normal random vectors to the case of complex proper normal vectors, using the results on the distribution of the generalized sample variance of complex proper normal vectors given in [Goodman (1963)]. To this end, we first formulate the null hypothesis and the statistic to evaluate it. We then present the asymptotic distribution for this statistic; and finally we build a confidence interval for this statistic based on its asymptotic distribution.

Let $N = n + 1$ independent vectors $\mathbf{x}_i$, each distributed according to $\mathcal{CN}_m(\mathbf{0}, \Sigma)$ $i \in [1, \dots, N]$ with an unbiased sample covariance matrix $\mathbf{S}$; let $\lambda_1 > \lambda_2 > \dots \lambda_m > 0$, be the eigenvalues of Σ and $l_1 > l_2 > \dots l_m > 0$, be the eigenvalues of the sample covariance matrix $\mathbf{S}$; also let $q = m - k$ be the number of smallest eigenvalues of Σ where $k \in [1, \dots, m - 1]$.

We want to know the ratio of contribution of the smallest eigenvalues of Σ with respect to the total contribution of all the eigenvalues of Σ . If this ratio is sufficiently small, then

we consider that only the first k eigenvalues are significant.

The null hypothesis to consider is

$$\mathcal{H}_q : \frac{\sum_{i=k+1}^m \lambda_i}{\sum_{i=1}^m \lambda_i} = h \quad (4.15)$$

Our target is to tune the values of h and q until the equality holds in (4.15). $\mathcal{H}_q$ is evaluated as follows. Let

$$C_q = \frac{\sum_{i=k+1}^m \lambda_i}{\sum_{i=1}^m \lambda_i} - h \quad (4.16)$$

and let the statistic that estimates C_q be given by

$$T_q = \frac{\sum_{i=k+1}^m l_i}{\sum_{i=1}^m l_i} - h \quad (4.17)$$

then the asymptotic distribution of g_q when $n \rightarrow \infty$, where g_q is given by

$$g_q = \sqrt{n} \left[T_q - C_q \right], \quad (4.18)$$

is $\mathcal{N}(0, \tau^2)$; with variance τ^2 equal to

$$\tau^2 = \left(\frac{h}{Tr(\Sigma)} \right)^2 \sum_{i=1}^k \lambda_i^2 + \left(\frac{1-h}{Tr(\Sigma)} \right)^2 \sum_{i=k+1}^m \lambda_i^2 \quad (4.19)$$

see proof in Appendix H.

We can now construct the following confidence interval for C_q

$$T_q - \sqrt{\frac{\tau^2}{n}} z^{1-\alpha} \leq C_q \leq T_q + \sqrt{\frac{\tau^2}{n}} z^{1-\alpha} \quad (4.20)$$

where $z^{1-\alpha}$ is the $100(1 - \alpha)$ percentage point of the standard normal distribution. When this interval is centered at zero and sufficiently small then there is no evidence to reject the hypothesis $\mathcal{H}_q$.

To determine the values of h and q for which equality holds in (4.15), we found more convenient to express (4.18) in terms of h so that we have

$$\tilde{g}_q = \sqrt{n} \left[\tilde{T}_q - \tilde{C}_q \right], \quad (4.21)$$

where $\tilde{T}_q = T_q/h$ and $\tilde{C}_q = C_q/h$. When $n \rightarrow \infty$ the asymptotic distribution of (4.21) is $\mathcal{N}(0, \tilde{\tau}^2)$ where

$$\tilde{\tau}^2 = \left(\frac{1}{Tr(\Sigma)} \right)^2 \sum_{i=1}^k \lambda_i^2 + \left(\frac{1-h}{h Tr(\Sigma)} \right)^2 \sum_{i=k+1}^m \lambda_i^2 \quad (4.22)$$

When equality holds in (4.15) the statistic $\tilde{T}_q$ is distributed according to $\mathcal{N}(0, \tilde{\tau}^2)$. Our target is to tune the values of h and q until we find that the confidence interval for $\tilde{C}_q$ is small enough and centered at zero in which case we no longer reject the hypothesis $\mathcal{H}_q$.

4.3.2 Application of the test

We perform this test in two steps. In the first step we estimate the confidence interval for $\tilde{C}_q(f)$ using the bootstrap *BCa* method, where $f \in [1, \dots, F = 11]$ is the index number for the independent frequency sub-band. In the second step we tune the values of both h and q until we can no longer reject $\mathcal{H}_q$.

Bootstrap confidence interval for $\tilde{T}_q$ [Efron (1993)] : Since we have eleven independent frequency sub-bands, we find $\tilde{T}_q(f)$ given in (4.21) for each of the frequency sub-bands.

We want to know how accurate these estimates are. To this end we use *BCa* method introduced in Appendix E to estimate the confidence interval for $\tilde{C}_q(f)$ as follows.

From the available data we have that for each frequency component f , there are $N = 17$ measurements of 4×4 MIMO channel matrices, which we denote by $\mathbf{H} = \langle \mathbf{H}(1), \dots, \mathbf{H}(N) \rangle$. The 4×4 bootstrap sample is denoted by $\mathbf{H}^*(b)$ and corresponds to the b^{th} element of the data set $\mathbf{H}$ where b is a uniform random index number in the range of $[1, \dots, N]$. The bootstrap realization $\mathbf{H}^*$ is a set of $N = 17$ bootstrap samples $\mathbf{H}^*(b)$ which are chosen by resampling uniformly with repetition the original data set $\mathbf{H}$. The bootstrap estimate at the f^{th} frequency, $\tilde{T}_q(f)^*$, is determined from each bootstrap realization. Finally we use $B = 10000$ bootstrap realizations to estimate the 90% confidence interval for $\tilde{C}_q(f)$ by using Equation (E.2) of Appendix E. We also evaluate the error of these confidence intervals by estimating the variance of the endpoints using Equation E.7 of Appendix E.

Finding the values of h and q : in order to find the value of h and the number of smallest eigenvalues q for which the hypothesis $\mathcal{H}_q$ is not rejected in (4.15), we need to look for a confidence interval for $\tilde{C}_q(f)$, which is of relative small length and is centered at zero. Since we have eleven independent frequency sub-bands to be tested we propose the following method to find this interval.

First, we estimate the overall zero centered confidence intervals using the following expression

$$\min_{h,q} \theta_{h,q} \quad (4.23)$$

where

$$\theta_{h,q} = \left(\frac{1}{2} \sum_{f=1}^F |\theta_{up}(f, q, h) + \theta_{lo}(f, q, h)| \right) \quad (4.24)$$

and where $\theta_{up}(f, q, h)$, $\theta_{lo}(f, q, h)$ correspond to the upper and lower limits of the 90% confidence interval for $\tilde{C}_q(f)$ for a given value of h .

Second, we verify that the length of the confidence interval associated with the values of h and q that minimize (4.23) is relatively small compared to the lengths of the confidence intervals associated with other values of h and q . Since there are eleven independent frequency sub-bands we estimate the smallest overall length of the confidence intervals with the following expression:

$$L\theta_{h,q} = \left(\frac{1}{2} \sum_{f=1}^F |\theta_{up}(f, q, h) - \theta_{lo}(f, q, h)| \right) \quad (4.25)$$

4.3.3 Test results

This section gives the results of the estimated 90% confidence interval for $\tilde{C}_q(f)$ corresponding to the transmit and receive covariance matrices of $[LO, SA]$ for which Kronecker was found to be a valid model.

We first evaluate the 90% confidence intervals for $\tilde{C}_q(f)$ statistic with *BCa* bootstrap method as indicated in the previous section. We then find the values of h and q that minimize (4.23) and (4.25) for which $\mathcal{H}_q$ is no longer rejected. In the following results $h_{dB} = 10 \log_{10}(h)$.

We have confirmed through empirical simulations that the asymptotic distribution in (4.21) when $q = 1$ is valid when the values of h_{dB} range from any desired small value up to a limit value of ≈ -11 .⁸ In the following results we evaluate the hypothesis $\mathcal{H}_q$, given in (4.12), with values of h_{dB} in the range of $[-35, -11]$ to indicate from non-significant to significant eigenvalues. We also evaluate the test for different number of smallest eigenvalues $q = [1, 2, 3]$ to indicate the degree of correlation in the elements of the MIMO channel.

⁸In general the asymptotic distribution is also valid for higher values of h_{dB} when $q = [2, 3]$

Figure 4.4: Finding the values of h and q that minimize $\theta_{h,q}$ to determine a zero-centered confidence interval for $\tilde{C}_q$

Figure 4.5: Finding the values of h and q that minimize $L\theta_{h,q}$ to determine a short-length confidence interval for $\tilde{C}_q$

Figure 4.4 shows $\theta_{h,q}$ vs. h_{dB} with different number of smallest eigenvalues q ; for the case of transmit and receive channel covariance matrices of *LO* and *SA* receive areas, respectively. We observe that the minimum of $\theta_{h,q}$ occurs at $h_{dB} \approx -11$ and $q = 1$ for the transmit and receive covariance matrices of *SA* and *LO*.

We also show in Figure 4.4 $L\theta_{h,q}$ vs. h_{dB} with different number of smallest eigenvalues q of the transmit and receive covariance matrices of *LO* and *SA* respectively. We observe that the values of h and q that minimize $\theta_{h,q}$ also minimize $L\theta_{h,q}$.

Figure 4.6 shows the estimated 90% confidence intervals for $\tilde{C}_q$ at each frequency sub-band i.e., $[\theta_{lo}(f), \theta_{up}(f)]$, for the case where the hypothesis cannot be rejected i.e., $h_{dB} \approx -11$ and $q = 1$ of the transmit and receive covariance matrices of *LO* and *SA*. We also show with dotted lines the standard deviations of the upper and the lower limits of these intervals. In general we see that the standard deviations of these limits are much smaller than the estimated limits, which indicate the good accuracy of the estimates.

From the previous results we obtain that we reject the hypothesis H_q for all values of h and q except for the case when $h_{dB} \approx -11$ and $q = 1$ for both the transmit and receive covariance matrices of *LO* and *SA*. To compare our results, we find h for the case of *iid* Rayleigh channels when $q = 1$. We generate the empirical distribution of h using $\mathbf{I}_4$ as the covariance matrix of the *iid* Rayleigh channels and find that the 90% confidence interval for h_{dB} is $\approx [-7, -6]$, which means that the estimated smallest eigenvalue of any of the covariance matrices of the measured channels is about 4 to 5 *dB* lower than that of the *iid* Rayleigh channels.

In the following section we show that this decrease in the smallest eigenvalue does not

Figure 4.6: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for $\tilde{C}_q(f)$ vs. frequency index, when $h_{dB} \approx -11$ and $q = 1$

cause a significant loss in capacity of the measured channels when compared to that of *iid* Rayleigh channels. As a result all the eigenvalues of the covariance matrices of the measured channels are significant, so that the elements of the MIMO measured channels corresponding to *LO* and *SA* can be considered spatially uncorrelated.

4.4 Discussion

In the following we illustrate with some examples the effect of different values in the smallest eigenvalue of the channel covariance matrix on the *cdf* of the channel capacity.

Figure 4.7 shows the *cdfs* of the channel capacities under equal power allocation for different cases of correlated 4×4 Rayleigh channels, at $SNR = 30$ dB. We express the correlation in the elements of the MIMO channel in terms of the eigenvalues of the transmit $\mathbf{R}_T$ and receive $\mathbf{R}_R$ channel covariance matrices i.e., $\text{eig}(\mathbf{R}_T) = \text{eig}(\mathbf{R}_R) = [\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4]$. In particular we illustrate the case when $\lambda_1 = \lambda_2 = \lambda_3$ and λ_4 is 0, 5, and 10 dB below any of the other eigenvalues.

We see that at $SNR = 30$ dB, the effect of correlation decreases channel capacity equally at any typical value of outage probability, this is reflected by the parallel left shift of the *cdf* of the capacity of the correlated channels with respect to the one of *iid* Rayleigh channels. We observe that when λ_4 is 5 and 10 dB below the other eigenvalues, there is respectively 5 and 10% decrease in channel capacity with respect to the capacity of *iid* Rayleigh channels.

This example is useful to illustrate the capacity loss of the measured channels with respect to that of *iid* Rayleigh channels due to spatial correlations. Given that the estimated smallest eigenvalue of any of the transmit and receive channel covariance matrices of [*LO*, *SA*]

is about 5 dB lower than that of *iid* Rayleigh channels and assuming the rest of eigenvalues have similar significance⁹ then we would expect about 5% loss in channel capacity with respect to that of *iid* Rayleigh channels. This indicates that there is no significant loss in channel capacity due to spatial correlations in the elements of the MIMO measured channels, so that we can consider that these channels are uncorrelated. High spectral efficiencies of outdoor to indoor measured channels due to richness in the multipath environment have also been reported by [Kermoal *et al.* (2002), Nguyen *et al.* (2004), Medbo *et al.* (2004)], where the authors show that the estimated capacities of the measured channels are at least 80% of those obtained by *iid* Rayleigh channels.

The reason the measured channels are uncorrelated is due to the large spacing of the antennas at the transmitter and receiver, which helps to make the measured MIMO channel elements less dependent. Also, the *NLOS* condition of the measured channels contributes to large angular spreads at the transmitter and at the receiver. Although we could not perform this analysis for the cases of the measured channels of $[DU, LI]$ due to the fact that these channels were affected by significant amounts of shadowing, we consider that the same results can be extended to these areas as well, since the mentioned channel conditions were the same for all measured receive areas.

Our results also indicate that the spatial correlations of the MIMO measured channels in *LO* and *SA* are similar. These results are in agreement with those found in [Nguyen *et al.* (2004)] where the authors show that the multipath richness of the measured outdoor to indoor channels is practically independent of the room where the receive unit is placed.

⁹We have confirmed the validity of this assumption based on a further evaluation of (4.12) for higher values of h when $q \in [2, 3]$. We found no significant differences between the estimated ratios h of the measured channels and those of the *iid* Rayleigh channels when $q \in [2, 3]$

Figure 4.7: *cdf* of the capacity of *iid* channels vs. that of correlated Rayleigh channels, assuming equal power allocation, $SNR = 30$ dB, and $f = 11$ frequencies

Chapter 5

MIMO single-user channel capacities

From the analysis results of the previous chapter we found that the measured MIMO channels for each receive area can be considered uncorrelated. This means that there would be no significant degradation in channel capacity at high SNR due to spatial correlations between the elements of the MIMO channel. In this chapter we analyze another important channel impairment that can significantly affect the performance of the measured channels per receive area. This channel impairment is related to large scale spatial variations in the receive power caused by:

- path loss differences due to different distances between the transmitter and the measured receive sub-areas of a building
- obstruction shadowing in a receive area

In this chapter we find estimates of the ergodic and the outage rates (spectral efficiencies) of the measured channels of a given receive area for single-user transmission with multiple transmit and receive antennas. We then compare the rates of these measured channels with the ones achieved by the theoretical *iid* Rayleigh channels and explain their differences based on the large scale variation parameters mentioned above.

The chapter is organized as follows: Section 5.1 explains the procedure we follow to estimate the maximum path loss difference within each of the measured receive areas (buildings). Section 5.2 describes the method we use to estimate the amount of shadowing in each of the measured receive areas. Section 5.3 gives a description of the mathematical channel model that we use to represent the wideband MIMO channels. It also introduces the mathematical formulation of the ergodic and the outage rates as well as the optimal power allocation algorithms for the case when transmitter knows the channel. Section 5.4 presents the method we use to estimate the ergodic and the outage rates of the measured channels under equal and optimal power allocations strategies. Section 5.5 compares the estimated capacities of the measured channels with respect to the capacities of the theoretical uncorrelated Rayleigh channels.

5.1 Estimated maximum path loss differences among the sub-areas of a building

Since our target is to estimate the single-user channel capacity of a given receive area, we want to determine how considerable are the path loss differences associated with the different measured sub-areas¹ of that receive area, since this may affect its ergodic and outage capacities.

Table 5.1 shows for each receive area the furthest (d_f) and the closest (d_c) distances from a measured sub-area to the transmitter. We observe that the maximum distance separation d_{dif} between sub-areas of a given building is 21 *m* and it occurs for *LI*. In the following two sections we estimate the maximum path loss difference among the measured sub-areas

¹We recall that each receive area is divided into three sub-areas, and each sub-area consists of five to six measured locations (see Section 3.1)

Table 5.1: Maximum distance separation $d_{dif} = d_f - d_c$ in m between two sub-areas of a given receive area

Receive area	$d_f(m)$	$d_c(m)$	$d_{dif}(m)$
DU	204	200	4
LI	281	260	21
LO	346	343	3
SA	365	350	15

of *LI*, since a considerable path loss difference could severely affect the performance of this receive area.

Note that the distance from the transmitter to *LI* is in between that from the transmitter to *DU* and that from the transmitter to *LO*. Also the measurements in these three buildings were all performed at the same relative floor height². We use these observations to estimate the maximum path loss difference among the sub-areas of *LI* using the estimated relative mean receive powers of *DU* and *LO*³.

Figure 5.1 shows the estimated 90% confidence interval for the relative mean receive powers of *DU* and *LO* (triangle markers) vs. the distance of each of these receive areas from the transmitter in a log-scale. The solid line connects the upper limit of the confidence interval for the mean receive power of *DU* and the lower limit of the confidence interval for the mean receive power of *LO*. We use this line to estimate the maximum path loss difference among the sub-areas of *LI* located at 260 m and 281 m (dashed vertical lines). We observe that there is a maximum path loss difference between these two sub-areas of about 2 dB . We will see later in this chapter that the effect of this path loss difference on the channel capacity of a receive area is not significant. As a result we can consider that there are no important

²Measurements in *SA* were performed two floors below the other receive areas as indicated in Section 3.1

³The estimated confidence intervals for the relative mean receive powers for each building are found in Section 6.1 of Chapter 6. We also recall that the estimated receive powers are all relative to some unknown reference as indicated in Section 2.3

Figure 5.1: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for r_{dB} of *DU* and *LO* (triangle markers) vs. distance in m in a *log* scale, used to estimate the path loss within *LI* (delimited by the vertical dashed lines)

differences in the path losses of the different locations of a measured receive area.

In the next section we analyze variations in the mean power experienced by the different measured locations of a receive area due to shadowing. The shadowing is caused by the presence of different interfering objects in the transmit-receive paths of the different measured locations.

5.2 Estimated shadowing losses for each building

In this section we describe the mathematical model to represent the variations in the mean power for a given receive area caused by shadowing.

At the l^{th} location (channel realization) the mean receive power of the channel with shadowing is given by

$$p(l) = r(l) s(l) \quad (5.1)$$

Goldsmith (1994); where $s(l)$ is the shadowing factor that is independent of $r(l)$ and is distributed according to $\mathcal{LN}(0, \sigma)$. We define $r(l)$ as the mean receive power of the Rayleigh fading channel at the l^{th} location given by

$$r(l) = (\gamma P)/(N_T N_R F) \sum_{i=1}^{N_T N_R} \sum_{f=1}^F |x_{i,f}(l)|^2 \quad (5.2)$$

where the average is performed over all F frequencies and $N_T N_R$ transfer functions of the MIMO channel; $x_{i,f}(l)$ is the fading channel gain that is distributed according to $\mathcal{CN}(0, 1)$; γ is the distance-dependent path loss that is considered to be constant for all locations of a given receive area and P is the transmitted power.

Alternatively we can express the mean receive power of the channel with shadowing, at the l^{th} location, on a log scale as

$$p_{dB}(l) = r_{dB}(l) + s_{dB}(l) \quad (5.3)$$

where $r_{dB}(l)$ and $s_{dB}(l)$ are respectively, the mean receive power of the Rayleigh fading channel and the shadowing factor in dB , at the l^{th} location. In (5.3) σ corresponds to standard deviation of $p_{dB}(l)$ around the mean $r_{dB} = E\{p_{dB}(l)\} = E\{r_{dB}(l)\}$.

Below, we describe the bootstrap method that we use to estimate the confidence interval for the parameter of interest.

Bootstrap confidence interval for σ for each receive area: we found in the previous section that there are no significant mean power variations in the different measured locations of a receive area caused by the different distance separations of these locations with respect to the transmitter. In this section we find that there may be significant variations in the mean powers of the different measured locations of a receive area caused by shadowing,

which we quantify by estimating the parameter σ for each receive area that characterizes (5.1) and (5.3). To estimate σ we use the bootstrap *BCa* method, introduced in Appendix E, and find the 90% confidence interval for this parameter.

For a given receive area we find the received power per location $p_{dB}(l)$ introduced in (5.3), which is written as

$$p_{dB}(l) = 10 \log_{10} \left(1/(N_T N_R F) \sum_{i=1}^{N_R} \sum_{j=1}^{N_T} \sum_{f=1}^F |H_{i,j,f}(l)|^2 \right)$$

where $F = 1000$ is the number of measured frequencies, $N_T N_R = 4 \times 4 = 16$ is the number of transfer functions and $H_{i,j,f}(l)$ are the elements of the transfer functions of the measured MIMO channel at the l^{th} location .

We then form the vector with the mean receive power for all locations of the given area $\mathbf{p}_{dB} = [p_{dB}(1), \dots, p_{dB}(l), \dots, p_{dB}(N)]$, where $N = 17$. Also let the estimate of the standard deviation of the shadowing factor s introduced in (5.1) be given by $\hat{\sigma} = \text{var}(\mathbf{p}_{dB})^{1/2}$, where $\text{var}(\cdot)$ is the unbiased estimator of the sample variance. We want to know how accurate this estimate is by finding the *BCa* confidence interval for σ as follows.

We denote the bootstrap sample by $p_{dB}^*(b)$, which corresponds to the b^{th} element of the data set $\mathbf{p}_{dB}$ where b is a uniform random index number in the range of $[1, \dots, N]$. The bootstrap realization: $\mathbf{p}_{dB}^*$ is a set of N bootstrap samples $p_{dB}^*(b)$ which are chosen by resampling uniformly with repetition the original data set $\mathbf{p}_{dB}$. The bootstrap estimate $\hat{\sigma}^* = \text{var}(\mathbf{p}_{dB}^*)^{1/2}$ is determined from each bootstrap realization. Finally we use $B = 10000$ bootstrap realizations to estimate the 90% confidence interval for σ using Equation (E.2) of Appendix E. Since this is an estimated interval we want to determine its accuracy by

Figure 5.2: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for σ (dB) (triangle markers), standard errors of the endpoints of this interval (+, × markers), and point estimate $\hat{\sigma}$ based on $N = 17$ (circle markers) for each receive area

evaluating the standard error of its endpoints using Equation (E.7) of Appendix E.

Figure 5.2 shows the point estimate (see Appendix E) $\hat{\sigma}$ (circle markers) for each of the receive areas based on $N = 17$ channel measurements. It also shows the 90% confidence interval for σ (triangle markers) along with the standard error for each of the endpoints of this interval (+, × markers). We observe in the figure that the estimated standard errors of these endpoints are small which shows that the estimated intervals are reasonably accurate.

In our measured channels we observe that shadowing was more significant when the measured receive area was large such as for the case of *LI* ($\sigma \in [3.6, 5.2]$ dB). For this receive area measurements were performed over a distance of ≈ 45 m. This contributed to the presence of shadowing caused by different interfering buildings in the transmit-receive paths of different sub-areas. In particular we observe that *DU* blocks the paths coming from the

transmitter to one of the sub-areas of *LI*. We think that *DU* most likely causes a diffraction loss that is reflected by the mean receive power of that sub-area being significantly less than those of the other two sub-areas of *LI* (see Figure 3.1 in Chapter 3).

For the other receive areas shadowing was less important. Nonetheless we make the following observation. We note that the closer the surrounding buildings were to the receive area the larger was the shadowing. We observe for instance that the effect of shadowing is the second largest for *DU* ($\sigma \in [1.6, 2.7]$ dB). This receive area is surrounded by Azrieli Pavilion and Tory buildings at relatively close distances of ≈ 25 m. The effect of shadowing decreases for *SA* ($\sigma \in [1.5, 2.5]$ dB) which was surrounded by *LI* and *LO* buildings at moderate distances of ≈ 35 m. Finally the effect of shadowing is the lowest for *LO* ($\sigma \in [1.4, 2.0]$ dB) which was surrounded by Paterson Hall and Tory buildings at relatively large distances of ≈ 50 m.

We emphasize that for the case of *LI* the confidence interval for σ is about $[1 - 3]$ dB higher than those for the other receive areas. We will show at the end of this chapter, the detrimental impact of shadowing on the ergodic and the outage rates of *LI*.

Before continuing with the presentation of the single-user channel and its associated ergodic and outage rates, we describe below the normalization procedure that we use with the measured single-user channels in order to find their estimated ergodic and outage rates given at the end of this chapter.

Normalization of measured channels: The measured transfer functions from different receive areas are scaled by different factors. More specifically, the measured transfer functions

of a given receive area are equally scaled and must satisfy

$$\frac{1}{FL} \sum_f \sum_l \text{Tr}(\mathbf{H}_f(l) \mathbf{H}_f^\dagger(l)) = N_T N_R \quad (5.4)$$

where $\mathbf{H}_f(l)$ are the measured MIMO channels with $N_T N_R$ transfer functions. The average is taken over all F frequencies and over all L locations of the given receive area. This normalization procedure preserves the mean receive power variations of the different measured locations of a receive area.

5.3 The MIMO single-user channel model and its ergodic and outage rates

We use the block fading model as described in [Biglieri *et al.* (2001)] where the channel is considered to be fixed within each block of length K symbols, and where the channel varies independently from block to block (F fading blocks). In our case we consider that each block is given by each of the independent frequency sub-bands of the MIMO channel that we explain below.

At the frequency sub-band $f \in [1, \dots, F]$ and at the symbol index $k \in [1, \dots, K]$, let the output of the channel be given by

$$\mathbf{y}_f(k) = \mathbf{H}_f \mathbf{x}_f(k) + \mathbf{n}_f(k) \quad (5.5)$$

where $\mathbf{y}_f(k)$ is the $N_R \times 1$ output vector; $\mathbf{H}_f$ is the $N_R \times N_T$ MIMO channel matrix; $\mathbf{x}_f(k)$ is the $N_T \times 1$ input vector and $\mathbf{n}_f(k)$ is $N_R \times 1$ noise vector distributed according to $\mathcal{CN}_{N_R}(\mathbf{0}, \mathbf{I}_{N_R})$.

We can find an equivalent representation to (5.5) in terms of the *SVD* of $\mathbf{H}_f$ as follows.

Let

$$\mathbf{H}_f = \mathbf{U}_f \mathbf{\Lambda}_f^{1/2} \mathbf{V}_f^\dagger \quad (5.6)$$

where $\mathbf{U}_f$ of size $N_R \times N_R$ and $\mathbf{V}_f$ of size $N_T \times N_T$ are unitary matrices, and $\mathbf{\Lambda}_f^{1/2}$ of size $N_R \times N_T$ is a diagonal matrix. If the m non-zero ordered eigenvalues of $\mathbf{H}_f \mathbf{H}_f^\dagger$ are denoted by $\lambda_{f,1} \geq \dots \geq \lambda_{f,m}$ then the main diagonal entries of $\mathbf{\Lambda}_f^{1/2}$ are given by the singular values $\lambda_{f,1}^{1/2} \geq \dots \geq \lambda_{f,m}^{1/2}$.

Also, let $\tilde{\mathbf{y}}_f(k) = \mathbf{U}_f^\dagger \mathbf{y}_f(k)$; $\tilde{\mathbf{x}}_f(k) = \mathbf{V}_f^\dagger \mathbf{x}_f(k)$; and $\tilde{\mathbf{n}}_f(k) = \mathbf{U}_f^\dagger \mathbf{n}_f(k)$ where $\tilde{\mathbf{n}}_f(k)$ is distributed according to $\mathcal{CN}_{N_R}(\mathbf{0}, \mathbf{I}_{N_R})$, then we can express (5.5) as

$$\tilde{\mathbf{y}}_f(k) = \mathbf{\Lambda}_f^{1/2} \tilde{\mathbf{x}}_f(k) + \tilde{\mathbf{n}}_f(k) \quad (5.7)$$

We can also write (5.7) component wise by: $\tilde{y}_{f,i}(k) = \lambda_{f,i}^{1/2} \tilde{x}_{f,i}(k) + \tilde{n}_{f,i}(k)$ when $i \in [1, \dots, m]$, and $\tilde{y}_{f,i}(k) = \tilde{n}_{f,i}(k)$ when $i \in [m+1, \dots, N_R]$. We observe that each of the first m components can be regarded as a spatial sub-channel attenuated by the corresponding singular value and corrupted by noise.

In the following, we consider that each channel realization consists of a set of F block fading channel matrices $\{\mathbf{H}_f\}_{f=1}^F$, i.e., each channel realization corresponds to a set of $m \times F$ sub-channels (in space and frequency). We now give some definitions necessary to find the ergodic and outage rates of the MIMO channel in (5.5).

Let $\mathbf{\Lambda} = [\lambda_1^T, \dots, \lambda_f^T, \dots, \lambda_F^T]$ be the $m \times F$ matrix whose elements are given by $\lambda_f = [\lambda_{f,1}, \dots, \lambda_{f,m}]$. Also, let $\mathbf{P} = [\mathbf{p}_1^T, \dots, \mathbf{p}_f^T, \dots, \mathbf{p}_F^T]$ be the $N_T \times F$ matrix of transmit powers for each frequency, where $\mathbf{p}_f = [p_{f,1}, \dots, p_{f,m}, \underbrace{0, \dots, 0}_{N_T-m}]$ is the $1 \times N_T$ vector whose elements correspond to the main diagonal of $E\{\mathbf{x}_f(k)\mathbf{x}_f^\dagger(k)\}$.

Then the maximum instantaneous mutual information in [Biglieri *et al.* (2001)], is given by

$$I_F(\mathbf{\Lambda}, \mathbf{P}) = \frac{1}{F} \sum_{f=1}^F \sum_{i=1}^m \log(1 + \lambda_{f,i} p_{f,i}) \quad (5.8)$$

where $p_{f,i}$ represents the power transmitted by the i^{th} antenna corresponding to the f^{th} frequency sub-band and its value depends on the input power constraints that we explain below.

Let the input short-term power constraint be given by

$$P_F(\mathbf{P}) = \frac{1}{F} \sum_{f=1}^F \sum_{i=1}^m p_{f,i} \leq P \quad (5.9)$$

this equation represents the instantaneous transmit power constraint, i.e. at each channel realization. Similarly, the long-term power constraint can be written as

$$E\{P_F(\mathbf{P})\} \leq P \quad (5.10)$$

and represents the average transmit power constraint i.e., averaged over all channel realizations. We note that the assumption of long-term input power constraint is not very realistic in practice, however it is very useful in interpreting physical results obtained as shown later in Section 5.3.2.

In the following expressions, the expected values are understood to be taken over all channel realizations.

5.3.1 Ergodic rates

The ergodic capacity when only the receiver knows the channel is given by

$$C^{eq} = E \left\{ \frac{1}{F} \sum_{f=1}^F \sum_{i=1}^m \log \left(1 + \frac{P}{N_T} \lambda_{f,i} \right) \right\} \quad (5.11)$$

We observe that the allocated power is the same for all frequencies, all transmit antennas and for all channel realizations [Paulraj *et al.* (2004)]. We note that, after normalizing the channel using (5.4), P in (5.11) also corresponds to the average SNR per receive antenna as presented in [Foschini and Gans (1998)] since the noise in (5.5) is assumed to have unit power spectral density.

The ergodic channel capacity when the transmitter and also the receiver know the channel under long-term input power constraint is given by

$$C = E \left\{ \frac{1}{F} \sum_{f=1}^F \sum_{i=1}^m \log(\mu - 1/\lambda_{f,i})_+ \right\} \quad (5.12)$$

where the water level μ is the solution to the long-term power constraint

$$E \left\{ \frac{1}{F} \sum_{f=1}^F \sum_{i=1}^m \left(\mu - \frac{1}{\lambda_{i,f}} \right)_+ \right\} = P \quad (5.13)$$

[Biglieri *et al.* (2001)]; where $(a)_+ = \max(a, 0)$. For the case of a short-term input power constraint, μ is the solution to

$$\frac{1}{F} \sum_{f=1}^F \sum_{i=1}^m \left(\mu - \frac{1}{\lambda_{i,f}} \right)_+ = P \quad (5.14)$$

In particular we note that under long-term input power constraint this solution corresponds to water-filling the power over all frequency and spatial sub-channels and over all channel

realizations. Under a short-term input power constraint this solution corresponds to water-filling the power over all frequency and spatial sub-channels of each channel realization.

5.3.2 Outage rates

By definition the outage probability is given by

$$P_{out}(R, P) = \text{Prob}(I_F(\mathbf{\Lambda}, \mathbf{P}) < R) \quad (5.15)$$

Assuming that the transmitter and the receiver both know the channel and that there is a short-term input power constraint we find that the outage capacity is related to the outage probability by

$$C_\epsilon = \sup_{\mathcal{P}(\mathbf{\Lambda})} \sup_R \{R : P_{out}(R, P) \leq \epsilon\} \quad (5.16)$$

where $\mathcal{P}(\mathbf{\Lambda})$ is the transmit power allocation function and C_ϵ is the maximum rate R with probability of outage less than or equal to ϵ when using an optimal power allocation strategy on the input such that the instantaneous transmit power is less than or equal to P . The outage capacity under long-term input power constraint can similarly be formulated.

When only receiver knows the channel, it is conjectured in [Telatar (1995)] that the optimal strategy is to allocate equal power to all the transmit antennas and to all the frequencies such as the total transmit power is less than or equal to P .

In the following we describe how to obtain the optimal power allocation first under short-term and then under long-term input power constraints when the transmitter and the receiver both know the channel. The authors in [Biglieri *et al.* (2001)] approach the solution by minimizing the outage probability with respect to the choice of transmit power allocation function $\mathcal{P}(\mathbf{\Lambda})$ assuming a fixed rate R transmission scheme. In their study they also assume that the block

fading channel matrices $\{\mathbf{H}_f\}_{f=1}^F$ in (5.5) are independent.

Short-term input power constraint minimization problem: the minimum outage probability optimization problem under short-term input power constraints is formulated as

$$\min_{\mathcal{P}(\mathbf{\Lambda})} \text{Prob}(I_F(\mathbf{\Lambda}, \mathbf{P}) < R) \quad (5.17)$$

$$\text{subject to:} \quad P_F(\mathbf{P}) \leq P \quad (5.18)$$

and the solution is presented in the following two statements [Biglieri *et al.* (2001)]

- (1) Let the optimal allocated power for the short-term constraint be given by

$$\mathcal{P}(\mathbf{\Lambda}) = \begin{cases} \mathcal{P}^{st}(\mathbf{\Lambda}, P) & \text{if } \mathbf{\Lambda} \in R_{on} \\ 0 & \text{if } \mathbf{\Lambda} \in R_{off} \end{cases} \quad (5.19)$$

where the power off region is given by the following set of channels

$$R_{off}(R, P) = \{\mathbf{\Lambda} : I_F(\mathbf{\Lambda}, \mathcal{P}^{st}(\mathbf{\Lambda}, P)) < R\} \quad (5.20)$$

and the power on region R_{on} is the complement of the R_{off} region.

- (2) $\mathcal{P}^{st}(\mathbf{\Lambda}, P)$ corresponds to power allocation that maximizes the instantaneous mutual information given an input power constraint, which can be found by solving the following maximization problem

$$\max_{\mathcal{P}^{st}} I_F(\mathbf{\Lambda}, \mathcal{P}^{st}) \quad (5.21)$$

$$\text{subject to:} \quad P_F(\mathcal{P}^{st}) \leq P \quad (5.22)$$

The solution to this problem is obtained by the water-filling algorithm, where the

component $p_{f,i}$ of $\mathcal{P}^{st}(\Lambda, P)$ is found using (5.13) with the water level μ given by

$$\mu = \frac{1}{\#(\Omega)} \sum_{(f,i) \in \Omega} \frac{1}{p_{f,i}} + \frac{F P}{\#(\Omega)} \quad (5.23)$$

where $\#(\mathbb{A})$ denotes the cardinality of the set $\mathbb{A}$, and Ω is the set of indexes (f, i) for which $p_{f,i}$ in (5.13) is positive.

In this solution when the channel is bad we do not transmit since an outage is unavoidable given the instantaneous power constraint, otherwise when the channel is good we transmit the total power that delivers the maximum instantaneous mutual information.

Given a target rate R and an instantaneous input power constraint P , this solution corresponds in our case to water-filling the total power over all frequency and spatial sub-channels for each channel realization when the channel is good.

Long-term input power constraint minimization problem: the minimum outage optimization problem under long-term input power constraints is formulated as

$$\min_{\mathcal{P}(\Lambda)} \text{Prob}(I_F(\Lambda, \mathbf{P}) < R) \quad (5.24)$$

$$\text{subject to:} \quad E\{P_F(\mathbf{P})\} \leq P \quad (5.25)$$

and the solution is presented in the following statements [Biglieri *et al.* (2001)]

- (1) Let the optimal allocated power for the long-term constraint be given by

$$\mathcal{P}(\Lambda) = \begin{cases} \mathcal{P}^{lt}(\Lambda, R) & \text{if } \Lambda \in R_{on} \\ 0 & \text{if } \Lambda \in R_{off} \end{cases} \quad (5.26)$$

where the power off region is given by the following set of channels

$$R_{off}(R, P) = \{\Lambda : P_F(\mathcal{P}^{lt}(\Lambda, R) > P^*)\} \quad (5.27)$$

and the power on region R_{on} may be interpreted as the set of good channels and it is the complement of the R_{off} region.

(2) The threshold $P^* > 0$ is the solution to

$$E\{P_F(\mathcal{P}^{lt}(\Lambda, R)) \mathbf{1}\{\Lambda \in R_{on}\}\} = P \quad (5.28)$$

where $\mathbf{1}\{\mathcal{A}\}$ is one if $\mathcal{A}$ is true, otherwise is zero and $\mathcal{P}^{lt}$ is the solution to the following optimization problem

$$\min_{\mathcal{P}^{lt}} P_F(\mathcal{P}^{lt}) \quad (5.29)$$

$$\text{subject to: } I_F(\Lambda, \mathcal{P}^{lt}) \geq R \quad (5.30)$$

The optimal solution of this problem corresponds to the water-filling solution, where the component $p_{f,i}$ of $\mathcal{P}^{lt}$ is found using (5.13) with the water level μ given by

$$\mu = \left(\frac{2^{N_R}}{\prod_{(f,i) \in \Omega} \lambda_{f,i}} \right)^{1/\sharp(\Omega)} \quad (5.31)$$

In this solution we transmit only when the channel is good and turn off the transmission when the channel is bad. The allocated power $\mathcal{P}^{lt}$ when the channel is good corresponds to the minimum power required to avoid an outage, which saves power for future transmissions.

Given a target rate R and an average power constraint P , this solution corresponds in our case to water-filling over all frequency and spatial sub-channels and over all channel

realizations; with the characteristic that the water level for each channel realization varies in order to achieve the targeted rate R .

We implement the algorithms that obtain the solutions mentioned above in Appendix I. We use those algorithms to estimate the outage capacities under optimal power allocation of the measured channels in the following section.

5.4 Estimated ergodic and outage rates of the measured MIMO single-user channels

In the previous chapter related to the study of the spatial correlations between the elements of the measured 4×4 MIMO channels, it was necessary to extract the independent frequency sub-bands of each of the sixteen measured transfer functions of the MIMO channel in order to perform the statistical analysis. There, we estimated that there are at least eleven independent frequency sub-bands in each transfer function. We selected those eleven frequency sub-bands from each transfer function to represent the wideband measured channels and performed our analysis.

In this study we use the same estimated independent frequency sub-bands. This step is required since the optimization procedure for the case of optimal outage rates under transmitter channel knowledge is based on the assumption of independent frequencies. Also, analyzing only the estimated eleven independent frequency sub-bands considerably reduces the processing time of the optimization routines used to find the optimal rates when the transmitter knows the channel.

We note that due to the limited number of measurements $N = 17$ as well as independent frequency sub-bands $F = 11$ there is an error associated with our point rate estimates. To see how accurate are these point estimates, we use the bootstrap *BCa* method to obtain the 90% confidence intervals for these rates as shown in Figures J.1 and J.2 of Appendix J.

In the subsequent analysis we use the rate estimates to compare the capacity achievable when using optimal and equal power allocation strategies; and we use the estimated confidence intervals for these rates to compare the capacity achievable on the measured channels and that achievable on theoretical channels under an equal power allocation strategy. Furthermore our analysis is focused on cases of high *SNR* where there is less variability of the estimated rates with respect to their expected values.

We have also confirmed in Figures J.3 and J.4 of Appendix J that the estimated eleven frequency sub-bands are a good representation of the measured wideband channels so that we can reliably base our analysis on these estimated frequency sub-bands.

In the following sections we present predictions of the capacity that would be expected on the measured channels in terms of ergodic and outage rates. In all the figures, *SNR* represents the average receive signal to noise ratio per receive antenna as described in Section 5.3.1. The ergodic and the outage rates correspond to the case where the transmission process takes place while the receiver moves through all measured locations in a building. Moreover, we find the rates under long-term power constraints assuming that we know all the channel realizations, so that power is water-filled over space frequency and channel realizations according to (5.13) and (5.31); we find the rates under short-term power constraints assuming that we only know one channel realization at a time, so that power is water-filled over space and frequency of that channel realization according to (5.14) and (5.23).

Figure 5.3: Point estimates of the ergodic rates $\hat{C}_{erg}$ under different power allocation strategies based on eleven independent frequencies and seventeen channel measurements

5.4.1 Ergodic rates

In this section we estimate the ergodic rates for the case when only the receiver knows the channel given by (5.11) and for the case when both the transmitter and the receiver know the channel given by (5.12). In the latter case we find the optimal power allocations under a short-term input power constraint and under a long-term input power constraint using the water-filling solution given in (5.13) to satisfy the corresponding input power constraints described by (5.9) and (5.10).

Figure 5.3 shows the point estimates of the ergodic rates in $b/s/Hz$ vs. SNR dB for the case of DU , LI , LO and SA respectively. We observe that the optimal rates under short-term and under long-term input power constraints are basically the same. There is only benefit at small SNR of having channel knowledge at the transmitter. However, this benefit is not significant since equal power allocation delivers nearly the same ergodic capacity. In this figure we use a reference dash-dotted line which represents the ergodic water-filling capacity of 4×4 iid Rayleigh channels under long-term input power constraints.

We also compare the ergodic capacity under equal power allocation of the measured channels with the one of iid Rayleigh channels. From the estimated 90% confidence interval for the ergodic rates under equal power allocation of the measured channels presented in Figure J.1 of Appendix J, we observe that at SNR = 30 dB there is a $[0 - 8]\%$ decrease in the ergodic rates of $[DU, LO, SA]$ and a $[5 - 17]\%$ decrease in the ergodic rate of LI with respect to the ergodic rates of iid Rayleigh channels given in Figure 5.4 (a). These percentage values are obtained as follows. We see that the 90% confidence intervals for the ergodic rates of $[DU, LI, LO, SA]$ are respectively $[32 - 34.5, 29 - 33, 32 - 33.5, 32.5 - 35]$ $b/s/Hz$ that correspond respectively to $[92 - 99, 83 - 95, 92 - 96, 93 - 100]\%$ of the ergodic rates of iid Rayleigh channels (34.9 $b/s/Hz$), so that the rates of $[DU, LO, SA]$ achieve $[0 - 8]\%$

Figure 5.4: *iid* Rayleigh channels with $F = 11$ independent frequencies

less than the ergodic rates of *iid* Rayleigh channels and *LI* achieves $[5 - 17]\%$ less than the ergodic rates of *iid* Rayleigh channels.

The estimated ergodic rates presented above can be achieved for the case of channels with no decoding delay constraints and represent the maximum achievable rates. In the next section we estimate the outage rates of the measured channels, which correspond to the case when there are decoding delay constraints.

5.4.2 Outage rates

The outage rate estimates that we present in this section are for the case of no channel knowledge at the transmitter where we use (5.16) under equal power allocation and also for the case of channel knowledge at both the transmitter and the receiver where we use the optimal power allocations indicated by the algorithms presented in Appendix I to find the outage rates under short-term and long-term input power constraints.

Figure 5.5 shows the point estimates of the outage rates at $(P_{out} = 6\%)^4$ for $[DU, LI, LO, SA]$ receive areas respectively. The dashed line denotes the ergodic capacity of the given receive area under optimal long-term power allocation (using water-filling) and represents its maximum achievable rate. We first note a major difference with respect to optimal power allocation of the ergodic rates. Although there is no benefit in optimal power allocation under short-term input power constraints and equal power allocation, there is a significant improvement when allocating the power optimally under long-term input power constraints, at $SNR = 30$ dB.

More specifically, based on the point estimates of the outage rates under long-term input power constraints, we observe that at $SNR = 30$ dB the outage rates of $[DU, LO, SA]$ achieve $[98 - 100]\%$ of their respective ergodic rates under optimal long-term power allocation while the outage rate of LI achieves 91% of its ergodic rate under optimal long-term power allocation. These percentage values are obtained from the point estimates of the optimal outage rates under optimal long-term power allocation of $[DU, LI, LO, SA]$ given by $[32.5, 28.4, 31.9, 33.5]$ b/s/Hz, and the point estimates of the ergodic rates under optimal long-term power allocation of $[DU, LI, LO, SA]$ given by $[32.8, 31.3, 32.5, 33.6]$ b/s/Hz.

We find that the outage rate under optimal long-term power allocation of LI (28.4 b/s/Hz) represents an improvement of $[7 - 23]\%$ over the estimated interval for its outage rate under equal power allocation given in Figure J.2 i.e., $[22 - 26.5]$ b/s/Hz. To put things in perspective, if the channels were *iid* Rayleigh we would only expect a 3% improvement from long-term optimal power allocation on the outage rates (34.9 b/s/Hz) with respect to the

⁴We have selected an arbitrary low outage probability value to analyze the outage capacity for illustration purposes (in our case this outage probability value means that one out of the seventeen measured channel realizations is in outage). However, when discussing our results we will generalize our observations to analyze the *cdf* of the channel capacity.

Figure 5.5: Point estimates of the outage rates $\hat{C}_{out}$ with $P_{out} = 6\%$ under different power allocation strategies based on eleven independent frequencies and seventeen channel measurements

respective outage rates with equal power allocation (33.8 b/s/Hz) as indicated by the simulated *iid* Rayleigh channels in Figure 5.4 (b).

We also note in Figure 5.5 that under optimal long-term power allocation the point estimates of the outage rates of all the receive areas are similar to the point estimates of their ergodic rates except for *LI*. We recall that a large number of the locations within *LI* were affected by the shadowing of *DU* building so that there are several channel realizations where the received power is low and which are consuming most of the available power in order to maintain a given fixed rate.

We now compare how significant is the decrease, under equal power allocation, of the outage rates of [*DU*, *LI*, *LO*, *SA*] given by $[28.5 - 31.5, 22 - 26.5, 28 - 31, 26.5 - 30.5] \text{ b/s/Hz}$ with respect to their respective ergodic rates given by $[32 - 34.5, 29 - 33, 32 - 33.5, 32.5 - 35] \text{ b/s/Hz}$. We base our comparison on the 90% confidence intervals for the ergodic rates given in Figure J.1, and those for the outage rates given in Figure J.2 of Appendix J. We observe that at $SNR = 30 \text{ dB}$ the outage rates of [*DU*, *LI*, *LO*, *SA*] achieve respectively, $[83 - 98]\%$, $[67 - 91]\%$, $[83 - 97]\%$ and $[76 - 94]\%$ of their corresponding ergodic rates. We see that for the case of *LI* there is a considerable decrease in the outage rates of $[9 - 33]\%$ when compared to the expected 3% decrease if the channels were *iid* Rayleigh ($33.8/34.9$), see Figure 5.4 (a).

From this analysis we see that especially for the case of *LI* there is a clear indication of a decrease of the ergodic and the outage rates of the measured channels with respect to *iid* Rayleigh channels. Furthermore, the considerable improvement in the outage rate when allocating the power optimally under long-term input power constraints (transmissions are performed only at favorable channel realizations) suggests that the receive power significantly

varies on each channel realization. In the following we illustrate with two examples how the rates (using equal power allocation) of uncorrelated channels are affected when there are variations in the received power due to path loss differences and shadowing similar to the ones experienced by the measured receive area of LI .

Effect of channels with path loss differences ≤ 2 dB on the *cdf* of the channel capacity:

We illustrate with an example the effect of channels with path loss differences on the *cdf* of the channel capacity under equal power allocation. We use a path loss difference of 2 dB which is the maximum value we estimated for LI (as described in Section 5.1). To resemble the measured scenarios, we consider that there are three sub-areas (where each sub-area consists of a set of measured locations or channel realizations) and that the path loss difference between two consecutive sub-areas is 1 dB, so that there is a maximum difference of 2 dB between the mean received powers of two sub-areas (where the average is taken over all the locations of a given sub-area).

Figure 5.6 (a) depicts the *cdf* of the capacity under equal power allocation of a 4×4 Rayleigh channel with eleven independent frequencies for the case where the three sub-areas have the same path loss (solid line) and for the case where the maximum path loss difference between sub-areas is 2 dB as described above (dashed line). We observe that at $SNR = 30$ dB, the ergodic rates of both channels are the same $\approx (34.9 \text{ b/s/Hz})$ and that the outage rate with $P_{out} = 6\%$ of the channel with path loss differences (32.8 b/s/Hz) achieve 97% of the capacity of the channel without path loss differences (33.8 b/s/Hz). Since these differences are reasonably small, we can consider that the estimated path loss differences within a building of the measured channels ≤ 2 dB are not significant, since the capacity of these channels is similar to that of the *iid* Rayleigh channels.

Figure 5.6: *iid* Rayleigh channels with $F = 11$ independent frequencies affected by different physical link factors

Effects of channels with shadowing ($\sigma = 3$) dB on the *cdf* of the channel capacity:

Figure 5.6 (b) shows the *cdf* of the capacity under equal power allocation of a 4×4 *iid* Rayleigh channels with $F = 11$ frequencies as well as the *cdf* of the capacity of the same channel with a shadowing factor s , as described in (5.1), where s is $\mathcal{LN}(0, 3)$ (similar to the minimum value found for LI in Section 5.2). As we can see, at $SNR = 30$ dB, the ergodic rates of the channels with shadowing (33.6 b/s/Hz) are about 4% less than the ergodic rates of the *iid* Rayleigh channels (34.9 b/s/Hz). This difference becomes more significant at lower outage probabilities where we see for instance that with $P_{out} = 6\%$ the outage rates of the channels with shadowing (27.3 b/s/Hz) are about 19% less than its ergodic rates (33.8 b/s/Hz). This is shown in this figure by the increase in the tail of the *cdf* of the capacity of the channels with shadowing with respect to that of the *iid* Rayleigh channels.

We observe that the simulated channels with shadowing and the measured channels of *LI* have similar capacities. To see this more clearly we depict in Figure 5.7 (a) the ergodic and

Figure 5.7: Uncorrelated channels with shadowing, s is $\mathcal{LN}(0, 3)$ and $F = 11$ independent frequencies

the outage rates, under equal power allocation, of the simulated channels with shadowing mentioned above as well as the 90% confidence intervals for the ergodic and the outage rates, under equal power allocation, of *LI* (dotted lines). The capacity curves of both channels have the same slope; also the parallel shift above by the curves of the simulated channels can be explained by the smaller value of standard deviation of the shadowing factor $\sigma = 3$ dB than the estimated one of *LI*, $\sigma \in [3.6, 5.2]$ dB.

Figure 5.7 (a) also shows that under equal power allocation the outage rates of the simulated channels with shadowing achieve $\approx 81\%$ of its ergodic rates. Figure 5.7 (b) shows that the outage rates under long-term optimal power allocation of the simulated channels with shadowing (32.5 b/s/Hz) achieve $\approx 96\%$ of its ergodic rates under equal power allocation (33.8 b/s/Hz). These results are in agreement with what we found in the measured channels and show the significant improvement in the outage rate when allocating the power optimally under long-term input power constraints. The significance of this improvement will depend on the degree of variation of the received SNR which in our case is caused by shadowing.

5.5 Discussion

Our results indicate that the significant decrease in the outage rate of LI is due to shadowing and not to path loss differences within this building. The estimated maximum path loss difference within a building of < 2 dB does not significantly affect the channel capacity at high SNR .

Previous studies on channel capacity of MIMO measured channels presented in [Bonek *et al.* (2003), Nguyen *et al.* (2004), Bonek (2006)], have also found that the variation of the mean receive power considerably affects the ergodic channel capacity. Channel capacity is practically unaffected by the variation in the multipath richness properties of the measured channels.

There are two main differences between our analysis and the previous studies that is explained as follows. First, we are able to specifically identify that shadowing is causing the variation of the mean receive power for each location and not the separation distance between transmitter and receive location of the measured single-user channels. We will see later that path loss differences are more important in the context of multi-user channels and that they will affect the sum capacity of the multi-user channels differently⁵. Second, we are able to establish the effect of shadowing not only in terms of ergodic capacity but also in terms of outage capacity of the measured single-user channels, which significantly varies for these two performance metrics.

To see how different degrees of shadowing affect the channel capacity under equal power allocation we present in Figure 5.8 the ergodic and the outage rates for uncorrelated channels with different values of σ . Based on this figure, at $SNR = 30$, we see that for the receive areas of $[DU, LO, SA]$, where the estimated standard deviation of the shadowing

⁵Section 6.6 of Chapter 6 shows the specific effect on the *cdf* of the sum capacity of having path loss differences among the largely spaced users of the multi-user channel

Figure 5.8: Uncorrelated channels with different degrees of shadowing; s is $\mathcal{LN}(0, \sigma)$, $F = 11$ and equal power allocation

factor is $\sigma \in [\approx 1.5, \approx 2.5] \text{ dB}$, we would expect a $[10 - 15]\%$ decrease in the outage rates ($[31.3, 28.8] \text{ b/s/Hz}$) with respect to their ergodic rates ($[34.6, 34] \text{ b/s/Hz}$) and a $[1 - 3]\%$ decrease in the ergodic rates with respect to the ergodic rates of *iid* Rayleigh channels.

For the receive area of LI , where $\sigma \in [\approx 3.5, \approx 5.5] \text{ dB}$, the corresponding percentages would be $[21 - 36]\%$ and $[5 - 13]\%$ respectively; based on the outage rates $[26.1, 19.5] \text{ b/s/Hz}$ and ergodic rates $[33.1, 30.4] \text{ b/s/Hz}$ shown in the same figure.

Chapter 6

MIMO downlink channel capacities

In the previous two chapters we analyzed the capacities of the measured MIMO single-user channels, where the base station (BS), equipped with multiple antennas, located on the rooftop of a building, transmits to one user terminal, also equipped with multiple antennas, located inside another building. We learned from our results of Chapter 4 that at high SNR there would be practically no degradation in channel capacity due to spatial correlations in the elements of the measured MIMO single-user channels. We also found in Chapter 5 that the measured channels were instead affected by shadowing, which would cause variations in the mean received power of the different sub-areas of a building. In particular we found that when the amount of shadowing is significant, such as for LI , there is a considerable degradation in performance, particularly in terms of outage rates under equal power allocation.

In this chapter we consider measured MIMO multi-user channels where BS and user terminals are similarly equipped as before. We are particularly interested to study the downlink channels in which the objective is to deliver the most balanced rates to all users¹. To this end we compare the relative performance in terms of ergodic and outage rates of the following

¹In this chapter we use the terms 'user', 'receive area' and 'building' interchangeably. This implies that there is no more than one user in each building occupying a given channel.

transmit-receive multi-user schemes:

- *OSDM* (orthogonal space division multiplexing)
- *TDM* (time division multiplexing)

When the *OSDM* schemes are employed, the *BS* uses all the system bandwidth to simultaneously transmit to all users located in different receive areas. This is achieved by projecting the user signals onto orthogonal sub-channels so that users are separated according to their spatial signatures. When the *TDM* scheme is employed, the *BS* sends information to each user in a different time slot, so that the multiple data MIMO sub-channels may be transmitted to each user terminal without interference to other user terminals. In the work to follow, we assume that the receiver knows the channel to each of its receiving antennas. We also assume that the transmitter knows the channel in the *OSDM* schemes and that the transmitter has limited channel knowledge in the *TDM* scheme, so that it can adjust the average transmission time allocation per user.

For this type of measured multi-user channels we predict a significant loss in channel capacity due to the presence of variations in the mean received power per user due to path loss differences between largely separated users.

This chapter is organized as follows: In Section 6.1, we estimate the path loss differences associated with the distance separations between the transmitter and each of the four users (buildings). We will use these estimates at the end of the chapter to explain the rates of the measured multi-user channels. In Section 6.2, we describe the fixed Gaussian channel model and two of the *OSDM* schemes presented by the authors in [Spencer *et al.* (2004)]. In Section 6.3, we extend these *OSDM* schemes to the case of fading frequency selective Gaussian channels. In Section 6.4, we propose an optimization criterion to find a set of

balanced user rates of the measured multi-user fading channels. In Section 6.5, we show a comparison of the ergodic and the outage rate point estimates of the measured multi-user channels operating with the *TDM* and the *OSDM* schemes. Finally in Section 6.6, we illustrate with an example the effect of path loss differences in the multi-user channel on the *cdf* of the sum channel capacity (assuming users can cooperate). We also present a simplified example to illustrate the effect of path loss differences and shadowing on the *cdfs* of the channel capacities when employing the *OSDM* and the *TDM* schemes.

6.1 Estimated mean receive power for each building

In this section we estimate the relative mean receive power for each user², $r_{dB} = E\{p_{dB}(l)\}$, where $p_{dB}(l)$ is the mean received power at the l^{th} location in a log scale, defined in (5.3). We describe below the bootstrap method to estimate the confidence interval for the parameter of interest.

Bootstrap confidence interval for r_{dB} at each receive area: we follow a similar bootstrap procedure as that described in Section 5.2 where we found the confidence interval for σ . We first evaluate the relative receive power $p(l)$, in a linear scale, for each location l using (5.1). We then obtain the corresponding relative mean receive powers per location $p_{dB}(l)$, in a log scale, and estimate the parameter of interest r_{dB} at each receive area by $\hat{r}_{dB} = (1/N) \sum_{l=1}^N p_{dB}(l)$ where N is the number of measured locations of the receive area.

Figure 6.1 shows the point estimate (see Appendix E) $\hat{r}_{dB}$ (circle markers) for each of the receive areas based on $N = 17$ measured samples $p_{dB}(l)$ ($\times$ markers). It also shows the 90% confidence interval for r_{dB} (triangle markers). For convenience we do not show in the figure

²We note that all the estimated receive powers are relative to some unknown reference as indicated in Section 2.3.

Figure 6.1: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for r_{dB} (triangle markers), point estimate $\hat{r}_{dB}$ (circle markers), and measured $p_{dB}(l)$ ($\times$ markers) in dB

the standard errors for each of the endpoints of the confidence interval that we found to be close to zero, which indicates the good accuracy of the estimated intervals.

We observe in the figure that the difference in the mean powers of the different receive areas are caused by their associated path losses. We see for instance that *DU*, which is the closest receive area to the transmitter, has the smallest path loss; and that *SA*, which is the furthest receive area to the transmitter, has the largest path loss. In particular we note that there is $[19 - 22]$ dB path loss difference between *SA* and *DU*. We also observe path loss differences of $[4 - 10]$ dB between *LI* and *DU*, $[2 - 7]$ dB between *LO* and *LI*, and $[7 - 10]$ dB between *SA* and *LO*. We note the large path loss difference between *LO* and *SA*, which are located relatively close to each other. This is explained by the fact that measurements in *SA* were performed two floors below the other receive areas, as indicated in Section 3.1.

Before describing the fixed Gaussian multi-user channel and the *OSDM* schemes introduced

by [Spencer *et al.* (2004)], we present below the normalization procedure that we use with the measured multi-user channels to find their ergodic and outage user rate point estimates given at the end of this chapter.

Normalization of measured channels: The measured transfer functions of all the receive areas are scaled by the same factor. More specifically, the measured transfer functions of all receive areas must satisfy

$$\frac{1}{FLJ} \sum_f \sum_l \sum_j \text{Tr}(\mathbf{H}_f(l, j) \mathbf{H}_f^\dagger(l, j)) = N_T N_R \quad (6.1)$$

where $\mathbf{H}_f(l, j)$ are the measured MIMO channels with $N_T N_R$ transfer functions. The average is taken over all F frequencies, over all L locations and over all J receive areas. This normalization procedure preserves the mean receive power variations of the different measured receive areas caused by user path loss differences as well as the mean receive power variations within a receive area caused by shadowing.

6.2 Downlink fixed channels with Gaussian noise

This section considers the case of downlink fixed Gaussian channels with one frequency component. Let the receive signal for the J users be given by

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} \quad (6.2)$$

where $\mathbf{y} = \langle \mathbf{y}_1^T, \dots, \mathbf{y}_j^T, \dots, \mathbf{y}_J^T \rangle^T$ is the $N_R \times 1$ receive vector, and $\mathbf{y}_j$ is the $N_R \times 1$ receive vector of the j^{th} user; $\mathbf{H} = \langle \mathbf{H}_1^T, \dots, \mathbf{H}_j^T, \dots, \mathbf{H}_J^T \rangle^T$ is the $N_R \times N_T$ downlink multi-user channel matrix, and $\mathbf{H}_j$ is the $N_R \times N_T$ channel matrix of the j^{th} user; $\mathbf{s}$ is the $N_T \times 1$ transmitted signal; and $\mathbf{n} = \langle \mathbf{n}_1^T, \dots, \mathbf{n}_j^T, \dots, \mathbf{n}_J^T \rangle^T$ is the $N_R \times 1$ noise vector distributed

according to $\mathcal{CN}_{N_R}(\mathbf{0}, \mathbf{I}_{N_R})$, and $\mathbf{n}_j$ is the $N_{R_j} \times 1$ noise vector of the j^{th} user.

Assuming perfect channel knowledge at the transmitter and at the receiver, we are interested in the MIMO transmission systems that perform linear pre-processing at the transmitter given by a modulation matrix $\mathbf{M}$ and linear post-processing at the receiver given by a demodulation matrix $\mathbf{D}$. Let m be the total number of sub-channels employed by all users to communicate with the transmitter i.e., $m = \sum_{j=1}^J m_j$, then the $m \times 1$ vector of the estimated data $\hat{\mathbf{x}}$ can be obtained from the received signal $\mathbf{y}$ by applying the $m \times N_R$ demodulation matrix $\mathbf{D}$

$$\hat{\mathbf{x}} = \mathbf{D} (\mathbf{H} \mathbf{s} + \mathbf{n}) \quad (6.3)$$

where $\mathbf{D} = \langle \mathbf{D}_1, \dots, \mathbf{D}_j, \dots, \mathbf{D}_J \rangle$ is composed of the individual user demodulation matrices $\mathbf{D}_j$, $\mathbf{s} = \mathbf{M} \mathbf{x}$ is the transmitted signal, $\mathbf{M} = \langle \mathbf{M}_1, \dots, \mathbf{M}_j, \dots, \mathbf{M}_J \rangle$ of size $N_T \times m$ is composed of the individual modulation matrices $\mathbf{M}_j$, and $\mathbf{x}$ is the $m \times 1$ transmitted information vector.

The signal at the j^{th} receiver can be expressed by

$$\mathbf{y}_j = \sum_{i=1}^J \mathbf{H}_j \mathbf{M}_i \mathbf{x}_i + \mathbf{n}_j \quad (6.4)$$

$$= \mathbf{H}_j \mathbf{M}_j \mathbf{x}_j + \mathbf{H}_j \sum_{i \neq j} \mathbf{M}_i \mathbf{x}_i + \mathbf{n}_j \quad (6.5)$$

The purpose of the user modulation matrices $\mathbf{M}_j$ is to eliminate the interuser interference, so that the product $\mathbf{H} \mathbf{M}$ in (6.3) is a $N_R \times m$ block diagonal matrix. In the next section we describe the *OSDM* multi-user schemes that find these user modulation matrices.

6.2.1 OSDM schemes for the downlink

In this section we describe some of the *OSDM* schemes presented in [Spencer *et al.* (2004)], which assume the channel model of the previous section. These methods are attractive to use for transmission on MIMO downlink channels because of the simplicity of their implementation as opposed to the implementation of the optimal solution achieved by the dirty paper region [Weingarten *et al.* (2005), Caire and Shamai (2003)]. Also, the sub-optimal channel capacity solution achieved by these methods is relatively easy to compute compared to the optimal channel capacity solution that requires computational demanding optimization algorithms presented in [Mohseni *et al.* (2006)].

In the *OSDM* schemes we use, we assume that the transmitter and the receiver have perfect channel knowledge. The strategy of these methods is to suppress the interuser interference at the transmitter by adding some constraints on each user channel, so that the resulting orthogonal channel can be optimized using the water-filling technique which corresponds to the optimal solution of MIMO single-user channels. Furthermore when there is more than one receive antenna per user, the authors in [Spencer *et al.* (2004)] proposed the so called coordinated transmit-receive processing technique in which the sub-channels, limited in number by the number of transmit antennas N_T at the *BS*, are shared among the J users in a non-overlapping fashion. To this end each user $j \in [1, \dots, J]$ employs its antennas to communicate with the transmitter using only m_j sub-channels such that $\sum_{j=1}^J m_j \leq N_T$. Bearing in mind the sub-channels that the user is likely to use, the transmitter adds constraints to each user channel such that it does not interfere with the channels of the other users.

We describe this technique in the following two steps. In the first step, we present the processing performed at the transmitter. In the second step, we describe the processing

performed at the receiver and we explain how this result is integrated to the solution of the previous step.

6.2.1.1 Step 1: transmit processing

In our study we use two of the *OSDM* schemes presented in [Spencer *et al.* (2004)]. The first one corresponds to the block diagonal method for the maximum throughput ($BD - \Sigma$), and the second one corresponds to the block diagonal method for power control ($BD - PC$). As we will see later, the $BD - \Sigma$ method is adequate for the cases where the statistics of the user channels are similar since it delivers fair or balanced rates to all users. However, if the user channel statistics are considerably different, this approach favors the users with strong channels (e.g., users located close to the *BS* whose channel gains are less attenuated than those of users located far from the *BS*) at the expense of delivering low rates to the users with weak channels. The $BD - PC$ method is more appropriate when one is interested in delivering the same quality of service per user. In this case the objective is to minimize the transmitted power such that the targeted rates per user are achieved.

Even though our target is to deliver fair rates to all users employing the $BD - PC$ method, we are still interested to find the maximum sum rates employing the $BD - \Sigma$ method, for comparison purposes. Below we give the details of each of these methods.

(a) sum rates using $BD - \Sigma$: the objective of this algorithm is to maximize the sum capacity given that the input transmit power is P . The strategy here is to add constraints on each user channel to eliminate the inter-user interference. This is achieved by finding the modulation matrix $\mathbf{M}$ such that the product $\mathbf{H} \mathbf{M}$ in (6.3) is a block diagonal matrix.

We first need to find the channel constraints for zero interuser interference. Let

$$\tilde{\mathbf{H}}_j = \langle \mathbf{H}_1^T, \dots, \mathbf{H}_{j-1}^T, \mathbf{H}_{j+1}^T, \dots, \mathbf{H}_J^T \rangle^T \quad (6.6)$$

Also let $\text{rank}(\tilde{\mathbf{H}}_j) = \tilde{L}_j$ and the *SVD* of $\tilde{\mathbf{H}}_j$ be given by

$$\tilde{\mathbf{H}}_j = \tilde{\mathbf{U}}_j \tilde{\Sigma}_j \langle \tilde{\mathbf{V}}_j^{(1)}, \tilde{\mathbf{V}}_j^{(0)} \rangle^\dagger \quad (6.7)$$

where $\tilde{\mathbf{V}}_j^{(1)}$ corresponds to the first $\tilde{L}_j$ columns of the right singular vectors of $\tilde{\mathbf{H}}_j$ and $\tilde{\mathbf{V}}_j^{(0)}$ corresponds the last $N_T - \tilde{L}_j$ columns of the right singular vectors of $\tilde{\mathbf{H}}_j$. Since $\tilde{\mathbf{V}}_j^{(0)}$ corresponds to the null space of $\tilde{\mathbf{H}}_j$ we choose these matrices as the channel constraints for user j that achieve zero interuser interference. The new constrained channel matrix is given by:

$$\mathbf{H}_c = \begin{bmatrix} \mathbf{H}_1 & \tilde{\mathbf{V}}_1^{(0)} & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{H}_J & \tilde{\mathbf{V}}_J^{(0)} \end{bmatrix} \quad (6.8)$$

The sum capacity using *BD* – Σ method is obtained by

$$\Sigma(\mathbf{H}, \mathbf{M}) = \max_{\mathbf{M}} (r_{\text{sum}}(\mathbf{H}, \mathbf{M})) \quad (6.9)$$

$$\text{subject to } \text{Tr}(\mathbf{M}\mathbf{M}^\dagger) \leq P \quad (6.10)$$

where

$$r_{\text{sum}}(\mathbf{H}, \mathbf{M}) = \log \left(\det \left\{ \mathbf{I}_{N_R} + \frac{1}{\sigma_n^2} \mathbf{H} \mathbf{M} \mathbf{M}^\dagger \mathbf{H}^\dagger \right\} \right) = \sum_{j=1}^J \log \left(\det \left\{ \mathbf{I}_{N_{R_j}} + \frac{1}{\sigma_n^2} \mathbf{H}_j \mathbf{M}_j \mathbf{M}_j^\dagger \mathbf{H}_j^\dagger \right\} \right);$$

the noise is assumed to have unit variance $\sigma_n^2 = 1$ as specified in (6.2); and $\mathbf{M}_j, j \in [1, \dots, J]$ are feasible user modulation matrices that we find below.

We remove the interuser interference of the channel $\mathbf{H}$ in (6.9) by using the channel constraints $\tilde{\mathbf{V}}_j^{(0)}$ given in (6.8), so that the optimal solution of the resulting constrained channel can be found by applying the water-filling technique of the MIMO single-user channels. This technique consists of transmitting power on the direction of the channel right singular vectors and by water-filling the power according to the corresponding singular values.

Also, since $\mathbf{H}_c$ is block diagonal we can find the *SVD* of each individual channel as follows

$$\mathbf{H}_j \tilde{\mathbf{V}}_j^{(0)} = \mathbf{U}_j \begin{bmatrix} \Sigma_j & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \langle \mathbf{V}_j^{(1)}, \mathbf{V}_j^{(0)} \rangle \quad (6.11)$$

where $\text{rank}(\Sigma_j) = L_j$ and $\mathbf{V}_j^{(1)}$ corresponds to the first L_j columns of the right singular vectors of $\mathbf{H}_j \tilde{\mathbf{V}}_j^{(0)}$.

We can then express the block diagonal modulation matrix $\mathbf{M}$ in (6.3) as the product of the channel constraint, the optimal transmit directions of the constrained channel, and a weighted transmitted power as follows

$$\mathbf{M} = \langle \tilde{\mathbf{V}}_1^{(0)} \mathbf{V}_1^{(1)}, \dots, \tilde{\mathbf{V}}_j^{(0)} \mathbf{V}_j^{(1)}, \dots, \tilde{\mathbf{V}}_J^{(0)} \mathbf{V}_J^{(1)} \rangle \mathbf{\Lambda}^{1/2} \quad (6.12)$$

where $\mathbf{\Lambda}^{1/2}$ is a diagonal matrix whose values are chosen by water-filling the singular values of $\mathbf{H}_c$, which are given by

$$\Sigma = \begin{bmatrix} \Sigma_1 & & \\ & \ddots & \\ & & \Sigma_J \end{bmatrix} \quad (6.13)$$

(b) user rates using *BD – PC*: the objective with this algorithm is to find a set of achievable rates $\mathbf{u} = [u_1, \dots, u_j, \dots, u_J]$ where the total transmitted power is minimized.

To this end we should solve the following optimization problem

$$P_{BD}(\mathbf{H}, \mathbf{M}) = \min_{\mathbf{M}} \text{Tr}(\mathbf{M} \mathbf{M}^\dagger) \quad (6.14)$$

$$\text{subject to } \mathbf{r} = \mathbf{u} \quad (6.15)$$

$$\text{Tr}(\mathbf{M} \mathbf{M}^\dagger) \leq P \quad (6.16)$$

where $\mathbf{r} = [r_1, \dots, r_J]$ with $r_j = \log \left(\det \left\{ \mathbf{I}_{N_{R_j}} + \mathbf{G}_j(\mathbf{H}_j \mathbf{M}_j \mathbf{M}_j^\dagger \mathbf{H}_j^\dagger) \right\} \right)$; $\mathbf{G}_j = \left(\sigma_n^2 \mathbf{I}_{N_{R_j}} + \sum_{i=1, i \neq j}^J \mathbf{H}_j \mathbf{M}_i \mathbf{M}_i^\dagger \mathbf{H}_j^\dagger \right)^{-1}$ is the noise plus interuser interference of user j ; and $\mathbf{M}$ is a set of feasible user modulation matrices.

If we use the same per user channel constraints of the previous algorithm i.e., $\tilde{\mathbf{V}}_j^{(0)}$ given in (6.8), we eliminate the interuser interference for each user channel so that this channel becomes equivalent to a MIMO single-user channel. Since for the MIMO single-user channel capacity is a monotonic function of the power, then minimizing the power to achieve a target rate can also be solved by water-filling [Spencer *et al.* (2004)]. To this end we just need to modify the previous algorithm by water-filling sufficient power to each user channel (instead of water-filling over the J users simultaneously the total available power) such that it satisfies the target rate for that user.

We find the user modulation matrices $\mathbf{M}_j$ using (6.12), the only difference is in the coefficients of $\Lambda^{1/2}$. To find these coefficients we first find $\Lambda_j^{1/2}$ by water-filling Σ_j of each user, such that satisfies the user rate u_j . Then we form the diagonal matrix

$$\Lambda^{1/2} = \begin{bmatrix} \Lambda_1^{1/2} & & \\ & \ddots & \\ & & \Lambda_J^{1/2} \end{bmatrix} \quad (6.17)$$

6.2.1.2 Step 2: transmit-receive processing

This step allows a given user j , who is equipped with multiple antennas, to limit the number of sub-channels m_j that uses to communicate with the transmitter. This will avoid the situation in which users with strong channels take over the limited number of spatial dimensions usually limited by N_T , the number of antennas at the BS . The authors in [Spencer *et al.* (2004)] propose that if user j uses m_j sub-channels, then a good choice for the beamformers that the user could employ is to select the m_j strongest left singular vectors of $\mathbf{H}_j$. Since the transmitter has knowledge of the user channels it can use this information to direct the transmission towards the beamformers of each of the users.

The following is the procedure to find these beamformers as well as the integration of this solution with that of the previous step.

Let the *SVD* of $\mathbf{H}_j$ be given by

$$\mathbf{H}_j = \mathbf{U}_j \mathbf{\Sigma}_j \mathbf{V}_j^\dagger \quad (6.18)$$

Let $\mathbf{W}_j$ be the first m_j columns of $\mathbf{U}_j$, then the new channel is given by

$$\bar{\mathbf{H}} = \langle (\mathbf{W}_1^\dagger \mathbf{H}_1)^T, \dots, (\mathbf{W}_J^\dagger \mathbf{H}_J)^T \rangle^T \quad (6.19)$$

In order to transmit to these new user channels we use $\bar{\mathbf{H}}$ in (6.19) in place of $\mathbf{H}$ in (6.2) and use the solutions of the previous step. It is noted by the authors in [Spencer *et al.* (2004)] that the beamformers $\mathbf{W}_j$ are only a guess that the transmitter makes about the user (receiver) structure. The actual receiver structure will correspond to the product of the first m_j columns of $\mathbf{U}_j$ from the previous optimization algorithms with $\mathbf{W}_j$.

In the following section we describe the *TDM* multi-user scheme that we are also interested

to study in this work.

6.2.2 TDM scheme for the downlink

In this scheme the *BS* sends information to each user in a different time slot, so that the multiple data MIMO sub-channels may be transmitted to each user terminal without interference to other user terminals. We assume that the receiver knows the channel to each of its receiving antennas and that the transmitter has limited channel knowledge so that it can adjust the transmission allocation time for each user. The achievable rate of this scheme is evaluated using Equation (5.8) under equal power allocation strategy. The *TDM* capacity region can be achieved by time sharing among the different users.

In the following example we illustrate the two-user capacity region of a downlink fixed channel using the *OSDM* and the *TDM* schemes described above.

6.2.3 Example: two user rate region of a downlink fixed channel

Let's consider the following 2×4 MIMO channels for each of two users, at one given frequency:

$$\mathbf{H}_1 = \begin{bmatrix} 0.008 + 0.805i & 0.570 - 0.914i & -0.700 - 0.234i & 0.205 + 0.352i \\ -0.456 - 0.484i & 0.164 - 0.052i & 0.947 - 0.597i & 1.046 + 1.053i \end{bmatrix}$$

$$\mathbf{H}_2 = \begin{bmatrix} -0.386 + 0.343i & -0.174 - 0.106i & -0.604 - 0.056i & -0.085 - 0.429i \\ -0.599 - 0.421i & 0.469 - 0.307i & -0.850 + 1.086i & -0.046 - 0.953i \end{bmatrix}$$

Assume that the base station dedicates two transmit sub-channels for each of the two users, so that $m_j = 2$, $j \in [1, 2]$.

Figure 6.2 shows the downlink two-user rate regions using the *OSDM* scheme (*BD-PC* method) given by the solid lines. These regions are obtained by different combinations of

Figure 6.2: Two-user capacity regions of a downlink channel with four antennas at the BS and with two antennas at each user terminal, when using the OSDM scheme (solid line) and the TDM scheme (dotted line)

power allocations to each user $\{p_1, p_2\}$ following the procedure described in Section 6.2.1. Figure 6.2 (a) shows the case where the user channels have the same SNR and Figure 6.2 (b) shows the case where the SNR of user one is attenuated by 10 dB with respect to the SNR of user two. In both cases, the rate regions correspond to an overall SNR = 10 dB i.e., mean receive power averaged over all users. The markers show the maximum sum rate points Σ achieved when using $BD - \Sigma$ method.

Figure 6.2 also shows the achievable two-user rate regions of the TDM scheme (dashed lines), which are obtained following the procedure described in Section 6.2.2.

We observe in Figure 6.2 (a) that when the user channels have same SNR, the BD method offers a good performance at high SNR especially for the point in the curve that achieves most balanced rates for both users, so that the maximum sum rate point Σ could be used as a good optimization criterion to achieve these rates.

Figure 6.3: Downlink fading channel

We observe in Figure 6.2 (b) that when the user channels have different SNR , the $BD - \Sigma$ method tends to be biased towards the user with stronger channel i.e., the sum rate point Σ favors user two and delivers lower rate to user one.

6.3 Downlink fading channels with Gaussian noise

In this section we describe the downlink channel models presented in [Mohseni *et al.* (2006)] for multiple antenna Gaussian channels. We illustrate this type of channel in Figure 6.3.

Let N_T be number of antennas of the multi-user channel i.e., number of antennas at the base station, and let N_R be the total number of receive antennas in the multi-user channel i.e., total number of antennas of all users. Also, let N be a finite number of channel realizations with $n \in [1, \dots, N]$.

At the n^{th} channel realization, let $\mathbf{s}(n) \in \mathbb{C}^{N_T \times 1}$ be the input vector and let $\mathbf{y}(n) \in \mathbb{C}^{N_R \times 1}$ be the output vector. Also let $\mathbf{n}(n) \in \mathbb{C}^{N_R \times 1}$ be the noise vector distributed according to $\mathcal{CN}_{N_R}(\mathbf{0}, \mathbf{I}_{N_R})$ and let $\mathbf{H}(n) \in \mathbb{C}^{N_R \times N_T}$ be the channel matrix of the J users, then the

channel output is given by:

$$\mathbf{y}(n) = \mathbf{H}(n) \mathbf{s}(n) + \mathbf{n}(n) \quad (6.20)$$

where $\mathbf{y}(n) = \langle \mathbf{y}_1(n)^T, \dots, \mathbf{y}_J(n)^T \rangle^T$, $\mathbf{n}(n) = \langle \mathbf{n}_1(n)^T, \dots, \mathbf{n}_J(n)^T \rangle^T$ and $\mathbf{H}(n) = \langle \mathbf{H}_1(n)^T, \dots, \mathbf{H}_J(n)^T \rangle^T$.

The signal received by the j^{th} user can be expressed by

$$\mathbf{y}_j(n) = \mathbf{H}_j(n) \mathbf{s}_j(n) + \sum_{i=1, i \neq j}^J \mathbf{H}_j(n) \mathbf{s}_i(n) + \mathbf{n}_j(n) \quad (6.21)$$

Similarly to the case of fixed Gaussian channels given in (6.5), the objective is to find the set of user modulation matrices such that the product $\mathbf{H}(n) \mathbf{M}(n)$ is a block diagonal matrix at every channel instantiation n . Equation (6.21) can then be written by

$$\mathbf{y}_j(n) = \mathbf{H}_j(n) \mathbf{M}_j(n) \mathbf{x}_j(n) + \sum_{i=1, i \neq j}^J \mathbf{H}_j(n) \mathbf{M}_i(n) \mathbf{x}_i(n) + \mathbf{n}_j(n) \quad (6.22)$$

The long-term and short-term transmit power constraints are given by $E\{P(n)\} \leq P$ and $P(n) \leq P$, respectively.

6.3.1 Extension of the *OSDM* schemes to fading channels

The *OSDM* schemes described in Section 6.2.1 were presented by [Spencer *et al.* (2004)]. In their paper, the authors assumed fixed channels with one frequency component. In this section, we extend these *OSDM* schemes to the case of fading frequency selective channels. Since the strategy of the *OSDM* schemes is to add constraints on the user channels to eliminate interuser interference, our approach is to optimize the resulting orthogonal fading channels by using the corresponding water-filling solutions of the MIMO single-user fading channels. More specifically, we consider that each channel realization of the fading channels

is a parallel Gaussian channel so that we can apply the same optimal solutions of the MIMO single-user fading channels presented in Sections 5.3.1, 5.3.2.

In the following we formulate the optimization problems to find the ergodic and outage channel rates of the *OSDM* schemes based on the corresponding optimal rate definitions of multi-user fading channels introduced by the authors in [Mohseni *et al.* (2006)]. We assume that the elements of the multi-user channel in (6.20) are jointly stationary and ergodic random processes. Also in the solution of the *BD – PC* method under short-term input power constraint, we assume that the maximum allocated powers to the J users, $\{p_1, \dots, p_J\}$, are determined at the beginning of the transmission and are held fixed during all the channel realizations.

6.3.1.1 Ergodic rates

The ergodic sum rate of *BD – Σ* method under a long-term input power constraint is given by

$$\Sigma^{erg}(P) = \max_{\mathbf{M}(n), n \in [1, \dots, N]} E \left\{ r_{sum}(\mathbf{H}(n), \mathbf{M}(n)) \right\} \quad (6.23)$$

$$\text{subject to } E\{Tr(\mathbf{M}(n)\mathbf{M}^\dagger(n))\} \leq P \quad (6.24)$$

where $n \in [1, \dots, N]$ is the index for each realization of the channel; $r_{sum}(\mathbf{H}(n), \mathbf{M}(n)) = \log \left(\det \left\{ \mathbf{I}_{N_R} + \frac{1}{\sigma_n^2} \mathbf{H}(n) \mathbf{M}(n) \mathbf{M}^\dagger(n) \mathbf{H}^\dagger(n) \right\} \right)$; and where $\mathbf{H}(n)$ and $\mathbf{M}(n)$ are respectively, the multi-user channel and the feasible set of user modulation matrices at the n^{th} channel realization.

The optimal solution of (6.23) can be obtained by finding the user modulation matrices as described in Section 6.2.1 for the case of the *BD – Σ* method. The difference in this case

is that we allocate the power according to the water-filling solution of the ergodic capacity of the MIMO single-user fading channels described in Section 5.3.1. More specifically we find the modulation matrix $\mathbf{M}(n)$ for every realization of the channel $\mathbf{H}(n)$ using (6.12), where the coefficients of $\mathbf{\Lambda}^{1/2}(n)$ are obtained using (5.12) and (5.13). In this solution water-filling of the power is performed over the spatial sub-channels of all users and over all channel realizations $n \in [1, \dots, N]$ ³.

The ergodic user rates of the *BD – PC* method under a long-term input power constraint are found by minimizing the total transmitted power to achieve the set of rates $\mathbf{u} = [u_1, \dots, u_J, \dots, u_J]$ as follows

$$BD^{erg}(P) = \min_{\mathbf{M}(n), n \in [1, \dots, N]} \sum_{n=1}^N Tr(\mathbf{M}(n)\mathbf{M}^\dagger(n)) \quad (6.25)$$

$$\text{subject to } E\{\mathbf{r}(n)\} = \mathbf{u} \quad (6.26)$$

$$E\{Tr(\mathbf{M}(n)\mathbf{M}^\dagger(n))\} \leq P \quad (6.27)$$

where:

$\mathbf{r}(n) = [r_1(n), \dots, r_J(n), \dots, r_J(n)]$, and where

$$r_j(n) = \log \left(\det \left\{ \mathbf{I}_{N_{R_j}} + \frac{1}{\sigma_n^2} \mathbf{H}_j(n) \mathbf{M}_j(n) \mathbf{M}_j(n)^\dagger \mathbf{H}_j(n)^\dagger \right\} \right) \quad (6.28)$$

The optimal solution of (6.25) can be obtained by finding the channel modulation matrices as described in Section 6.2.1 for the case of the *BD – PC* method, with the difference that we allocate the power using the water-filling solution of the ergodic capacity of the MIMO single-user fading channels described in Section 5.3.1. More specifically, we find $\mathbf{M}(n)$ for every channel $\mathbf{H}(n)$ using (6.12), and the coefficients of $\mathbf{\Lambda}_j^{1/2}(n)$ in (6.17) are found using

³We recall that since each user channel realization consists of a set of spatial sub-channels, the water-filling process is also performed over the spatial sub-channels of each user channel realization

(5.12) and (5.13). In this solution the water-filling process is performed over the spatial sub-channels of user j and over all channel realizations $n \in [1, \dots, N]$.

Ergodic rates under a short-term input power constraint In order to obtain the ergodic user rates under a short-term input power constraint we should replace (6.24) and (6.27) by $Tr(\mathbf{M}(n)\mathbf{M}^\dagger(n)) \leq P$. The difference with the previous water-filling solutions to find the ergodic rates is that for this case the water-filling process is given by (5.14), so that water-filling is no longer performed over the channel realizations $n \in [1, \dots, N]$.

6.3.1.2 Outage rates

The outage sum rate of $BD - \Sigma$ is given by the maximum rate R_{sum} that can be achieved with a probability of outage given by $P_{out} = Prob(r_{sum}(\mathbf{H}(n), \mathbf{M}(n)) < R_{sum})$, where $r_{sum}(\mathbf{H}(n), \mathbf{M}(n))$ is the sum rate at the n^{th} channel realization defined in (6.23).

In order to find the outage rates we use an optimization procedure similar to the one described in Section 5.3.2, where the outage rate of the MIMO single user channel is maximized by minimizing the probability of outage. Given a target rate R_{sum} and a long-term input power constraint the problem is stated as follows:

$$\Sigma^{out}(R_{sum}, P) = \min_{\mathbf{M}(n), n \in [1, \dots, N]} P_{out} \quad (6.29)$$

$$\text{subject to } E\{Tr(\mathbf{M}(n)\mathbf{M}(n)^\dagger)\} \leq P \quad (6.30)$$

To find the solution of (6.29) we obtain the modulation matrices following the same procedure described in Section 6.2.1 for the case of $BD - \Sigma$ method. The difference in this case is that the water-filling process is performed as described in Section 5.3.2 for the case of long-term input power constraint. More specifically the modulation matrices $\mathbf{M}(n)$ are

found using (6.12) where the coefficients of $\Lambda^{1/2}(n)$ are obtained using the optimal power allocation algorithms of Appendix I. In this solution we transmit enough power to avoid an outage only to the users with strong channels, i.e., we waterfill enough power to avoid an outage over the spatial sub-channels of the strong users and over all the channel realizations. Otherwise transmission is turned off for the users at the channel realizations where channels are weak and power is saved for future transmissions.

In the *BD – PC* method, the outage rate of user j is given by the rate u_j that can be achieved with probability of outage $p_{out,j} = \text{Prob}(r_j(n) < u_j)$. $r_j(n)$ is the rate of the j^{th} user that is achieved at the n^{th} channel realization defined in (6.28). Let the set of target rates be given by $\mathbf{u} = [u_1, \dots, u_j, \dots, u_J]$ and the associated outage probabilities be given by $\mathbf{p}_{out} = [p_{out,1}, \dots, p_{out,j}, \dots, p_{out,J}]$. We formulate the following optimization problem

$$BD^{out}(P) = \min_{\mathbf{M}(n), n \in [1, \dots, N]} \sum_{n=1}^N \text{Tr}(\mathbf{M}(n)\mathbf{M}^\dagger(n)) \quad (6.31)$$

$$\text{subject to } \text{Prob}(\mathbf{r}(n) < \mathbf{u}) = \mathbf{p}_{out} \quad (6.32)$$

$$E\{\text{Tr}(\mathbf{M}(n)\mathbf{M}(n)^\dagger)\} \leq P \quad (6.33)$$

We solve (6.31) by finding the modulation matrices $\mathbf{M}(n)$ following the same procedure described in Section 6.2.1 for the case of the *BD – PC* method with the difference that the coefficients of $\Lambda^{1/2}(n)$ in (6.17) are obtained by using the water-filling process described in Section 5.3.2 for the case of long-term input power constraint. The modulation matrices $\mathbf{M}(n)$ are found using (6.12) and the coefficients of $\Lambda^{1/2}(n)$ in (6.17) are obtained using the optimal power allocation algorithms of Appendix I. In this solution we waterfill the power over the spatial sub-channels of each user separately and over all channel realizations. When the channel is good we transmit enough power to avoid an outage for that user at

that channel realization. Otherwise transmissions are turned off for that user at the channel realization where channel is weak and we save the power for future transmissions to that user.

Outage rates under a short-term input power constraint: when there is a short-term input power constraint, we should replace the long-term input power constraints in (6.30) and (6.33) by $\text{Tr}(\mathbf{M}(n)\mathbf{M}(n)^\dagger) \leq P$. The difference with the previous water-filling solutions to find the outage rates is that we use the water-filling process described in Section 5.3.2 for the case of short-term input power constraint. For this case, the strategy is to turn off transmissions to the users at the channel realizations where channels are weak, otherwise we transmit the total power to the users at the channel realizations with strong channels.

Frequency selective channels: when the MIMO channel $\mathbf{H}_f(n)$ has $f \in [1, \dots, F]$ independent frequency sub-bands, we can build a new channel matrix $\mathbf{H}_{fs}(n)$ whose block diagonal entries correspond to each of the channel sub-bands as follows

$$\mathbf{H}_{fs}(n) = \begin{bmatrix} \mathbf{H}_1(n) & & \\ & \ddots & \\ & & \mathbf{H}_F(n) \end{bmatrix} \quad (6.34)$$

so that the water-filling power process is also performed in frequency domain.

6.3.2 TDM scheme for fading channels

As previously indicated, in this scheme the *BS* sends information to each user in a different time slot, so that the multiple data MIMO sub-channels may be transmitted to each user terminal without interference to other user terminals.

We assume receiver knows the channel and transmitter has limited channel knowledge so

that it can adjust the average transmission time allocation per user. The ergodic user rates of this scheme are evaluated using (5.11) and the outage rates are evaluated using (5.16) under equal power allocation. Also, the respective user-rate regions are achieved by time sharing among the different users.

6.4 Finding a set of balanced rates for all users

Since our objective is to deliver balanced rates to all users, we need to look for the corresponding user power allocation that achieves the point in the capacity region where the rates are balanced for all users.

For the case of fixed Gaussian channels we propose to find a set of balanced rates $\mathbf{r}^g = [r_1^g, \dots, r_J^g]$ by using the following optimization criterion

$$\mathbf{r}^g = \min_{\{p_1, \dots, p_J\}} \sum_{i \neq j \in [1, \dots, J]} |1 - r_i/r_j| \quad (6.35)$$

$$\text{subject to: } \sum_{j=1}^J p_j \leq P \quad (6.36)$$

where J is the number of users; r_i, r_j are the rates of i^{th}, j^{th} users given by the *OSDM* scheme in (6.14); and $\{p_1, \dots, p_J\}$ is the set of allocated powers for the J users, whose sum satisfies the transmitted power constraint in (6.36).

In order to find the solution of (6.35) we compute the capacity region for different combinations of user power allocations $\{p_1, \dots, p_J\}$. We then look for the set of balanced rates $\mathbf{r}^g$ associated to the user power allocation $\mathbf{p}^g = [p_1^g, \dots, p_J^g]$ that minimizes this equation.

For the case of fading Gaussian channels we propose to find a set of balanced ergodic rates

$\mathbf{r}^{erg} = [r_1^{erg}, \dots, r_J^{erg}]$ by rewriting (6.35) as

$$\mathbf{r}^{erg} = \min_{\{\bar{p}_1, \dots, \bar{p}_J\}} \sum_{i \neq j \in [1, \dots, J]} |1 - E\{r_i(n)\}/E\{r_j(n)\}| \quad (6.37)$$

$$\text{subject to: } \sum_{j=1}^J \bar{p}_j \leq P \quad (6.38)$$

where J is the number of users; $r_i(n)$, $r_j(n)$ are the rates of i^{th} , j^{th} users at the n^{th} channel realization that are achieved by the *OSDM* scheme in (6.28); and $\{\bar{p}_1, \dots, \bar{p}_J\}$ is the set of allocated mean powers for the J users, with $\bar{p}_j = E\{p_j(n)\}$, whose sum satisfies the power constraint in (6.38).

In order to find the solution of (6.37), we first find the ergodic capacity region with different combinations of user mean power allocations $\{\bar{p}_1, \dots, \bar{p}_J\}$. We then look for the set of balanced ergodic rates $\mathbf{r}^{erg}$ associated with the mean user power allocation $\bar{\mathbf{p}}^{erg} = [\bar{p}_1^{erg}, \dots, \bar{p}_J^{erg}]$, that minimizes this equation.

For the case of outage rates we recall that the optimization problem is formulated in terms of minimizing the outage probability per user $p_{out,j}$ given a target user rate u_j and a user power constraint p_j (see Section 5.3.2). Since the formulation of this optimization problem lets one choose the target rates beforehand and since our objective is to deliver balanced rates to all users, we propose that the target rate be the same for all users i.e., $u_j = u \forall j$, such that $p_{out,j} \leq \epsilon \forall j$. This process will guarantee that all users are served with the same user rates where the associated outage probabilities are less than or equal to ϵ .

Now under this condition different user power allocations will lead to different solutions of target rates u that can be achieved, and we would like to find the user power allocation

that maximizes this target rate. Since it is computationally intensive to generate the outage capacity region, we suggest that we use the same user power allocation $\bar{\mathbf{p}}^{erg}$ of the ergodic case in order to find the point in the capacity region where we expect the target rate u is close to its maximum.

6.5 Estimated downlink rates of the measured channels: *OSDM* vs. *TDM*

In this section we compare the relative capacity of the measured multi-user channels when using the *OSDM* schemes presented above with that obtained using the *TDM* scheme. We assume that in the *OSDM* schemes the transmitter and the receiver know the channel; while in the *TDM* scheme, receiver knows the channel and transmitter has limited channel knowledge so that it can adjust the average transmission time allocation per user.

Furthermore, in the *OSDM* methods we use the strategy described in Section 6.2.1 where each user only uses one sub-channel to communicate with the transmitter. In the *TDM* scheme each user at a given time uses all four sub-channels to communicate with the transmitter.

Our results correspond all to point estimates (see Appendix E) based on seventeen channel measurements and eleven independent frequency sub-bands for the receive areas (users) of $[DU, LI, LO, SA]$. In the measurements *BS* is equipped with four antennas and each of the four users is also equipped with four antennas. The data from these multi-user channels are normalized according to (6.1). The overall *SNR* in the figures, corresponds to the average signal-to-noise-ratio per receive antenna over all users i.e., 4×16 multi-user channel, which includes the effects of path loss differences and shadowing.

In the following we present the point estimates of the ergodic rates and outage rates with $P_{out} = 0.06$ of the *OSDM* schemes using *BD-Σ*, *BD-PC* methods and the *TDM* scheme. When evaluating the ergodic rates of the *OSDM* schemes (*BD-PC*, *BD-Σ* methods) we use (6.23) and (6.25) respectively. When evaluating the outage rates of the *OSDM* schemes (*BD-PC*, *BD-Σ* methods) we use (6.29) and (6.31) respectively. The ergodic and outage rates of the *TDM* schemes are obtained using respectively (5.11) and (5.16) under equal power allocation. Moreover, the *TDM* scheme and the *OSDM* scheme with power control (*BD-PC* method), use the optimization criteria presented in Section 6.4 to find the set of balanced ergodic and outage user rates.

6.5.1 Ergodic and outage rates

Figure 6.4 shows the point estimates of the user rates of *BD-Σ* vs. the overall *SNR*. These rates are achieved with *BD-Σ* method under long-term input power constraints, $[DU_{lt}, \dots, SA_{lt}]$, and under short-term power constraints, $[DU_{st}, \dots, SA_{st}]$. We also show in the figure the point estimates of the total sum ergodic rates divided by the number of users J that we denote by $1/J \Sigma_{lt}^{erg}$ for the case of long-term input power constraints and $1/J \Sigma_{st}^{erg}$ for the case of short-term input power constraints. Ideally we would like that all the users perform the same as the dashed line (long-term sum ergodic rate), since this means that the maximum sum rate can be achieved fairly by adding all user's rates.

Figure 6.4 (a) shows the point estimates of the user ergodic rates of *BD-Σ* method under short-term and long-term input power constraints vs. the overall *SNR*. We observe that there is practically no difference between long-term and short-term input power allocations. We also see that when the overall *SNR* is 30 dB the user ergodic rates $[DU, LI, LO, SA]$ achieve respectively $\approx [146, 109, 89, 56]\%$ of $1/J \Sigma_{lt}^{erg}$ (dashed line).

Figure 6.4: Point estimates of the sum rates under long-term and short-term input power constraints of $J = 4$ users $[DU, LI, LO, SA]$.

Figure 6.4 (b) shows the point estimates of the outage rates of $BD - \Sigma$ vs. the overall SNR . The point estimates of the sum outage rates under long-term input power constraints Σ_{lt}^{out} is the same as Σ_{lt}^{erg} . When the overall SNR is 30 dB the sum outage rate under short-term input power constraints Σ_{st}^{out} is about 10% less than Σ_{lt}^{erg} . More specifically, the outage rates of $[DU, LI, LO, SA]$ achieve respectively $\approx [128, 84, 71, 40]\%$ of $1/J\Sigma_{lt}^{erg}$ under long-term input power constraints, and $\approx [118, 74, 60, 30]\%$ of $1/J\Sigma_{lt}^{erg}$ under short-term input power constraints. Similarly to what occurred for the case of single-user channels of the previous chapter, the significant improvement of the outage rate under long-term input power constraint with respect to the one under short-term input power constraint is due to shadowing.

It is clear that the $BD - \Sigma$ method favors the users with stronger channels. In both ergodic and outage rates, we see that the priority of delivering rates to users is proportional to the estimated user path loss found in Section 6.1. There we found that DU had the smallest path loss and SA the largest one. As an example, when the overall SNR is 30 dB, we see that the outage rate under short-term input power constraints of DU is almost four times that of SA .

We finally note that when the overall SNR is ≥ 25 dB the slope of all user rate curves is similar to that of $1/J \Sigma_{tt}^{erg}$. This indicates that maximum benefit for all users is obtained when this scheme operates at high levels of the overall SNR .

Figure 6.5 is similar to Figure 6.4 but for the case of balanced rates. Figure 6.5 (a) shows the point estimates of the ergodic rates of the $BD - PC$ method under long-term and short-term input power constraints vs. the overall SNR . We see that there is no difference between short-term and long-term input power allocation in the ergodic rates of this scheme. The dashed line in the figure is shown as a reference and it represents the point estimate of the ergodic long-term sum rate $1/J \Sigma_{tt}^{erg}$ of the $BD - \Sigma$ method. When the overall SNR is 30 dB, the ergodic rates of the users $[DU, LI, LO, SA]$ achieve respectively $\approx [112, 76, 79, 77]\%$ of $1/J \Sigma_{tt}^{erg}$ (dashed line).

Figure 6.5 (b) shows the point estimates of the outage rates of the $BD - PC$ method under long-term and short-term input power constraints. We observe that given a common user outage probability $P_{out} \leq 0.06$, the outage rates are the same for all users as described in Section 6.4. When the overall SNR is 30 dB, the outage user rates under long-term and short-term input power constraints achieve respectively 72% and 64% of $1/J \Sigma_{tt}^{erg}$ (dashed line).

Figure 6.5: Point estimates of the balanced rates under long-term and short-term input power constraints of $J = 4$ users $[DU, LI, LO, SA]$

We also note that the slope of all user rate curves of $BD - PC$ method is similar to the one of the dashed line when the overall SNR is ≥ 20 dB, which indicates that maximum benefit for all users is attained by operating at those levels of the overall SNR .

Figure 6.5 (c) shows the point estimates of the ergodic rates of the TDM scheme vs. the overall SNR . We see that when using equal power allocation, the ergodic rates of $[DU, LI, LO, SA]$ achieve respectively $\approx [87, 81, 85, 96]\%$ of $1/J \Sigma_{tt}^{erg}$ (dashed line), when the overall SNR is 30 dB.

Figure 6.5 (d) shows the point estimates of the outage rates of the TDM scheme vs. the overall SNR . We see that when using equal power allocation, the outage rates of $[DU, LI, LO, SA]$ achieve respectively $\approx [79, 72, 78, 74]\%$ of $1/J \Sigma_{tt}^{erg}$ (dashed line) when the overall SNR is 30 dB.

In general we see that all users are allocated similar rates for all levels of the overall SNR for the cases of ergodic and outage rates. We also note that the slope of all user rate curves is similar to the one of the sum long-term ergodic rates $1/J \Sigma_{tt}^{erg}$ when the overall SNR is ≥ 15 dB, which indicates that maximum benefit of this scheme can be achieved for all users by operating at those levels of the overall SNR .

From the previous results we see that the TDM scheme performs better than the $OSDM$ schemes in terms of both ergodic and outage rates. We also note that the superior performance of TDM over $OSDM$ is more significant in terms of outage rates than in terms of ergodic rates.

In the following section we show results related to the power, time and SNR allocation

per user that are used by the *OSDM* and the *TDM* schemes in order to achieve the rates obtained in this section.

6.5.2 Transmit power, time allocation and *SNR* for each user

Figure 6.6 shows the point estimate of the user transmit power allocation vs. the overall *SNR* of the *OSDM* schemes for the case of ergodic rates under long-term input power constraints. It also shows the point estimate of the user transmission time allocation vs. the overall *SNR* of the *TDM* scheme for the case of ergodic rates. Finally, this figure also shows the point estimate of the receive *SNR* for each user vs. the overall *SNR* for each of these schemes under the same conditions.

Figure 6.6 (a) shows the point estimate of the user transmit power allocation vs. the overall *SNR* of *BD* – Σ method. We observe that when the overall *SNR* = 0 dB the power allocated to $[DU, LI, LO, SA]$ is respectively [70, 25, 5, 0]% of the total power; and when the overall *SNR* = 30 dB the power allocated to $[DU, LI, LO, SA]$ is respectively [26, 26, 26, 22]% of the total power. This user power allocation is explained since when the overall *SNR* is low the sum rate is maximized by water-filling most of the transmit power to the users with strong channels i.e., small path losses, while when the overall *SNR* is high there is enough transmit power to waterfill to the channels of all users.

Figure 6.6 (b) shows the point estimate of the *SNR* per user vs. the overall *SNR* of the *BD* – Σ method. We see that when the overall *SNR* = 30 dB, the receive *SNR* of *DU* is ≈ 33 dB, which is ≈ 20 dB higher than that of *SA* i.e., *SNR* ≈ 13 dB. We note that the receive *SNR* of *DU* is about 3 dB higher than the overall *SNR* while the receive *SNR* of *SA* is about 17 dB lower than the overall *SNR*.

Figure 6.6 (c) shows the point estimate of the user transmit power allocation vs. the overall SNR of the $BD - PC$ method. We see that at almost all levels of overall SNR , the allocated power per user is fixed. The power allocated to $[DU, LI, LO, SA]$ is respectively $[5, 5, 15, 75]\%$ of the total power. Since the objective is to deliver fair rates to all users, most of the transmit power is allocated to the user with the largest path loss e.g., SA (see Section 6.1); and the remaining power is allocated to users with smaller path losses i.e., $[LO, LI, DU]$.

Figure 6.6 (d) shows the point estimate of the SNR for each user vs. the overall SNR for the $BD - PC$ method. We note that the gap SNR difference between DU and SA is reduced from ≈ 20 dB of the $BD - \Sigma$ method to ≈ 6 dB of the $BD - PC$ method. We also note that the receive SNR of any of the users is lower than the overall SNR , for instance when the overall $SNR = 30$ dB, the receive SNR of DU is 25 dB while the receive SNR of SA is 18 dB.

Figure 6.6 (e) shows the point estimate of the user transmission time allocation vs. the overall SNR of the TDM scheme. We see that when the overall $SNR = 0$ dB, the time allocated to $[DU, LI, LO, SA]$ is respectively $[5, 10, 20, 65]\%$ of the total transmission time. When the overall $SNR = 30$ dB, the time allocated to $[DU, LI, LO, SA]$ is respectively $[15, 20, 25, 40]\%$ of the total transmission time. We note in this scheme that transmission time is allocated proportional to the path loss experienced by each of the users found in Section 6.1, so that most of the transmission time is allocated to the users with large path losses. At high transmit powers the transmission time allocation order is still maintained, however the allocation of transmission time is distributed more uniformly among all users. This can be explained since capacity grows slowly when the overall SNR is low, so that more transmission time needs to be allocated to the users with weaker channels i.e., large

(a) Transmit power $BD - \Sigma$ (b) $SNR BD - \Sigma$ (c) Transmit power $BD - PC$ (d) $SNR BD - PC$ (e) Transmit power TDM (f) $SNR TDM$ Figure 6.6: Point estimates of the transmit power, time allocation and SNR per user with $OSDM$ and TDM schemes

path losses, in order to obtain balanced rates for all users. When the overall SNR is high instead, capacity grows faster so that less transmission time can be allocated to users with weaker channels in order to obtain balanced rates for all users.

Figure 6.6 (f) shows the point estimate of the SNR for each user vs. the overall SNR for the TDM method. We see that there is a maximum difference in the user SNR of ≈ 20 dB due to the path loss difference between DU and SA i.e., the receive SNR of DU is 35 dB (which is 5 dB higher than the overall SNR) while the receive SNR of SA is 15 dB (which is 15 dB lower than the overall SNR).

We illustrate with an example in next section the effect of user path loss differences and shadowing on the *cdf* of the capacity of MIMO single-user channels (which can be considered as an upper bound for the rates of multi-user channels).

6.6 Discussion

We are particularly interested to analyze the sum rates of the following 4×4 downlink cooperative multi-user channels, i.e., assuming users can cooperate, using equal power allocation strategy at the transmitter.

- (i) *Rayleigh*: We assume a 4×4 channel with $F = 11$ independent frequencies. Channel entries are distributed according to $\mathcal{CN}(0, 1)$
- (ii) *Path loss*: We use the same type of channel as the one described in item(i) but where receive antennas have different path losses which are given by $PL = [0, 6, 12, 18]$ dB. (In the multi-user scenario using the *OSDM* schemes, each receive antenna represents a user. All users are geographically separated from each other so that they experience different path losses and therefore different mean receive powers. Also note that the

relative path loss values PL are within the same range of the user path losses found in Section 6.1)

- (iii) *Shadowing*: We use the channel given in item(i) and consider an independent shadowing factor for each user given by s , where s is $\mathcal{LN}(0, 4.5)$ as described in (5.1). (The multi-user measured channels using the *OSDM* schemes are also affected by shadowing, assuming the user moves over the different locations of a receive area. We note that the value $\sigma = 4.5$ dB is in the range of estimated values found for LI in Section 5.2).
- (iv) *Path loss + shadowing*: We use the channel given in item(ii) with independent shadowing factor s for each user, where s is $\mathcal{LN}(0, 4.5)$

For all the channels mentioned above we follow the same normalization procedure given by (6.1). We also assume that transmitter does not know the channel so that a reasonable strategy is to use equal power allocation.

Figure 6.7 (a) shows the ergodic sum rates of each of the channels of interest. We observe that at $SNR = 30$ dB, the channel with path loss and shadowing achieves $\approx 95\%$, the channel with path loss achieves $\approx 85\%$ and the channel with shadowing achieves $\approx 97\%$ of the ergodic sum rates of the Rayleigh channels. We see that the effect of path loss differences is more significant than the effect of shadowing in terms of ergodic sum rates.

Figure 6.7 (b) shows the outage sum rates of these channels. We observe that at $SNR = 30$ dB, the channel with path loss differences and shadowing achieves $\approx 76\%$, the channel with shadowing achieves $\approx 82\%$ and the channel with path loss differences achieves $\approx 82\%$ of the ergodic sum rates of the Rayleigh channels. We note that effect of shadowing is similar to the effect of path loss differences in terms of sum outage rates.

(a) Ergodic sum rates

(b) Outage sum rates with $[P_{out} = 0.06]$.

Figure 6.7: Ergodic and outage sum rates of 4×4 cooperative multi-user channels using equal power allocation strategy and with $F = 11$ independent frequencies

Figure 6.8: *cdfs* of the sum capacities of 4×4 cooperative multi-user channels using equal power allocation strategy and with $F = 11$ independent frequencies

We observe in both figures that the effect of path loss differences on the ergodic and outage sum rates increases at high $SNRs$, while the effect of shadowing on ergodic and outage sum rates is the same at all $SNRs$.

Figure 6.8 shows the *cdfs* of the sum capacities of each of the channels mentioned above, at $SNR = 30$ dB. We see in the figure that the *cdf* of the sum capacity of the channel with path loss differences (dashed line) is shifted to the left of that of the Rayleigh channel (thin solid line). We note that the *cdf* of the sum capacity of the channel with shadowing (dotted line) extends over a larger range of sum capacity values than that of the Rayleigh channel. Also, we note that the *cdf* of the sum capacity of the channel with path loss differences and shadowing (thick dotted line) is similar to that of the channel with shadowing.

In general we see that shadowing causes a larger variation on the sum capacity with respect to that of Rayleigh channels, this fact particularly affects the sum capacity of the channels with shadowing at low outage probabilities.

For the case of channels with path loss differences among users, we observe that the mean power differences among the MIMO channel elements of different receivers cause an unbalance in the eigenvalues of the channel covariance matrix. The effect of having unbalanced eigenvalues in the channel covariance matrix is a decrease in sum capacity, which is reflected by the *cdf* of the sum capacity of the channels with path loss differences being shifted to the left of that of Rayleigh channels, i.e. path loss differences have a similar effect on the sum capacity at any outage probability.

As a result we observe that the sum rates of the multi-user channels will similarly be affected by shadowing and path loss differences and that their sum rates will be upper bounded by those of the cooperative multi-user channels explained in this example.

Finally based on an example that we show in Appendix K, we are able to specifically illustrate the effect of path loss differences and shadowing of the multi-user channels on the *cdf* of the channel capacity when using the *OSDM* and the *TDM* schemes. In our example we consider a two-user fading channel with one frequency component. Our simulation results show that, path loss differences affect more the ergodic and outage rates of the *OSDM* scheme than those of the *TDM* scheme. Our results also show that, shadowing decreases more the outage rates of the *TDM* scheme than those of the *OSDM* scheme. However, the fact that our measured multi-user channels were more severely affected by path loss differences than by shadowing resulted in the overall superior rates of the *TDM* scheme with respect to those of the *OSDM* scheme.

A previous work presented in [Jindal and Goldsmith (2005)] compared the capacities of the *TDM* and dirty paper coding schemes for the case of MIMO downlink channels. The following are the differences of our study with respect to the previous study. First, we study the capacity of the *TDM* and *OSDM* schemes based on real world channel measurements. The measured user channels are in fact non-symmetrical, which is not considered in the previous theoretical study where user channel statistics are assumed to be symmetrical. Second, we are not interested to maximize the throughput performance but instead to deliver fair rates to all users. The results of the previous work favors the performance of dirty paper coding over *TDM*. In our study, we find that the asymmetries among the measured user channels due to path loss differences among users degrades more the capacity of *OSDM* scheme than that of *TDM* scheme.

Chapter 7

Conclusions

In this section we present first the results of our work, the list of contributions, followed by the conclusions and the list of suggested topics for future work.

The objective of this thesis was to establish a performance comparison of outdoor-to-indoor measured MIMO multi-user channels, in terms of ergodic and outage rates, using the orthogonal space division multiplexing (*OSDM*) schemes and the time division multiplexing (*TDM*) scheme for the downlink. The approach we followed to achieve this objective is described as follows. We first built a (4×4) MIMO broadband channel sounder. We then used this equipment to perform channel measurements in the scenario of interest. We obtained the channel impulse response estimates from the collected data and used them in our analysis. We did this analysis using tools from multivariate statistics as well as bootstrap methods. The latter methods were useful tools that allowed us perform the statistical analysis of the data based on a reduced number of measurements. Finally, we performed our analysis in two stages. We first analyzed the measured MIMO single-user channels. We then extended the insight gained in this stage, to analyze the measured MIMO multi-user channels in the next stage. The results of our analysis are summarized below.

Summary of Results:

- (1) We found that spatial correlations are not a major factor in reducing capacities in the measured channels. This is mainly due to the large antenna separation at the transmitter and receiver, the *NLOS* condition of the measured channels and the rich scattering environment. We also found that variations in the mean receive power could affect considerably the performance of the measured channels. More specifically, the capacity of the single-user channel is significantly decreased when the amount of shadowing in the channel is considerable. Also, the capacity of the multi-user channel was significantly decreased due mainly to the presence of path loss differences, and to a lower extent, to the presence of shadowing in the channel. These results indicate that when estimating the channel capacity of a given area, shadowing should be considered in small coverage areas; while path loss differences and shadowing should be considered in larger coverage areas.
- (2) We also showed that shadowing and path loss differences affect differently the *cdf* of the channel capacity. While path loss has a similar effect on the channel capacity at any outage probability, shadowing affects mostly the channel capacity at low outage probabilities. If the performance metric of interest is ergodic capacity, then shadowing is not expected to be of much concern.
- (3) We found that at high levels of the overall *SNR* (averaged over all users), the ergodic and outage rates for the *TDM* scheme are superior to those of the *OSDM* schemes. This is mainly due to the significant path loss differences among users. The advantage of *TDM* over *OSDM* is explained as follows. In order to deliver fair rates to all users, *OSDM* compensates for possible differences in the *SNRs* of each user by allocating most of the transmit power to the users whose *SNRs* are low (this is especially noticeable when there are considerable differences in the *SNRs* of the different users).

As a result, the SNR for each user may be much lower than the overall SNR , so that this scheme is not able to exploit the linear capacity growth of the MIMO channel. In the TDM scheme instead, this compensation for the differences in the $SNRs$ of the different users is performed by allocating more transmission time to the users with low $SNRs$. This means that most of the transmission time is allocated to the users whose $SNRs$ are much lower than the overall SNR , so that they cannot exploit either the linear capacity growth of their MIMO channels. It also means that a fraction of transmission time is allocated to the users with high $SNRs$. Since the $SNRs$ of these users are higher than the overall SNR , they are able to exploit the linear capacity growth of their MIMO channels when the overall SNR is high. We note that the fraction of time allocated to the users with high $SNRs$ becomes significant when the overall SNR is high. This results in the superior performance of TDM .

- (4) We illustrated with a two user example the effect of shadowing and path loss differences on the *cdfs* of channel capacities when using the $OSDM$ schemes and the TDM scheme. We found that, when there are no path loss differences or shadowing in the channel, the outage rates of TDM are better than those of $OSDM$. However, the ergodic rates of $OSDM$ are better than those of TDM . We also found that path loss differences affect more the outage and ergodic rates of $OSDM$ than those of TDM . Finally, we found that shadowing has a higher impact on the outage rates of TDM than on those of $OSDM$. The effect of shadowing on the ergodic rates is similar for both TDM and $OSDM$. The fact that our measured channels were severely affected by path loss differences resulted in the overall superior rates of the TDM scheme.
- (5) When there are large power variations (e.g., due to path loss differences among largely spaced users) in some of the sub-channels of the MIMO multi-user channel and when the objective is to deliver fair rates to all users, we obtain the following insight. It

is better to use a multi-user scheme that transmits on a subset of balanced-power sub-channels at each channel realization (so that the large power variations of the sub-channels occur between channel realizations), than to use a multi-user scheme that transmits on a subset of unbalanced-power sub-channels all the time (so that there are no significant power variations in the sub-channels of different channel realizations).

List of Contributions:

- (1) Development of the channel sounder and performance of channel measurements
- (2) Novel use of bootstrap methods to estimate parameters of the measured channels
- (3) Analysis of the capacity of measured channels in terms of:
 - Spatial correlations: developed a new method to estimate the degree of correlation of the elements of the measured MIMO channels by estimating the number of non-significant eigenvalues of their covariance matrices
 - Shadowing
 - Path loss differences
- (4) Performance comparison of *OSDM* and *TDM* of real world measured channels

Conclusions

- (1) Prior knowledge of spatial correlations, shadowing and path loss can be used to:
 - Predict realistic limits on the capacity that can be achieved in measured channels
 - Suggest transmit receive schemes that best exploit the gains of measured channels
- (2) In multi-user channels symmetries among users may not be a realistic assumption

- (3) The main detrimental factor on the capacity of multi-user channels is path loss differences when users are largely spaced. Also when users are moderately spaced, shadowing may affect the capacity of multi-user channels.

Topics for future work:

- (1) We analyzed the performance of the measured channels in terms of capacity. An extension to this study could be done by analyzing the performance of the measured channels using a given transmit receive architecture in terms of *BER*.
- (2) When studying the performance of the measured channels using spatial multiplexing multi-user schemes, we assumed perfect channel knowledge at the transmitter which is reasonable for the case of fixed or slow moving users. However, when considering mobile users a perfect channel estimation may not be possible due to the fast varying channel, so that a more suitable assumption would be to consider imperfect channel knowledge at the transmitter.
- (3) We studied the performance of a given configuration of downlink channel. One could also investigate the performance of different configurations of measured channels for different number of transmit antennas and receive antennas as well as different number of users.
- (4) We analyzed the performance of spatial multiplexing schemes for the case of measured downlink channels. A similar analysis can be extended to analyze the performance of these schemes for the case of measured uplink channels, which are further limited by a power constraint for each user.
- (5) We evaluated the performance of linear spatial multiplexing schemes based on zero-forcing methods, which suppress all the inter-user interference by adding constraints

to the channel. A further study is needed in order to evaluate the performance of other linear schemes (e.g., schemes based on minimum mean square error methods) as well as the performance of non-linear schemes, which may offer superior performance at the expense of higher implementation costs.

- (6) We established a comparison between spatial and time multiplexing schemes assuming high values of SNR . A further analysis could be done by establishing this comparison considering low values of SNR .
- (7) In our analysis we only consider the case of one cell scenario. A next step could be implemented by considering multiple cells so that there is also co-channel interference.

Appendix A

Background on radio channels

To study radio channels, one can use the physical models derived on the basis of electromagnetic theory and associated fundamental equations. For most wireless applications analyzing such a model could become very complex and would not be practical. Instead of using accurate physical models one can approximate them by mathematical models which are much more tractable analytically.

We start with an appropriate deterministic mathematical representation for the physical radio channel then we further simplify this mathematical approximation by finding first its lowpass equivalent and then its discrete time representation. We continue building our model by adding noise to the discrete representation and finally we complete this modelling by incorporating its statistical properties.

In the following we discuss an approximation of the physical wireless channel based on the input output relationship between the transmit and receive antennas which is represented as the impulse response of a time varying linear filter.

Multipath mathematical model: we first assume a deterministic multipath channel without noise which can be modeled as a linear time varying filter. The multipath term is in reference to multiple signals caused for example by reflections from objects. The bandpass

output of a radio channel can be described by

$$y_r(t) = \Re \left\{ \sum_{i=1}^L a_i(t) x(t - \tau_i(t)) e^{j 2\pi [(f_c + f d_n(t)) (t - \tau_i(t))]} \right\} \quad (\text{A.1})$$

where

- L is the number of different multipath components taken by the transmitted signals as they travel between transmitter and receiver
- $a_i(t)$ is the attenuation along the i^{th} multipath component at time t , and $\tau_i(t)$ is the associated delay.
- $f d_n(t)$ is the Doppler shift for each multipath component at time t
- $x(t)$ is the lowpass input signal

Lowpass representation: Equation (A.1), where the spectrum of input signal is centered around carrier frequency f_c , is called the bandpass representation. Analyzing bandpass signals is more cumbersome than analyzing baseband signals and since f_c does not contribute any information, we can use the equivalent low pass representation of the signals which is nothing but shifting the original signal spectrum from the carrier frequency f_c to the zero frequency. Without loss of generality, based on what said above, from this point on we assume lowpass representation for signals.

Using the mathematical description of a lowpass signal [Haykin (1994)], we write (A.1) as

$$y_r(t) = \Re \left\{ \sum_{i=1}^L a_i(t) e^{-j\phi_i(t)} x(t - \tau_i(t)) \right\} e^{j 2\pi f_c t} \quad (\text{A.2})$$

where

$$\phi_i(t) = 2\pi [(f_c + f d_i(t)) \tau_i(t) - f d_i(t) t] \quad (\text{A.3})$$

The complex valued lowpass impulse response can be written as

$$\tilde{h}(\tau, t) = \sum_{i=1}^L \alpha_i(t) \delta(\tau - \tau_i(t)) \quad (\text{A.4})$$

[Stuber (2001)]; where the impulse response coefficients are given by

$$\alpha_i(t) = a_i(t) e^{-j\phi_i(t)} \quad (\text{A.5})$$

and finally the lowpass received signal is given by the convolution of the lowpass representation of the input signal with the lowpass model; i.e.,

$$y(t) = \int_{-\infty}^{\infty} \tilde{h}(\tau, t) x(t - \tau) d\tau \quad (\text{A.6})$$

Discrete time channel model: in practice because of widely used computational engines, known as DSPs, in wireless communication systems, most of the algorithms and post processing are done on discrete time signals rather than their continuous time counterparts. Because of this an equivalent discrete time channel model is used to further simplify the analysis done on discrete time received signals. In the following we present this equivalent discrete time representation.

Based on the Nyquist sampling theorem, a band-limited signal $y(t)$ with bandwidth of $W/2$ can be reconstructed from its equally spaced samples at time intervals $T = \frac{1}{W}$. Due to spreading function $\tilde{h}(\tau, t)$ in (A.6), the bandwidth of output signal $y(t)$ may be greater than the bandwidth of input signal $x(t)$, and so $x(t)$ can be reconstructed from its samples spaced at the same time interval T as the output samples using

$$x(t) = \sum_{m=-\infty}^{\infty} x(mT) \phi(t - mT) \quad (\text{A.7})$$

where $\phi(t - mT)$ is the interpolation function and is given by

$$\phi(t - mT) = \frac{\sin(\pi W(t - mT))}{\pi W(t - mT)} \quad (\text{A.8})$$

[Proakis (2001)]. Substituting (A.7) in (A.6) we get

$$y(t) = \sum_{m=-\infty}^{\infty} x(mT) \int_{-\infty}^{\infty} \tilde{h}(\tau, t) \phi(t - \tau - mT) d\tau \quad (\text{A.9})$$

The discrete time representation of the output signal then is given as

$$y(kT) = \sum_{m=-\infty}^{\infty} x(mT) \int_{-\infty}^{\infty} \tilde{h}(\tau, kT) \phi((k - m)T - \tau) d\tau \quad (\text{A.10})$$

or

$$y(kT) = \sum_{m=-\infty}^{\infty} x(mT) h((k - m)T, kT) \quad (\text{A.11})$$

where

$$h((k - m)T, kT) = \sum_{i=1}^L \alpha_i(kT) \phi((k - m)T - \tau_i(kT)) \quad (\text{A.12})$$

[Telatar (2003)]. Equation (A.12) is obtained using (A.4). Also note that the frequency spectrums of $\tilde{h}(\tau, t)$ and of $h((k - m)T, kT)$ are the same in the frequency interval $|s| \leq \frac{W}{2}$.

Additive white Gaussian noise : generally the source of noise that predominates in wireless systems at high frequencies, such as $> 2 \text{ GHz}$, is thermal noise which is caused by the agitation of electrons of the conductors inside wireless receivers. Based on this we can assume there are many independent noise sources such that when all added together, based on the central limit theorem, the resulting noise can be modeled as Gaussian distributed. In the following we first find the lowpass representation of the bandpass noise signal, we then describe its statistical properties and finally we extend discrete time channel model to

include this lowpass noise representation.

- (1) Statistical characterization of lowpass noise signal: let $n_r(t)$ be *wss* narrowband band-pass Gaussian noise with zero mean with power spectral density

$$S_{n_r}(s) = \begin{cases} \frac{N_o}{2} & ||s| - f_c| \leq \frac{W}{2} \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.13})$$

then its lowpass equivalent representation $n(t)$ is *wss* proper and zero mean complex random process with spectral density

$$S_n(s) = \begin{cases} N_o & |s| \leq \frac{1}{2}W \\ 0 & |s| > \frac{1}{2}W \end{cases} \quad (\text{A.14})$$

[Neeser and Massey (1993)]; This random process is completely defined by the autocorrelation function

$$R_n(\tau) = \frac{1}{2} E\{n(t+\tau) n(t)^*\} \quad (\text{A.15})$$

$$= N_o \frac{\sin \pi W \tau}{\pi \tau} \quad (\text{A.16})$$

where in the limit, when $W \rightarrow \infty$, $R_n(\tau)$ is given by

$$R_n(\tau) = N_o \delta(\tau) \quad (\text{A.17})$$

Furthermore for complex proper random Gaussian processes the autocorrelation and cross correlation functions of $n_i(t)$ and $n_q(t)$ should satisfy the following properties

$$R_{n_i}(\tau) = R_{n_q}(\tau), \quad (\text{A.18})$$

$$R_{n_i n_q}(\tau) = -R_{n_q n_i}(-\tau) \quad (\text{A.19})$$

(2) Discrete time noise signal: the noise signal $n(t)$ can be reconstructed as shown in (A.7) and (A.8) from the discrete samples $n(mT)$. We also know that the new discrete random process is proper and complex since the continuous process is also proper [Neeser and Massey (1993)].

We can write now the discrete channel model with noise as

$$y(kT) = \sum_{m=-\infty}^{\infty} x(mT) h((k-m)T, kT) + n(kT) \quad (\text{A.20})$$

and replacing $k-m$ by i

$$y(kT) = \sum_{i=-\infty}^{\infty} x((k-i)T) h(iT, kT) + n(kT) \quad (\text{A.21})$$

Finally adopting the following notation $y_k = y(kT)$. The output from a discrete time Gaussian channel (DTGC) can then be described as

$$y_k = \sum_{i=-\infty}^{\infty} h_i x_{k-i} + n_k \quad (\text{A.22})$$

[Neeser and Massey (1993)].

Statistical channel model: since the characteristics of radio channels often change with time and frequency, it is useful to incorporate these variations into the model. Here we distinguish between the small scale variations, which are rapid fluctuations in the strength of the signal over very short distances, and the large scale variations, which are slow variations of the signal strength over arbitrary distances between transmitter and receiver.

Generally the parameters that measure these variations depend on the particular channel

and are obtained from measurements.

Small scale variations: These variations take place at the scale of the wavelength of the carrier frequency. The receive power may fluctuate by 30 or 40 dB when the receiver changes position by a fraction of a wavelength [Rappaport (1996)].

These rapid power variations can be caused by: the multipath components arriving at the receiver with different phases; Doppler frequency shifts caused by relative motion of the transmitter, the receiver and/or surrounding objects which determines how fast the channel changes in time; and the relative signal bandwidth with respect to the channel coherence bandwidth that determines how quickly the signal changes in frequency.

The measurements of small scale variations are important for example if we want to find the number of taps to model the channel or if we want to know how quickly these taps change with time and frequency.

(1) *Small scale fading based on the time delay spread*

(a) *Frequency flat fading channels:* for flat fading channels or channels with a very narrow bandwidth the multipath delay spread, which is the time spread of the propagation of the multipath components, is much smaller than the data symbol duration; i.e., $T_m \ll T$. If we assume in (A.12) that the number of multipath components L is large and that the phase of each component is uniformly distributed over the interval $[0, 2\pi]$ then the magnitude of the envelope $|h_k|$ can be modeled as Rayleigh distributed. For the case where there are specular or fixed components then the envelope has non-zero mean and the magnitude of the envelope $|h_k|$ can be modeled as Ricean distributed. See [Stuber (2001)] for the description of these distributions.

(b) *Frequency selective channels:* these channels are generated when the multipath delay spread is much longer than the data symbol duration ($T_m \gg T$) and so

they cause intersymbol interference.

Referring to (A.12), when the number of multipath components is sufficiently large and the phase of each component is uniformly distributed then the received envelope can still be approximated as a complex Gaussian process with a Rayleigh distributed envelope. As we increase the bandwidth we achieve a better tap resolution, however the Rayleigh assumption becomes weaker since there are fewer multipath components arriving to each of the taps which now have a narrower time window.

(2) *Small scale fading based on the Doppler spread*

Although here we focus on the more general frequency selective channels but some parameters of interest can also be calculated from observing flat fading channels. We assume that the multipath fading channel given in (A.4) is wide sense stationary with uncorrelated scatterers (*wssus*). More specifically we assume that phases and attenuations corresponding to different Doppler shifts are uncorrelated and that phases and attenuations associated with different multipath component delays are uncorrelated or equivalently it is wide sense stationary (*wss*) in the frequency domain [Bello (1963)]. Furthermore we assume that the channel given in (A.4) is a complex Gaussian process with zero mean.

The author in [Bello (1963)] presents a statistical characterization of this complex random process by four autocorrelations functions. One of them is defined below

$$R_h(k, n) = W E\{h^*(k; m) h(k; m + n)\} \quad (\text{A.23})$$

which is basically the autocorrelation function of each tap k as it evolves with time m . Note that in deriving (A.23), the assumptions used are that the random process describing the channel is *wss* and that it does not depend on time m , specifically taps

k are uncorrelated for different delays n .

Similarly the coherence bandwidth B_c and the coherence time T_c are related to the autocorrelation function $R_h(k, n)$ and are obtained as follows.

- Let L be the number of taps of $h(k, n)$; we have the average delay μ_k and the RMS delay spread σ_k given in terms of $R_h(k, 0)$ by

$$\mu_k = \frac{\sum_{k=0}^{L-1} k R_h(k, 0)}{\sum_{k=0}^{L-1} R_h(k, 0)} \quad (\text{A.24})$$

$$\sigma_k = \sqrt{\frac{\sum_{k=0}^{L-1} (k - \mu_k)^2 R_h(k, 0)}{\sum_{k=0}^{L-1} R_h(k, 0)}} \quad (\text{A.25})$$

The quantity we are interested to measure is the coherence bandwidth, which is given by

$$B_c \propto \frac{1}{\sigma_k} \quad (\text{A.26})$$

- Let the DFT of $R_h(0, n)$ with respect to n be given by

$$S_\nu(0, N) = \sum_{n=0}^{P-1} R_h(0, n) e^{-j2\pi Nn/P}, \quad N = 0, 1, \dots, P-1 \quad (\text{A.27})$$

The range of values for which $S_\nu(0, N)$ is significant is the Doppler spread and the coherence time T_c is given by

$$T_c \approx \frac{1}{B_d} \quad (\text{A.28})$$

[Goldsmith (2003)].

Large scale variations: the measurements of these variations are based on the local average of the signal strength. These measurements are generally used to determine the coverage area of the transmitter and they are also used to analyze the channel capacity.

Some general propagation loss models for free space, power loss exponent and shadowing presented in [Goldsmith (2003)] are given below.

- (1) *Free space path loss*: it is the minimum power loss expected in a communication system. It can be used as a reference when calculating the path loss from measurements. Let P_t be the transmitted power; P_r the received power; d the distance between transmitter and receiver; and G_T , G_R the transmitter and receiver antenna gains respectively. The path loss in free space is given by

$$\frac{P_r}{P_t} = \left(\frac{\sqrt{G_T} \sqrt{G_R} \lambda}{4\pi d} \right)^2 \quad (\text{A.29})$$

where λ is the wavelength in meters, d is the separation between transmitter and receiver in meters. We can see that the received power in free space reduces by the square of the distance per meter.

- (2) *Power loss exponent*: this model is convenient to use due to its simplicity and calculates the path loss based on the power loss exponent which is obtained from measurements. Let $PL = \left(\frac{P_r}{P_t} \right)_{dB}$ be the power ratio between transmitter and receiver in dB; $\overline{PL}(d_0)$ be the average path loss in dB at d_0 ; d_0 be a reference distance in the antenna far field and $d > d_0$ be the separation distance between transmitter and receiver, then the power loss is given by

$$PL = \overline{PL}(d_0) + 10\gamma \log_{10} \left(\frac{d_0}{d} \right)^{\frac{d_0}{d}} \quad (\text{A.30})$$

where γ is the path loss exponent that typically ranges from 2 which is the power loss in free space to 6 which is the power loss that may be given by obstructions such as buildings. When plotting path loss vs. d in a log to log scale then γ corresponds to a slope equal to 10γ dB per decade.

- (3) *Shadowing*: it is a power loss which is caused when the radiated energy by the transmitted is absorbed by the scattering objects, like buildings or trees, before it reaches the receiver. The overall received power is not only attenuated by the inverse of the distance d^2 but also there is an exponential attenuation given by the depth of the obstacle. So we can have for the same distance separation d but for a different location a different average of signal strength, this variation is what is called shadowing. According to measurements shadowing can be modelled as a random variable that is normally distributed in dB .

Let Ω be the averaged received power in dB when there is shadowing, then it can be modelled by the Gaussian distribution as follows

$$P_{\Omega}(x) = \frac{1}{\sqrt{2\pi}\sigma_{\Omega}} \exp \left[\frac{-(x - \mu_{\Omega})^2}{2\sigma_{\Omega}^2} \right] \quad (\text{A.31})$$

where the mean power μ_{Ω} depends on the path loss and building properties of the area of interest and is given by

$$\mu_{\Omega} = E\{\Omega\} \quad (\text{A.32})$$

If the path loss model includes path loss and shadowing then $\mu_{\Omega} = 0$. The values for the shadow standard deviation σ_{Ω} range from 4 to 12 dB with 8 dB been the typical value.

- (4) *Path loss and shadowing*: this model includes path loss as well as shadowing. The variables are defined as in (A.30) and (A.31).

$$PL_{shadow} = \overline{PL}(d_0) - 10\gamma \log_{10} \frac{d}{d_0} + \Omega \quad (\text{A.33})$$

where PL_{shadow} is the path loss with shadowing in dB , Ω is $\mathcal{N}(0, \sigma_{\Omega}^2)$ and the rest of variables are defined as in (A.30).

Appendix B

Complex proper normal distribution

Complex normal random vectors: we introduce the definitions of properness and circular symmetry when describing one class of complex normal random vectors as follows

Complex proper normal: If $\mathbf{x}$ is an m dimensional complex proper normal vector then its *pdf* is defined in terms of the mean vector μ and the non-singular covariance matrix $\Sigma = E\{(\mathbf{x} - \mu)(\mathbf{x} - \mu)^\dagger\}$ as follows

$$P_{\mathbf{x}}(x) = \frac{1}{\pi^n \det(\Sigma)} e^{-(\mathbf{x} - \mu)^\dagger \Sigma^{-1} (\mathbf{x} - \mu)} \quad (\text{B.1})$$

then the m -variate distribution of the random vector $\mathbf{x}$ is denoted by $\mathcal{CN}_m(\mu, \Sigma)$ and we will write $\mathbf{x}$ is $\mathcal{CN}_m(\mu, \Sigma)$.

Complex circularly symmetric normal: a complex zero mean random variable x whose real and imaginary components are *iid* and each of the real and imaginary parts have half of the variance σ^2 is circularly symmetric and is denoted by x is $\mathcal{CN}(0, \sigma^2)$.

Let the complex n dimensional random vector $\mathbf{x}$ be written as

$$\mathbf{x} = \begin{bmatrix} x_i(1) \\ \vdots \\ x_i(N) \\ x_q(1) \\ \vdots \\ x_q(N) \end{bmatrix} = \begin{bmatrix} \mathbf{x}_i \\ \mathbf{x}_q \end{bmatrix} \quad (\text{B.2})$$

$\mathbf{x}$ is circularly symmetric normal vector if and only if the following conditions are met.

$$E\{\mathbf{x}\} = \mathbf{0} \quad (\text{B.3})$$

$$E\{\mathbf{x} \mathbf{x}^T\} = E \begin{bmatrix} \mathbf{x}_i \mathbf{x}_i^T & \mathbf{x}_i \mathbf{x}_q^T \\ \mathbf{x}_q \mathbf{x}_i^T & \mathbf{x}_q \mathbf{x}_q^T \end{bmatrix} \quad (\text{B.4})$$

$$(\text{B.5})$$

$$\text{where } E\{\mathbf{x}_i \mathbf{x}_i^T\} = E\{\mathbf{x}_q \mathbf{x}_q^T\} \quad (\text{B.6})$$

$$\text{and } E\{\mathbf{x}_i \mathbf{x}_q^T\} = -E\{\mathbf{x}_q \mathbf{x}_i^T\} \quad (\text{B.7})$$

[Telatar (1995)]. Random vectors that are complex circularly symmetric normal distributed are also zero mean complex proper normal distributed [Neuser and Massey (1993)].

Appendix C

Empirical distributions of the test statistics

Below we find the empirical distributions of the different statistics we use in our analysis. These distributions are obtained with Monte Carlo simulations and the corresponding parameters are described in each of the figures.

Distribution of the statistic for testing independence of two frequencies: Figure C.1 shows the empirical distribution of $-2\ln(\lambda)$ where λ is given in (4.3). We evaluate this distribution with 10^5 simulations and the following parameters: $p = 1$, which is the number of transfer functions of the MIMO channel; $q = 2$, which is the pair of frequency components of the transfer function; and $N = 17$, which is the number of independent samples

Distribution of the statistic for testing independence of the transmit vectors: Figure C.2 shows the empirical distribution of $-2\ln(\lambda)$ where λ is given in (4.3). We evaluate this distribution with 100000 simulation points and with the following parameters $q = 4$ (four transmit vectors), $p_1 = \dots = p_4 = 4$ (number of variables in each transmit vector), $N = 17$ (number of samples).

Figure C.1: Number of occurrences vs. statistic for testing independence. Empirical distribution of $\tilde{\lambda} = -2 \ln \lambda$, with parameters $q = 2$, $p = 1$, $N = 17$.

Figure C.2: Number of occurrences vs. statistic for testing independence. Empirical distribution of the statistic $\tilde{\lambda} = -2 \ln \lambda$, with parameters $q = 4$, $p_i = 4$, $i \in [1 \cdots 4]$, $N = 17$.

Figure C.3: Number of occurrences vs. statistic for testing multiple homogeneity. Empirical distribution of the statistic $-2 \ln \lambda_5$ with parameters $p = 4$; $q = 4$; $N_i = n_i + 1 = 17$ $i \in [1, \dots, 4]$; $n_{qj} = 1$, $4 \times n_i = 64$

Distribution of the multiple homogeneity statistic for validating Kronecker model:

The empirical distribution of $\tilde{\lambda}_5 = -2 \ln(\lambda_5)$ where λ_5 is given in (4.13). Figure C.3 shows the empirical distribution of $-2 \ln(\lambda_5)$ with 10^6 simulation points and with the following parameters: $q = 4$, which is the number of covariance matrices; $p = 4$, which is the number of variables of each independent random vector; $i = 4$, which is the number of transmit random vectors; $N_i = N_0 = 17$, $i \in [1 \dots 4]$ is the size of independent samples for each of the four random vectors, $n_i = n_0 = N_0 - 1$ is the degrees of freedom and the index $n_q = 4 \times n_0$.

Distribution of the small test statistic for testing normality: Figure C.4 shows the empirical distribution for the sample skewness and sample kurtosis parameters assuming normality, a sample of size $N = 17$ and 10^6 simulation points.

Distribution of the Shapiro-Wilk test statistic for testing normality: Figure C.5 shows the empirical distribution of the W statistic for normal variables of samples size $N = 17$ using 10^6 simulation points.

Figure C.4: Number of occurrences vs. statistics for skewness and kurtosis. Empirical distributions of $b_1^{1/2}$ and b_2 for a normal variable with a sample size $N = 17$.

Figure C.5: Number of occurrences vs. Shapiro-Wilk statistic. Empirical distribution of W for normal variables of sample size $N = 17$.

Appendix D

Multivariate normality validation

Results on capacity are generally based on the assumption that the distribution in space of the elements of the MIMO channel is joint complex circular multivariate normal. This becomes the first motivation for validating this distribution since we will be able to find whether this assumption is appropriate for the measured data; otherwise based on this measured data we will indicate possible conditions under which this assumption cannot be made. Another important reason for doing this validation is that available statistical tools that may be used for analyzing the data are based on multivariate normal data, therefore to make an appropriate use of them it is important to check beforehand the data distribution. In the following we present some relevant techniques for testing multivariate normality as well as the procedure we follow to apply them to our measured data.

There is no well established theory to determine which is the best method to test multivariate normality, however two recommended techniques are the following: one is based on methods for testing univariate normality; and the other one is based on methods for assessing multivariate normality as proposed by [Looney (1995)]. It is important to note that all these theories are developed for testing normality of real random variables while our case

involves complex random variables. We use these theories to test normality of the real and imaginary parts of a complex random variable separately since this is a necessary condition for a complex normal random variable.

For the case of assessing multivariate normality using multivariate normal tests, the author in [Looney (1995)] recommends a battery of four tests: Royston's H test, Small's test, Srivastava and Hui's test and Srivastava's test. We have found that these tests are rather limited by the vector size which usually involves a limited number of variables e.g., $p \leq 8$ and also by the large number of samples required to perform the test e.g., $N \geq 50$. Both of these factors prevent us from using these tests in our analysis. Therefore we focus our study of goodness of fit based on the following univariate normal tests.

Mathematical framework of univariate normal tests: one necessary but not sufficient condition for a vector to be jointly normal distributed is that the marginal distributions should be univariate normal. This is the reason why is generally recommended to check this condition first, since there is a more developed theory for testing univariate normality than for testing multivariate normality. [D'Agostino and Stephen (1986), Small (1980), Gnanadesikan (1986)].

The following are commonly approaches used to test univariate normality: the skewness and kurtosis tests; and the Shapiro-Wilk W test for samples of size $N \leq 50$ [Shapiro and Wilk (1965)] (the Shapiro-Wilk W test for larger samples is given in [Royston (1982)]). It is generally recommended to use this set of tests all in conjunction since none of them is stronger than the other in detecting non univariate normality [D'Agostino and Stephen (1986), Looney (1995)].

Sample skewness and kurtosis parameters: Skewness is a parameter that describes the asymmetry of a random variable's probability distribution. This asymmetry is indicated by a positively skewed distribution or a negatively skewed distribution. The normal probability distribution has a zero skewness parameter showing that it is symmetric.

Let the random sample $X_1, X_2, \dots, X_n$; with sample mean $\bar{X} = (1/N) \sum_{j=1}^N X_j$, and sample variance $\bar{\sigma}^2 = 1/N \sum_{j=1}^N (X_j - \bar{X})^2$. The sample skewness may be defined as

$$b_1^{1/2} = \frac{(1/N) \sum_{j=1}^N (X_j - \bar{X})^3}{\bar{\sigma}^3} \quad (\text{D.1})$$

[D'Agostino and Stephen (1986)].

Kurtosis is a parameter that describes the shape of a random variable's probability distribution. In particular it is related to the peakedness and the tails density of the probability distribution. The normal probability distribution has a kurtosis parameter of 3. If the distribution's kurtosis is greater than 3 then it is said to be leptokurtic i.e., the distribution is peaked with fat tails. If the distribution's kurtosis is less than 3 then it is said to be platykurtic i.e., the distribution is less peaked with thin tails. The sample kurtosis may be defined as

$$b_2 = \frac{(1/N) \sum_{j=1}^N (X_j - \bar{X})^4}{\bar{\sigma}^4} \quad (\text{D.2})$$

[D'Agostino and Stephen (1986)].

The skewness and kurtosis parameters given above apply to any normal population, irrespective of its mean μ and its variance σ^2 .

Let the null hypothesis $\mathcal{H}_o$ be given by:

$$\mathcal{H}_o: X_1, \dots, X_n \text{ are distributed according to } \mathcal{N}(\mu, \sigma^2) \quad (\text{D.3})$$

where μ and σ^2 are unknown. We reject $\mathcal{H}_o$ if the calculated skewness parameter $b_1^{1/2}$ falls under the lower tail or the upper tail of the distribution of $b_1^{1/2}$, specified by the corresponding values of level of significance α , and sample size N .

We also reject $\mathcal{H}_o$ if the calculated kurtosis parameter b_2 falls under the lower tail or the upper tail of the distribution of b_2 , specified by α and N .

Shapiro - Wilk W test statistic: let the random sample be given by $X_1, X_2, \dots, X_N$; we denote its order statistics as $X_{(1)} < X_{(2)} < \dots < X_{(N)}$. Also let

$$S^2 = \sum_{i=1}^N (X_i - \bar{X})^2 \quad (\text{D.4})$$

Then we define

$$Y = \sum_{i=1}^r a_{N+1-i} (X_{(N+1-i)} - X_{(i)}) \quad (\text{D.5})$$

where $r = N/2$ when N is even; and $r = (N - 1)/2$ when N is odd; tabulated values for the coefficients $a_{(N+1-i)}$ are found in [Shapiro and Wilk (1965), D'Agostino and Stephen (1986)]. The Shapiro - Wilk W statistic is given by

$$W = Y^2/S^2 \quad (\text{D.6})$$

The distribution of W under the null hypothesis depends on N , but not on the parameters μ or σ^2 . We test the null hypothesis $\mathcal{H}_o$ given in (D.3) and we reject $\mathcal{H}_o$ if the values of W are less than the values specified by the lower tail of the W test statistic's distribution for the corresponding values of level of significance α , and sample size N .

Application of the tests: in the following we specify first the empirical distributions of the selected tests in order to be able to apply these tests to the measured data; we also specify the parameters required to find the adjusted level of significance as explained in Section

4.1.2 of Chapter 4. Then using the empirical distributions found previously we determine from each test statistic whether to reject or not reject the specified hypothesis $\mathcal{H}_o$. If $\mathcal{H}_o$ is not rejected for all the test statistics, then this will be an indication that the data analyzed comes from a normal population.

Some percentage points for the empirical distributions of skewness and kurtosis parameters are tabulated in [D'Agostino and Stephen (1986)]. Since neither the percentage points nor the sample size that we need in our analysis are specified in this table, we proceed to find the empirical distributions of each of these two parameters for normal samples. We then determine Monte Carlo percentage points of the skewness and kurtosis parameters for the sample size of interest. Figure C.4 shows this distribution using 1000000 simulation points, where we assumed normality and a sample size $N = 17$.

Also some tabulated values of W can be found in [Shapiro and Wilk (1965), D'Agostino and Stephen (1986)], however the percentage points of interest are not specified in the tabulated tables. We find the empirical distribution of the W test statistic in order to determine these percentage points. Our results are in good agreement with the results given in [Shapiro and Wilk (1965)]. Figure C.5 shows the empirical distribution of the W statistic using 1000000 simulation points, assuming normal variables of samples size $N = 17$.

The following are the parameters that we use when applying these tests. First of all, our results are based on a sample size of $N = 17$ i.e., there are seventeen MIMO channel measurements per building. We will determine the adjusted level of significance α for each of the tests based on the following parameter definitions: let $s = 2$ be the number of simultaneous tests which correspond to: skewness and kurtosis test; and Shapiro-Wilk W test; also let $F = 11$ be the number of independent frequencies over which we are going to perform the

test; let $w = 2$ correspond to the real and imaginary parts of the complex random variable; and let $m = 16$ be the number of channel elements of the channel matrix specified in (F.1) that will be simultaneously tested; finally let the value for the nominal level of significance be $\alpha^* = 0.1$.

Tests' results: we first evaluate the two univariate tests separately and then we give the overall results.

Results of skewness and kurtosis test: we use this test to evaluate whether each of the components of MIMO channel matrix $\mathbf{H}_f$ defined in (F.1) is normal; we test separately the real and imaginary parts of each of the random variables; We perform this test over $F = 11$ independent frequencies, and apply it to the real and imaginary parts of the sixteen transfer functions of the MIMO channel matrix. The adjusted significance level α for the skewness parameter is given by:

$$\alpha = \frac{\alpha^*}{m \times F \times w \times s} = \frac{0.1}{16 \times 11 \times 2 \times 2}$$

and for the kurtosis parameter is given by:

$$\alpha = \frac{\alpha^*}{m \times F \times w \times s} = \frac{0.1}{16 \times 11 \times 2 \times 2}$$

Figures D.1, D.2 show respectively the skewness and kurtosis test results vs. channel elements per building. The abscissa shows for each building each of the eleven measured frequencies. Each frequency consists of 32 elements of the MIMO channel (with the real and imaginary parts specified separately). There is a total of 352 elements per building. The horizontal lines in the figures are found from the empirical distributions given in Figure C.4. For the skewness test we obtain that $[c_{1-\alpha/2}, c_{\alpha/2}] = [-2.2807, 2.2909]$; for the kurtosis test

Building name	DU	LI	LO	SA
skewness	1	4	0	0
kurtosis	1	5	0	0
skewness or kurtosis	1	5	0	0

Table D.1: Number of rejections of the skewness and kurtosis tests with a nominal level of significance $\alpha^* = 0.1$

$[c_{1-\alpha}] = [8.6402]$. Any calculated values, of the corresponding tests, that lie outside these limits are rejected. We consider that the test fails whenever the skewness parameter or kurtosis parameter fail; we see in the figure that the points of the evaluated test that lie outside the boundaries correspond to *DU* and *LI*. We also show the maximum p-values corresponding to each receive area¹. We observe that for the receive areas of *DU* and *LI* the p-values are ($< \alpha$). The results of this test are summarized in Table D.1, where we present the name of the building and the corresponding number test rejections over a total number of eleven evaluated frequencies.

Results of Shapiro-Wilk test: we use this one-sided test in a similar manner as the skewness and kurtosis test. The adjusted level of significance is given by:

$$\alpha = \frac{\alpha^*}{m \times F \times w \times s} = \frac{0.1}{16 \times 11 \times 2 \times 2}$$

Again we apply this test to each the real and imaginary parts of the sixteen transfer functions of the MIMO channel matrix over the $F = 11$ frequencies. Figure D.3 shows the Shapiro-Wilk test vs. measured channel elements per building. The abscissa shows for each building each of the eleven measured frequencies. Each frequency consists of 32 elements of the MIMO channel (with the real and imaginary parts specified separately). There is a total of 352 elements per building. The thick horizontal line specifies the lower limit c_α that is

¹The p-values are evaluated according to the definition introduced in Section 4.2.2

Figure D.1: Skewness test results (32 variables and 11 frequencies) vs. channel elements per building. The horizontal lines represent $[c_{\alpha/2}, c_{1-\alpha/2}] = [-2.2807, 2.2909]$ with a nominal level of significance $\alpha^* = 0.1$.

Figure D.2: Kurtosis test results (32 variables and 11 frequencies) vs. channel elements per building. The horizontal line represents $[c_{1-\alpha}] = [8.6402]$ with a nominal level of significance $\alpha^* = 0.1$.

Figure D.3: Shapiro-Wilk W test results (32 variables and 11 frequencies) vs. channel elements per building. The horizontal line at 0.7312 is the lower limit for the one sided test with a nominal level of significance $\alpha^* = 0.1$. The calculated values of the test that lie below the lower limit are rejected.

determined from the empirical distribution of the Shapiro-Wilk W statistic given in Figure D.3. From this figure we observe that $c_\alpha = [0.7312]$. Any calculated value of the statistic that lies below this limit is rejected. Again, we see in the figure that the points that lie outside the boundary correspond to the cases of DU and LI . We also show in the figure the maximum p-values corresponding to each receive area². We note that these p-values are ($< \alpha$) for the cases of DU and LI . The results of this test are summarized in Table D.2, which presents the name of the building and the number of test rejections over the eleven evaluated frequencies. We observe that the rejected frequencies by this test coincide with the rejected frequencies by the skewness and kurtosis tests. Table D.3 shows the overall test results. We consider that the overall test fails for a given frequency when any of the previous tests fails at that frequency.

²The p-values were evaluated according to the definition introduced in Section 4.2.2

Building name	DU	LI	LO	SA
Shapiro-Wilk	1	5	0	0

Table D.2: Number of rejections of the Shapiro-Wilk W test with a nominal level of significance $\alpha^* = 0.1$

Building name	DU	LI	LO	SA
All tests	1	5	0	0

Table D.3: Number of rejections of all the tests with a nominal level of significance $\alpha^* = 0.1$

From these results we observe that in all the univariate tests that we used the maximum p-values per receive area were ($< \alpha$) for the cases of DU and LI . Therefore we reject the hypothesis of multivariate normality for the case of DU and LI since some of their evaluated test points lie outside their boundaries.

As we will confirm later in Chapter 5 the reason for rejecting multivariate normality assumption for DU and LI is because these two receive areas were the most affected by shadowing.

For the case of DU we think that shadowing is caused by two nearby buildings Azrieli Pavilion and Tory building (see Figure 3.1), which are obstructing the direct paths coming from the transmitter.

For the case of LI the amount of shadowing is more considerable. We think the reason is the large extension of the corridor where measurements were performed i.e., approximately 45 m, which contributed to the presence of different obstructing buildings in the transmit-receive paths of the different measured sub-areas. As an example we can see in Figure 3.1 that Dunton tower DU is clearly shadowing only one of the sub-areas of the library LI where measurements were performed. This is most likely causing a diffraction loss that is reflected by the considerable reduction of the mean receive power of that sub-area.

Appendix E

Bootstrap methods

A comprehensive introduction to bootstrap methods is given by the author in his book [Efron (1993)]. In the following we present with an example how bootstrap methods work and give a brief summary of the *BCa* confidence interval which is the accuracy measure we use in our analysis.

Let our data be drawn from a random sample $X(1), \dots, X(i), \dots, X(N)$ whose elements are *iid* according to an unknown probability distribution $\mathcal{F}$. Given that we observe the following sample $\mathbf{x} = [X(1) = x(1), \dots, X(N) = x(N)]$ we want to estimate a given parameter θ by using the estimator $\hat{\theta}$ and we want to determine how accurate it is the calculated estimate $\hat{\theta}(\mathbf{x})$ based on the observed data $\mathbf{x}$. We will refer to $\hat{\theta}(\mathbf{x})$ as the point estimate.

The most common measure of accuracy of the estimator is its associated standard error. However, it is not easy to obtain the standard errors of most of the estimators since there are no formulas available to compute them (excepting the standard error formula of the sample mean estimator) [Efron and Tibshirani (1986)]. In 1979 Efron introduced bootstrap as a computer method to estimate the standard error $\hat{se}$ of virtually any estimator with the

advantage that there is no formula required.

An important application of the estimated standard errors $\hat{se}$ is that they can be used to build standard confidence intervals which are approximate confidence intervals for θ and are given by

$$\theta \in [\hat{\theta} \pm \hat{se} z^\alpha] \quad (\text{E.1})$$

where z^α is the $(100 \alpha^{th})$ percentile point of the standard normal distribution. An interval with confidence level $x\% = 100(1 - 2\alpha)\%$ means that $x\%$ of the time a confidence interval built in this way will include the parameter of interest θ . A confidence interval together with the point estimate tells us what is the best guess for θ and how far in error that guess might reasonably be [Efron (1993)].

Standard confidence intervals work well when the number of samples is large; however when the number of samples is small estimated confidence intervals based on bootstrap methods have shown to be better approximations to the true intervals than standard confidence intervals. We are particularly interested in using *BCa* (bias-corrected and accelerated) intervals because of their good accuracy and transformation respecting properties. A small summary of *BCa* intervals is given at the end of this appendix.

Let any of the random observations given above, $X(i)$, be denoted by X with mean $E\{X\} = \mu_{\mathcal{F}}$ and variance $var\{X\} = \sigma_{\mathcal{F}}^2$. In this example we want to estimate the mean $\theta = \mu_{\mathcal{F}}$ by the sample mean $\hat{\theta} = 1/N \sum_{i=1}^N x(i)$. As the number of samples grow large the distribution of the sample mean is approximately normal with the same mean and variance divided by the number of samples i.e., $\hat{\theta} \sim \mathcal{N}(\mu_{\mathcal{F}}, \sigma_{\mathcal{F}}^2/N)$.

Let the random sample of size $N = 13$ be given by

[0.8315	-1.2291	1.0176	0.8362	0.1877	1.1722	0.0779
-1.6054	-0.5893	1.4661	-1.1532	0.5811	-0.1525]	

In this example we have assumed that the data comes from a standard normal population i.e., X is $\mathcal{N}(0, 1)$. The point estimate of the sample mean is $\hat{\theta} = 0.1108$, but we are interested to know how accurate this point estimate is with respect to θ . In what follows we describe the procedure we use to estimate the 90% confidence interval using *BCa* methods.

Let the bootstrap sample be given by $x^*(b)$ which corresponds to the b^{th} element of the data set $\mathbf{x}$ where b is a uniform random index number in the range of $[1, \dots, N]$. Let the bootstrap realization $\mathbf{x}^*$ be the set of N bootstrap samples $x^*(b)$ which are chosen by resampling uniformly with repetition the original data set $\mathbf{x}$. Let the bootstrap estimate $\theta^* = 1/N \sum_{i=1}^N x^*(i)$ be the sample mean for each bootstrap realization.

We are interested to find the 90% confidence interval $[\theta_{lo}, \theta_{up}]$ based on $B = 5000$ bootstrap realizations using the *BCa* method. We find this interval by using (E.2) and we find the standard errors of the endpoints of this interval using (E.7).

Figure E.1 shows the histogram of the bootstrap estimate θ^* using $B = 5000$ bootstrap sample realizations. We observe that the *pdf* of θ^* is very close to the theoretical *pdf* of the sample mean $\hat{\theta} \sim \mathcal{N}(0, 1/N)$ which is shown as the dotted line. The vertical dashed lines indicate the estimated 90% *BCa* confidence interval $\theta \in [-0.37, 0.53]$. The calculated standard errors of the endpoints are 0.09 and 0.08 respectively.

We can also find the standard 90% confidence interval given in (E.1) by using the unbiased estimate of the variance $\sigma_{\mathcal{F}}$ so that $\hat{se} = \sigma_{\mathcal{F}}/\sqrt{N} = 0.2769$ to obtain $\theta \in [-0.34, 0.57]$. We compare these intervals with respect to the exact 90% confidence interval $\theta \in [-0.35, 0.57]$.

Figure E.1: Histogram of the bootstrap sample mean (solid line) vs. theoretical *pdf* of the sample mean (solid line); estimated 90% *BCa* confidence interval for θ with $N = 13$ and $B = 5000$ (dashed line).

We observe in this example that both the standard interval and the *BCa* interval are good approximations to the exact interval. The reason standard intervals perform well here is that the normal approximation for the distribution of the sample mean is quite accurate. In general however when the distribution of the estimator is not necessarily symmetric *BCa* intervals will outperform standard intervals [Efron (1993)]. In the following we give a brief description of the *BCa* intervals presented in [Efron (1993)].

***BCa* intervals:** there are two kinds of *BCa* methods, one is the parametric and the other the non parametric method. We will be dealing with the non parametric method where we do not presuppose anything about the underlying distribution of the data. *BCa* methods estimate a true parameter by finding an approximate confidence interval, so from now on we will refer to these methods as *BCa* intervals. In this section we give a more detailed

description of *BCa* intervals along with some error measures that determine how close a *BCa* interval is to a true interval. The following are some of the characteristics of *BCa* intervals:

- (1) Transformation respecting: *BCa* intervals are transformation respecting i.e., they do not depend on the scale where the samples were measured. *BCa* intervals will automatically find the best scale to obtain the estimate.
- (2) Central confidence intervals: *BCa* intervals give a central confidence interval, which means that if the *BCa* interval is $[\theta_{lo}, \theta_{up}]$, a $(1 - 2\alpha)\%$ confidence interval will have α probability of not covering the true value of the parameter θ from above or below i.e., $P(\theta < \theta_{lo}) = \alpha$ and $P(\theta > \theta_{up}) = \alpha$.
- (3) Bias: *BCa* intervals are nearly unbiased.
- (4) Variability: *BCa* intervals may have considerable variability which depends not only on the number of samples but also on the number of bootstrap samples.
- (5) Accuracy: consider for convenience one of the endpoints of the *BCa* interval, $\theta[\alpha]$. Let $O(\cdot)$ denote the order of magnitude level, which is a scaled value proportional to $(\cdot)$; then we say that: an interval is first order accurate if $P(\theta < \theta[\alpha]) = \alpha + O(N^{-1/2})$; an interval is second order accurate if $P(\theta < \theta[\alpha]) = \alpha + O(N^{-1})$. *BCa* intervals are second order accurate.
- (6) Correctness: it evaluates how well the approximate endpoint of an interval $\hat{\theta}[\alpha]$ matches an exact endpoint of an interval $\Theta[\alpha]$. An interval is first order correct if $\theta[\alpha] = \Theta[\alpha] + O(N^{-1})$. An interval is second order correct if $\theta[\alpha] = \Theta[\alpha] + O(N^{-3/2})$. *BCa* intervals are second order correct.

Error assessment of BCa intervals: two important ways to assess the errors in the *BCa* intervals are given by the variance of each of the endpoints of the *BCa* interval itself; and also by the variance of the standard deviation of the *BCa* interval. Following we describe how to obtain these measures.

Variance of the BCa interval: let the estimate of the interval $[\Theta_{lo}, \Theta_{up}]$ be the *BCa* interval given by

$$[\theta_{lo}, \theta_{up}] = [\hat{\theta}^{*\alpha 1}, \hat{\theta}^{*\alpha 2}] \quad (\text{E.2})$$

where $\hat{\theta}^{*\alpha j}$ indicates the $100\alpha j^{\text{th}}$ percentile of B ordered bootstrap replications $\hat{\theta}^*(1), \hat{\theta}^*(2), \dots, \hat{\theta}^*(B)$; and

$$\alpha_1 = \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha)})} \right) \quad (\text{E.3})$$

$$\alpha_2 = \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha)})} \right) \quad (\text{E.4})$$

where $\Phi(x)$ is the *cdf* of the standard normal distribution at x and $\hat{z}_0$ is the bias parameter given by

$$\hat{z}_0 = \Phi^{-1} \left(\frac{\#\theta^*(b) < \hat{\theta}}{B} \right) \quad (\text{E.5})$$

where $\Phi^{-1}(x)$ is the inverse of the *cdf* of the standard normal distribution at probability x ; $\hat{\theta}$ is the original estimate of θ ; and $\hat{a}$ is the acceleration parameter given by

$$\hat{a} = \frac{\sum_{i=1}^N (\hat{\theta}_{(.)} - \hat{\theta}_{(i)})^3}{6 \left(\sum_{i=1}^N (\hat{\theta}_{(.)} - \hat{\theta}_{(i)})^2 \right)^{3/2}} \quad (\text{E.6})$$

where $\hat{\theta}_{(i)}$ is the estimate of the parameter θ where the i^{th} sample has been removed; and $\hat{\theta}_{(.)} = 1/N \sum_{i=1}^N \hat{\theta}_{(i)}$.

We estimate the variance of the first endpoint using the jackknife method as follows

$$var_{jack}(\hat{\theta}^{*\alpha 1}) = \frac{1}{N-1} \sum_{i=1}^N (\hat{\theta}^{*\alpha 1(i)} - \hat{\theta}^{*\alpha 1(\cdot)})^2 \quad (\text{E.7})$$

where $\hat{\theta}^{*\alpha 1(i)}$ is the estimate of the endpoint θ_{lo} with the i^{th} sample removed; $\alpha 1(i)$, $\bar{z}_{0(i)}$ and $\bar{a}(i)$ are defined in the same manner as the parameters given in (E.3), (E.5) and (E.6) respectively, with the i^{th} sample removed. The variance for the second endpoint is defined similarly.

Variance of the standard deviation of the BCa interval $var(\hat{se})$: the variability of $\hat{se}$ is used to determine the number of necessary bootstrap samples B . Let the estimate of the bootstrap standard error be given by

$$\hat{se} = \frac{\hat{\theta}^{*\alpha} - \hat{\theta}^{*(1-\alpha)}}{2z^\alpha} \quad (\text{E.8})$$

Then the estimated standard deviation of the *BCa* confidence interval $\hat{se}$ has a variance which depends on the sample size N and also on the bootstrap sample size B i.e.,

$$var(\hat{se}) \approx \frac{c1}{N^2} + \frac{c2}{N B} \quad (\text{E.9})$$

We calculate the variance of the $\hat{se}$ using the jackknife method as follows

$$var_{jack}(\hat{se}) = \frac{N-1}{N} \sum_{i=1}^N (\hat{se}_{(i)} - \hat{se}_{(\cdot)})^2 \quad (\text{E.10})$$

where $\hat{se}_{(\cdot)} = \sum_{i=1}^N \hat{se}_{(i)} / N$ and $\hat{se}_{(i)}$ is the estimated standard error of the sample without considering the i^{th} sample.

Appendix F

Estimation of the power ratio

Since we are interested to study the channel in frequency domain, we need to assure that it is a dual representation of the measured channel in time domain. To satisfy this property the channel needs to be zero mean wide sense stationary with uncorrelated scatterers (*wssus*) in both domains [Bello (1963)]. We guarantee the zero mean condition by performing the statistical analysis on the random component of the measured channel.

Also it is useful to know the power ratio between the random and fixed components of the channel since this will give an indication of the significance of our analysis. Next we give a mathematical description of the power ratio parameter we want to estimate as well as the description of the procedure we follow for its estimation.

Mathematical description of the power ratio: Let the MIMO channel matrix of the channel be denoted by $\mathbf{H}_f$ where f denotes the index for the independent frequency $f \in [1 \cdots F]$. The matrix channel elements are denoted by $H_{i,j,f}$, where $i \in [1 \cdots N_R]$ corresponds to the receiver index and $j \in [1 \cdots N_T]$ denotes the transmitter index, so we

write

$$\mathbf{H}_f = \begin{bmatrix} H_{1,1,f} & \cdots & \cdots & H_{1,N_T,f} \\ \vdots & \ddots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ H_{N_R,1,f} & \cdots & \cdots & H_{N_R,N_T,f} \end{bmatrix} \quad (\text{F.1})$$

Also let

$$\mathbf{h}_f = \text{vec}(\mathbf{H}_f) \quad (\text{F.2})$$

Now we define the mean vector

$$\mu_f = E\{\mathbf{h}_f\} \quad (\text{F.3})$$

and the channel covariance matrix as

$$\mathbf{R}_f = E\{\mathbf{h}_f \mathbf{h}_f^\dagger\} \quad (\text{F.4})$$

In particular if the components of the MIMO channel are assumed to be joint complex proper normal distributed, then we can obtain a complete characterization of the channel in terms of the spatial covariance matrix $\mathbf{R}_f$ and the mean vector μ_f for each of the independent frequencies.

Now we define the power of the deterministic component for each frequency component f as the square of the vector norm of the mean i.e.,

$$P_f = \|\mu_f\|_2^2 \quad (\text{F.5})$$

and the power of the random component for each frequency component f as the trace of the covariance matrix i.e.,

$$\tilde{P}_f = \text{Tr}(\mathbf{R}_f) \quad (\text{F.6})$$

We then obtain the power ratio between the fixed component and the variable component, for each frequency f , given by

$$\theta_f = \frac{P_f}{\tilde{P}_f} \quad (\text{F.7})$$

We analyze the power ratio θ_f for each frequency since the performance of the channel measured in terms of capacity is evaluated based on the power of the random component for each frequency.

Estimation of the power ratio using bootstrap: our measurements have the following specifications. The number of transmitters and receivers $N_T = N_R = 4$ respectively, the number of frequency components $F = 11$ ¹ and the number of independent samples $N = 17$. Let $\mathbf{H}_f(l)$ denote the l^{th} sample of the transfer function of the channel where $l \in [1 \cdots N]$, and $\mathbf{h}_f(l) = \text{vec}(\mathbf{H}_f(l))$. Also, let $\mathbf{W}_f = \langle \mathbf{h}_f(1), \dots, \mathbf{h}_f(N) \rangle$ be a 16×17 matrix for the f^{th} frequency. We then define the unbiased estimate of the sample covariance matrix as

$$\mathbf{S}_f = \frac{1}{N-1} \mathbf{W}_f \mathbf{W}_f^\dagger \quad (\text{F.8})$$

and the sample mean as

$$\mathbf{m}_f = \frac{1}{N} \sum_{l=1}^N \mathbf{h}_f(l) \quad (\text{F.9})$$

Let the estimate of the power ratio of (F.7) be given by $\hat{\theta}_f = \|\mathbf{m}_f\|_2^2 / \text{Tr}(\mathbf{S}_f)$ for each frequency $f \in [1, \dots, F]$. We want to know how accurate these estimates are. To this end we use the *BCa* method introduced in Appendix E to find the confidence interval for each θ_f as follows.

Let the 16×1 bootstrap sample for the f^{th} frequency be denoted by $\mathbf{h}_f(b)^*$ where b is a uniform random index number in the range of $[1, \dots, N]$. We form a data set $\mathbf{W}_f^*$ which

¹We use the estimated eleven independent frequency components found in Section 4.1 of Chapter 4

corresponds to $N = 17$ bootstrap samples $\mathbf{h}_f(b)^*$ that are chosen by resampling uniformly with repetition the columns of the original data set $\mathbf{W}_f$. The bootstrap estimate for the f^{th} frequency, $\hat{\theta}_f^*$, is determined from each bootstrap realization. Finally we use $B = 10000$ bootstrap realizations to estimate the 90% confidence interval for each θ_f by using Equation (E.2) of Appendix E.

Results of estimated confidence interval of the power ratio: we see in Figure F.1 that the estimated confidence interval of the power ratio θ_f varies for each frequency. For each of the buildings, which are arranged according to the separation distance with respect to the transmitter, we observe that in average the estimated confidence interval of θ_f tends to decrease as the distance between transmitter and receiver increases.

If we let $\theta_f^{dB} = 10 \log_{10}(\theta_f)$, then we see that the maximum upper limit for the confidence interval for θ_f^{dB} is about -10 for *DU* and it gradually decreases to a value of -12 for *SA*. The figure also shows that the bootstrap standard error estimate ($\hat{se}$) of θ_f , which is evaluated based on Equation (E.8) of Appendix E, has no significant variation which gives us confidence in the accuracy of these estimates.

Based on the estimated results we can see that for each frequency component at least 90% of the total power lies on the random component of the channel and it is worthwhile to focus our analysis on the random component of the channel.

Figure F.1: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for the power ratio θ_f (solid lines with markers) vs. frequency index, for each of the measured buildings; standard errors of the endpoints of the interval (dot markers)

Appendix G

Independence of transmit vectors

We want to evaluate the independence among the four 4×1 transmit vectors of interest $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$ and $\mathbf{h}_4$, which correspond to each of the four receive antennas at a given frequency. To this end we use the hypothesis $\mathcal{H}_0 : \Sigma_{ij} = \mathbf{0} ; i \neq j = [1 \cdots , q]$ given in Section 4.1.1.

The statistic that tests this hypothesis is given in (4.3) with tabulated points given in [Krishnaiah *et al.* (1976)]. Note that there are no tabulated points for our parameters of interest so that we find its empirical distribution with the following parameters: $q = 4$, which corresponds to the number of transmit random vectors, $p_i = p_0 = 4; i \in [1, \cdots , 4]$, which corresponds to the number of variables of each of the random vectors and $N = 17$, which is the number of independent samples.

Figure C.2 shows the empirical distribution of $-2\ln(\lambda)$ with 10^5 simulation points. The nominal level of significance that we use with this independence test is $\alpha^* = 0.05$. Since we apply this test over 11 independent frequencies simultaneously we adjust the level of significance to $\alpha = \alpha^*/11$ (see Section 4.1.2). From the empirical distribution given in Figure C.2 we obtain the following percentage point $c_{1-\alpha} = 38.0425$ which is the upper limit of the non

Figure G.1: Independent test statistic vs. channel element sample index. The thick horizontal line is the upper limit of the one sided test $c_{1-\alpha} = 38.0425$ with nominal level of significance $\alpha^* = 0.05$.

	Building name	LO	SA
$\alpha^* = 0.05$	Independence test	0	0

Table G.1: Number of rejections of independence test for transmit vectors with $\alpha^* = 0.05$

rejection region of the hypothesis $\mathcal{H}_0$.

Figure G.1 shows the evaluated test points using our data. The thick horizontal line represents the percentage point $c_{1-\alpha}$ mentioned above; any evaluated point above this limit will contribute to the rejection of $\mathcal{H}_0$. We see that all the frequency components of $[LO, SA]$ lie below $c_{1-\alpha}$. Furthermore we also show in the figure the maximum p-values for each receive area ¹. Table G.1 summarizes the results of this test. We observe that for both receive areas the p-values are ($> \alpha$) so that there is no evidence to reject $\mathcal{H}_0$.

¹The p-values were evaluated according to the definition introduced in Section 4.2.2

Appendix H

Principal components test - complex case

In this appendix we show the proof to find the variance of the statistic T_q which is given in (4.18). We will use a result on the distribution of the sample generalized variance of complex Gaussian random vectors given in [Goodman (1963)] and will follow similar steps as the ones given, on page 265, in [Anderson (1984)] to prove the following.

Let the sample generalized variance $\det(\mathbf{S})$ be defined as

$$\det(\mathbf{S}) = \det \left(\frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})^\dagger \right) \quad (\text{H.1})$$

[Anderson (1984)], where $\mathbf{x}_i$ are $p \times 1$ random vectors which are independent distributed, each according to $\mathcal{CN}(\mu, \Sigma)$.

We then obtain the following result

$$\sqrt{n} \left(\frac{\det(\mathbf{S})}{\det(\boldsymbol{\Sigma})} - 1 \right) \quad (\text{H.2})$$

is asymptotically normally distributed according to $\mathcal{N}(0, p)$.

Proof:

From [Anderson (1984)] we find that the distribution of the sample generalized variance is

$$\det(\mathbf{S}) \sim \det(\mathbf{A}) / (n)^p \quad (\text{H.3})$$

where $n = N - 1$ and $\mathbf{A}$ is given by

$$\mathbf{A} = \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^\dagger \quad (\text{H.4})$$

where $\mathbf{z}_1, \dots, \mathbf{z}_n$ are $p \times 1$ random vectors which are independent distributed, each according to $\mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$.

Let

$$\mathbf{z}_i = \mathbf{C} \mathbf{y}_i, \quad (\text{H.5})$$

where $i \in [1, \dots, n]$, $\mathbf{C}\mathbf{C}^\dagger = \boldsymbol{\Sigma}/2$, and $\mathbf{y}_i$ is a $p \times 1$ random vector where each vector is independent distributed according to $\mathcal{CN}_p(0, I)$, we then have

$$\mathbf{B} = \sum_{i=1}^n \mathbf{y}_i \mathbf{y}_i^\dagger \quad (\text{H.6})$$

We now can write $\det(\mathbf{A})$ as

$$\det(\mathbf{A}) = \det(\mathbf{C}) \det(\mathbf{B}) \det(\mathbf{C}^\dagger) = \det(\mathbf{B}) \det(\boldsymbol{\Sigma}) / 2^p \quad (\text{H.7})$$

Using (H.3) in (H.7), we then find that

$$\frac{\det(\mathbf{B})}{(2n)^p} \sim \frac{\det(\mathbf{S})}{\det(\mathbf{\Sigma})} \quad (\text{H.8})$$

We also know that

$$(2n)^p \frac{\det(\mathbf{S})}{\det(\mathbf{\Sigma})} \quad (\text{H.9})$$

is distributed as the product of p independent χ^2 chi-squared random variables with $2n, 2(n-1), \dots, 2(n-p+1)$ degrees of freedom respectively [Goodman (1963)].

We now proceed to find the asymptotic distribution of the sample generalized variance, we have that

$$\frac{\det(\mathbf{B})}{(2n)^p} = V_1(n) V_2(n) \cdots V_p(n) \quad (\text{H.10})$$

where $V_i(n)$ $i \in [1, \dots, p]$ are independent identically distributed and

$$2n V_i(n) \sim \chi_{2(n-p+i)}^2 \sim \sum_{\alpha=1}^{2(n-p+i)} W_{\alpha}^2 \quad (\text{H.11})$$

where W_{α} is $\mathcal{N}(0, 1)$; by the central limit theorem we have that

$$\frac{2n V_i(n) - 2(n-p+i)}{\sqrt{4(n-p+i)}} = \frac{\sqrt{n}(V_i(n) - 1 + (p-i)/n)}{\sqrt{(1 - (p-i)/n)}} \quad (\text{H.12})$$

is asymptotically distributed according to $\mathcal{N}(0, 1)$; and $\sqrt{n}(V_i(n) - 1)$ is asymptotically distributed according to $\mathcal{N}(0, 1)$.

Now let

$$\mathbf{u}(n) = \begin{bmatrix} V_1(n) \\ \vdots \\ V_p(n) \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad (\text{H.13})$$

so that $\sqrt{n}(\mathbf{u}(n) - \mathbf{b})$ has limiting distribution $\mathcal{N}(\mathbf{0}, \mathbf{I}_p)$.

Applying Theorem (4.2.3), (p. 121) [Anderson (1984)]¹, and letting

$$\frac{\det(\mathbf{B})}{(2n)^p} = f(u_1, \dots, u_p) = u_1 u_2 \cdots u_p,$$

$$\mathbf{T} = \mathbf{I}_p,$$

$$\phi_{b_i} = \frac{\partial f}{\partial u_i}|_{\mathbf{u}=\mathbf{b}} = 1 \quad \forall i, \text{ where } i \in [1 \cdots p],$$

so that $\phi_b = \mathbf{1}$, and

$$\phi_b^\dagger \mathbf{T} \phi_b = p,$$

we find that

$$\sqrt{n} \left(\frac{\det(\mathbf{B})}{(2n)^p} - 1 \right) \tag{H.14}$$

is $\mathcal{N}(0, p)$ distributed.

Finally from (H.8) we have

$$\sqrt{n} \left(\frac{\det(\mathbf{S})}{\det(\mathbf{\Sigma})} - 1 \right) \tag{H.15}$$

is $\mathcal{N}(0, p)$. In particular we get that

$$\sqrt{n} \left(\frac{l_i}{\lambda_i} - 1 \right) \tag{H.16}$$

is $\mathcal{N}(0, 1)$.

We use this result and follow same steps as the ones given in [Anderson (1984)] p.p. 474–475;

¹The theorem states that the limiting distribution of $\sqrt{n}(f[\mathbf{u}(n)] - f[\mathbf{b}])$ is $\mathcal{N}(\mathbf{0}, \phi_b^\dagger \mathbf{T} \phi_b)$ distributed; where $\mathbf{T}$ is the variance of the limiting distribution of $\sqrt{n}(\mathbf{u}(n) - \mathbf{b})$ and $\phi_{b_i} = \partial f(\mathbf{u}) / \partial u_i |_{\mathbf{u}=\mathbf{b}}$

writing (4.16) as a function of λ

$$f(\lambda) = \frac{\sum_{i=k+1}^m \lambda_i}{\sum_{i=1}^m \lambda_i} - h \quad (\text{H.17})$$

and using the fact that the partial derivatives of $f(\lambda)$ are given by

$$\frac{\partial f(\lambda)}{\partial \lambda_i} = -\frac{\lambda_{k+1} + \cdots + \lambda_m}{(\lambda_1 + \cdots + \lambda_m)^2}, \quad i = 1, \dots, k, \quad (\text{H.18})$$

and

$$\frac{\partial f(\lambda)}{\partial \lambda_i} = \frac{\lambda_1 + \cdots + \lambda_k}{(\lambda_1 + \cdots + \lambda_m)^2}, \quad i = k+1, \dots, m. \quad (\text{H.19})$$

We obtain the asymptotic generalized variance τ^2 given in (4.19).

Appendix I

Optimal power allocation algorithms

In the following we present the algorithms that implement the solutions presented in Section 5.3.2 of Chapter 5.

Outage capacity given a short-term input power constraint: for a set of L channel realizations $\{\Lambda_1^*, \dots, \Lambda_l^*, \dots, \Lambda_L^*\}$ ¹, a number of independent frequencies F , a power constraint P and a target rate R . We find the probability of outage P_{out} as follows:

¹ $\Lambda_l^* = \text{vec}(\Lambda_l)$ is a $Fm \times 1$ vector

```

step 1:    Find  $I_M$  for each channel realization
by water-filling over space and frequency:
     $\mapsto P_T = (FP)$ 
     $\mapsto$  initialize the following variables to zero
    number_trans_off,    number_trans_on
     $\mapsto$  for each channel realization  $l = 1, \dots, L$  do
         $\mapsto$  find  $\mu$  using water-filling algorithm
        with parameters  $(\Lambda_l^*, P_T)$ 
         $\mapsto$  calculate  $I_M$  with parameters  $(\Lambda_l^*, \mu)$ 
    step 2:    Turn transmission on or off depending
    on the channel realization:
         $\mapsto$  if  $(I_M \geq R)$   $\mapsto$  then
             $\mapsto$  number_trans_on ++
         $\mapsto$  else
             $\mapsto$  number_trans_off ++
         $\mapsto$  end
     $\mapsto$  end
step 3:    find outage probability:
     $\mapsto P_{out} = \text{number\_trans\_off} / L;$ 

```

Outage capacity given a long-term input power constraint: for a set of channel realizations $\{\Lambda_1^*, \dots, \Lambda_l^*, \dots, \Lambda_L^*\}$, a number of independent frequencies F , a power constraint P and a target rate R we find the probability of outage P_{out} as follows:

```

step 1:  calculate the minimum allocated power per channel to
avoid an outage:
    ↪  $P_T = (LFP)$ 
    ↪ for each channel realization  $l = 1, \dots, L$  do
        ↪  $P_{alloc}(l) = \text{call minpower}(\Lambda_l^*, R)$ 
    ↪ end
step 2:   select the set of good channels:
    ↪ sort in ascending order the allocated powers:
         $P_{sort} = \text{sort}(P_{alloc}) \uparrow$ 
    ↪ initialize variables: set to zero  $P_{cum}$  and set to one  $i$ 
    ↪ while ( $i \leq L$ ) and ( $P_{cum} \leq P_T$ )
        ↪ water-fill in time by allocating power to the good channels:
            ↪  $P_{cum} = P_{cum} + P_{sort}(i)$ 
            ↪  $i++$ 
        ↪ end
step 3:  calculate the threshold based on the set of good channels:
    ↪  $\varphi^* = P_{sort}(i)$ ;
step 4:  find outage probability:
    ↪  $P_{out} = (L - i)/L$ ;

```

minpower (Λ_1^*, R)

→ sort the eigenvalues of the channel in descending order:

$$\Lambda_{\text{sort}}^* = \text{sort}(\Lambda_1^*) \downarrow$$

step 1: tune value of water level μ over space and frequency of the channel:

→ set water level to cover the worst noise element of sorted channel:

$$\mu = 1/\lambda_{\text{sort}}^{\min}$$

→ find allocated power for all eigenvalues of the channel:

$$P_{\text{alloc}} = \left(\mu - 1/\Lambda_1^* \right)_+$$

→ find I_M with parameters (μ, Λ_1^*)

→ if $I_M > R$ then

→ lower the water level by discarding the smallest eigenvalue of the sorted channel Λ_{sort}^*

→ go to step 1

→ end

step 2: find exact water level value μ to get target rate R :

→ update channel Λ^* by discarding eigenvalue elements with zero allocated power

→ solve μ for the updated channel with parameters: (Λ_1^*, R)

→ find allocated power: $P_{\text{alloc}} = \left(\mu - 1/\Lambda_1^* \right)_+$

→ return: P_{alloc}

Appendix J

Accuracy of the capacity estimates

Variation of the capacity estimates based on 17 samples and the estimated eleven independent frequency sub-bands: We are interested to estimate the ergodic and outage rates with $P_{out} = 6\%$, under equal power allocation strategy, per receive area. To this end we use the bootstrap *BCa* method introduced in Appendix E to find the 90% confidence interval for these parameters.

Let $i(l)$ be the instantaneous mutual information at the l^{th} realization of the channel, given in (5.8), where the average is taken over $F = 11$ frequencies and where we assume equal power allocation strategy. Let $\mathbf{i} = [i(1), \dots, i(l), \dots, i(N)]$, where $N = 17$ corresponds to the number of measured locations per receive area. Also, let the estimate of the ergodic rate be given by $\hat{C}_{erg} = (1/N) \sum_{l=1}^N i(l)$ and let the estimate of the outage rate be given by $\hat{C}_{out,\alpha} = \mathbf{i}_\alpha$ where $\alpha = 0.06$ ($\mathbf{i}_\alpha$ corresponds to the α^{th} percentile point of $\mathbf{i}$). We want to know how accurate these estimates are by finding the *BCa* confidence intervals for C_{erg} and $C_{out,\alpha}$ as follows.

We denote the bootstrap sample by $i^*(b)$ which corresponds to the b^{th} element of the data

Figure J.1: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for C_{erg} and point estimate $\hat{C}_{erg}$ under equal power allocation, based on $F = 11$ eleven independent frequencies and $N = 17$ channel measurements

set $\mathbf{i}$ where b is a uniform random index number in the range of $[1, \dots, N]$. The bootstrap realization $\mathbf{i}^*$ is a set of N bootstrap samples $i^*(b)$ which are chosen by resampling uniformly with repetition the original data set $\mathbf{i}$. The bootstrap estimates $\hat{C}_{erg}^* = (1/N) \sum_{l=1}^N i(l)^*$ and $\hat{C}_{out,\alpha}^* = \mathbf{i}_\alpha^*$ are determined from each bootstrap realization. Finally we use $B = 10000$ bootstrap realizations to estimate the 90% confidence interval for C_{erg} and for $C_{out,\alpha}$ by using Equation (E.2) of Appendix E. We also evaluate the accuracy of these intervals of confidence by estimating the standard error of the endpoints using Equation (E.7) of Appendix E.

Figure J.1 shows the 90% confidence interval for the ergodic rates of $[DU, LI, LO, SA]$ respectively, based on $F = 11$ estimated eleven independent frequency sub-bands and $N = 17$ channel measurements. Similarly we show the 90% confidence intervals for the outage rates with $P_{out} = 6\%$ of the same receive areas $[DU, LI, LO, SA]$ in Figure J.2 respectively, with the same number of estimated independent frequency sub-bands F and the same number of measurements N . We also show in the figures the estimated standard errors (lower dotted lines) for each confidence limit. We first note that the standard errors of these interval limits are very small which shows the reasonably good accuracy of these confidence intervals.

We observe more variability of the outage rate estimates with respect to the ergodic rate estimates, which is shown by the larger confidence interval lengths of the former ones. We also observe more variability in the estimates of LI given by the larger confidence interval lengths of this receive area with respect to the ones of the other receive areas. Finally, we see that at high SNR the variations of the estimates are less significant since the confidence intervals are relatively small with respect to their rate estimates (dashed lines). We will use these confidence intervals in our analysis to compare the performance of the measured channels with the one of theoretical channels, and we will focus our analysis in the high SNR regime.

Figure J.2: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for $C_{out,0.06}$ and point estimate $\hat{C}_{out,0.06}$ under equal power allocation, based on $F = 11$ eleven independent frequencies and $N = 17$ channel measurements

Rate estimates under equal power allocation based on 11 independent frequencies vs. 1000 measured frequencies: We present in Figure J.3 the estimates of the 90% confidence intervals for the ergodic rates (based on the estimated eleven independent frequency sub-bands) vs. the ergodic rates based on the 1000 measured frequencies (solid lines with no marker) of $[DU, LI, LO, SA]$ respectively. Similarly we present in Figure J.4 the estimates of the 90% confidence intervals for the outage rates $P_{out} = 0.06$ (based on the estimated eleven independent frequency sub-bands) vs. the outage rates based on the 1000 measured frequencies (solid lines with no marker) of $[DU, LI, LO, SA]$ respectively.

As we can observe for all the cases the 90% confidence intervals for the rates of each of the measured areas include the corresponding estimated rate based on the 1000 measured frequencies. This confirms that the estimated eleven frequency sub-bands can be used as a good approximation to the 40 MHz bandwidth measured channels (given by the 1000 measured frequencies).

Figure J.3: Estimated 90% confidence interval $[\theta_{l0}, \theta_{up}]$ for C_{erg} under equal power allocation based on $F = 11$ eleven independent frequencies and ergodic rate point estimate $\hat{C}_{erg}$ under equal power allocation based on 1000 measured frequencies and $N = 17$ channel measurements

Figure J.4: Estimated 90% confidence interval $[\theta_{lo}, \theta_{up}]$ for $C_{out,0.06}$ under equal power allocation based on $F = 11$ eleven independent frequencies and outage rate point estimate $\hat{C}_{out,0.06}$ under equal power allocation based on 1000 measured frequencies and $N = 17$ channel measurements

Appendix K

Example: path loss and shadowing

We illustrate with an example the effect of path loss differences and shadowing on the *cdfs* of the channel capacities of the orthogonal space division multiplexing (*OSDM*) and time division multiplexing (*TDM*) downlink schemes presented in Chapter 6.

Problem specifications: let's consider a downlink channel where the base station equipped with two antennas transmits to two user terminals, where each terminal is also equipped with two antennas.

The channel of user $i \in [1, 2]$ is given by a $N_{R_i} \times N_T = 2 \times 2$ matrix $\mathbf{H}_i(n)$, where $n \in [1, \dots, N]$ indicates the index for the channel realization. The multi-user channel at the n^{th} channel realization is given by the $N_R \times N_T = 4 \times 2$ channel matrix denoted by $\mathbf{H}(n)$. For convenience, we also consider that the user channels are frequency flat, and that there is a short-term input constraint in the transmit power. For the case of *OSDM*, we assume that the transmitter and the receiver know the channel. For the case of *TDM*, we assume that the receiver knows the channel and that the transmitter has a limited channel knowledge so that it can adjust the average transmission allocation time for each user.

The objective is to find the maximum balanced user rates that these two schemes can deliver

for the following types of multi-user channels:

- (i) no path loss differences & no shadowing: entries of user channels $\mathbf{H}_1(n)$ and $\mathbf{H}_2(n)$ are distributed according to $\mathcal{CN}(0, 1)$
- (ii) path loss differences & no shadowing: user channels are the same as in item (i) but power of channel of user one, i.e., $p_1 = E\{Tr(\mathbf{H}_1(n)\mathbf{H}_1^\dagger(n))\}$, is attenuated by 10 dB with respect to that of user two
- (iii) shadowing & no path loss differences: user channels are the same as in item (i) and each user channel is affected independently by a shadowing factor s , where s is $\mathcal{LN}(0, \sigma)$
- (iv) path loss differences & shadowing: user channels are the same as in item (ii) and each user channel is affected independently by a shadowing factor s , where s is $\mathcal{LN}(0, \sigma)$

Path loss differences are indicated by the ratio between the two user powers $p_2/p_1(\text{dB})$. In our example we consider power ratios of $[0, 10]\text{dB}$.

When considering channels with shadowing we assume that each user is independently affected by a shadowing factor s , where s is $\mathcal{LN}(0, 3)$.

Formulation of the solution: for the case of *OSDM* schemes, we use transmit and receive beamforming by applying the block diagonal method with power control (*BD-PC*) explained in Sections 6.2, 6.3. For the case of *TDM* we use for the 2×2 channel of each user, the MIMO single-user channel capacity expression given by (5.8).

To deliver fair rates to both users in both schemes (see Section 6.4) we proceed as follows. In the *OSDM* scheme, we first find the allocated power for user one at the n^{th} channel realization, which is given by:

$$P_{T_1}(n) = \frac{P_T \lambda_2(n)}{\lambda_1(n) + \lambda_2(n)} \quad (\text{K.1})$$

where P_T is the total transmitted power, $\lambda_i(n)$ $i \in [1, 2]$ is the eigenvalue of

$$\hat{\mathbf{H}}_i(n)\hat{\mathbf{H}}_i^\dagger(n)$$

where $\hat{\mathbf{H}}_i(n)$ is the 1×2 effective non interfering channel of the i^{th} user after using $BD-PC$ at the n^{th} channel realization. The allocated power for user two at the n^{th} channel realization is $P_{T_2}(n) = P_T - P_{T_1}(n)$.

In the TDM scheme, we calculate the average time allocation as follows. Let μ be the fraction of transmission time user one uses and $(1 - \mu)$ the fraction of transmission time user two uses the channel. Then we solve

$$\mu = \frac{\alpha}{(1 + \alpha)} \quad (\text{K.2})$$

where $\alpha = E\{C_2(n)\}/E\{C_1(n)\}$, and where $C_i(n)$, $i \in [1, 2]$ is the channel capacity of the i^{th} user, at the n^{th} channel realization, with equal power allocation strategy.

Simulation results: in our results the overall SNR corresponds to the average signal-to-noise-ratio per receive antenna over all users, i.e., 4×2 multi-user channel which includes the effects of path loss differences and shadowing. Also, the SNR for user i , corresponds to the average signal-to-noise ratio per receive antenna of the i^{th} user.

As a reference we show the *cdf* of the cooperative multi-user channel, which we denote by (4×2) coop. This *cdf* of the sum capacity is obtained using the corresponding type of 4×2 multi-user channel described above, where we assume users can cooperate and transmitter uses equal power allocation (i.e., MIMO single-user capacity under equal power allocation).

We first show the results for the TDM scheme in terms of the *cdf* of the channel capacity, SNR per user and transmission time allocation per user. We then show the results of the

(a) no path loss differences & no shadowing

(b) path loss differences & no shadowing

(c) shadowing & no path loss differences

(d) path loss differences & shadowing

Figure K.1: *cdfs* of the channel capacities of *TDM* scheme when the channel is affected by path loss differences and or shadowing

OSDM scheme in terms of the *cdf* of the channel capacity, *SNR* per user, and average transmit power allocation per user.

***TDM* scheme:** Figure K.1 shows the *cdfs* of the channel capacities for each of the four cases mentioned above, at $SNR = 30$ dB.

Figure K.1 (a) shows the case when there are no path loss differences and there is no shadowing. (4×2) coop (thick dashed line) corresponds to the *cdf* of the sum capacity of the 4×2 multi-user channel assuming users can cooperate and using equal power allocation at the transmitter. This figure also shows the *cdf* of the sum channel capacity of *TDM* scheme that delivers fair rates to both users (thin dash-dotted line). Also as a reference, this figure shows the *cdf* of the channel capacity of *TDM* for each user (thin solid line and thick dotted line) assuming all transmission time is allocated to that user. Finally, this figure shows the *cdf* of the sum channel capacity of *OSDM* that delivers fair rates to both users (dotted line). When comparing the *cdfs* of the sum channel capacities of *TDM* and *OSDM* schemes, we observe that in terms of outage rates *TDM* is better than *OSDM* and that in terms of ergodic rates *OSDM* is better than *TDM*.

Figure K.1 (b) shows the *cdf* of the sum channel capacity of *TDM* when there are path loss differences and there is no shadowing. We observe that compared to the previous case, there is a larger difference between the *cdf* of the sum channel capacity of *TDM* and the *cdf* of the sum capacity of the cooperative multi-user channel (4×2) coop (thick solid line) due to the path loss differences. We see that if we transmit only to the strongest user, we get a similar performance to that of the cooperative multi-user channel (4×2) coop. We also note that the ergodic and outage rates of the *OSDM* scheme are more sensitive to these path loss differences than those of the *TDM* scheme.

Figure K.1 (c) shows the *cdf* of the sum channel capacity of *TDM* when there is shadowing and there are no path loss differences. (4×2) coop (thick dash-dotted line) corresponds to

the *cdf* of the sum capacity of this type of multi-user channel assuming users can cooperate and using equal power allocation. We observe that the outage rates of the *TDM* scheme are more sensitive to shadowing than those of the *OSDM* scheme.

Figure K.1 (d) shows the *cdf* of the sum channel capacity of *TDM* when there are path loss differences and there is shadowing. (4×2) coop (thick solid line) is the sum capacity for this type of multi-user channel assuming users can cooperate and that there is equal power allocation. Again we observe that the ergodic and outage rates of *TDM* are better than those of *OSDM*. We note however, that the *TDM* advantage in terms of outage rates for this case is slightly lower than that found for the case when there was no shadowing.

Tables K.1 and K.2 quantify the observations we have made. Table K.1 shows the percentage of the ergodic and outage sum rates with $P_{out} = 0.06$ of *TDM* with respect to that of the corresponding cooperative multi-user channel (4×2) coop. Table K.2 shows the percentage of the ergodic and outage sum rates with $P_{out} = 0.06$ of *OSDM* with respect to that of the corresponding cooperative multi-user channels (4×2) coop.

We see that when there are no path loss differences and there is no shadowing the ergodic sum rates of *OSDM* achieve 2% more than those of *TDM*. Also, the outage sum rates of *TDM* achieve 15% more than those of *OSDM*. We also see that both shadowing and path loss differences decrease the channel capacity of *TDM* and *OSDM* schemes. More specifically, we see that *OSDM* is more sensitive to path loss differences since the outage rate disadvantage of *OSDM* with respect to *TDM* increases even further. Also the ergodic sum rates of *OSDM* are now lower than those of *TDM*. If there is shadowing, we observe that it has a higher impact on the outage sum rates of *TDM* than on those of *OSDM*, since the outage sum rate advantage of *TDM* with respect to *OSDM* is slightly decreased.

We show in Figure K.2 the percentage of transmission time allocated to each user vs. the

Table K.1: % sum channel capacity of *TDM* with respect to the sum capacity of the cooperative multi-user channel (4×2) coop

σ (dB)	p_2/p_1 (dB)	<i>TDM</i> (ergodic)	<i>TDM</i> (outage)
0	0	84%	80%
0	10	79%	75%
3	0	82%	77%
3	10	76%	67%

Table K.2: % sum channel capacity of *OSDM* with respect to the sum capacity of the cooperative multi-user channel (4×2) coop

σ (dB)	p_2/p_1 (dB)	<i>OSDM</i> (ergodic)	<i>OSDM</i> (outage)
0	0	86%	65%
0	10	49%	29%
3	0	85%	66%
3	10	75%	50%

overall *SNR*. We show this for the four cases described above. We note that transmission times for the users are very similar when there are no path loss differences (left hand side). However these transmission times are very disparate when there are path loss differences (right hand side). Furthermore, we note in the latter case that when the overall *SNR* is low only a small fraction of transmission time is given to the strongest user. We also see that this fraction of allocated time to the strongest user becomes more significant when the overall *SNR* is high.

Figure K.3 shows the *SNR* per user of *TDM* vs. the overall *SNR*. We show the *SNR* per user, for the four cases described above. We observe that when there are no path loss differences (left hand side) the *SNRs* for each user are similar to the overall *SNR*. When there are path loss differences (right hand side) we observe that the *SNR* for user one is about 10 dB less than the *SNR* for user two, this happens since *TDM* does not compensate for this power difference directly, but instead it allocates a larger fraction of transmission time to user one.

Spatial multiplexing (*OSDM*) scheme: Figure K.4 shows the *cdf* of the sum channel capacity of *OSDM*. This figure is similar to Figure K.1 the only difference is that we show here, the *cdf* of the channel capacity of *OSDM* for each user.

Figure K.5 shows the transmit power allocations per user vs. the overall *SNR*. Again, we show this for the four cases of interest. We see that when there are no path loss differences (left hand side) the power allocations for both users are similar. When there are path loss differences (right hand side) we see that almost all the power is allocated to user one.

Finally, Figure K.6 shows the *SNR* per user vs. the overall *SNR*. We show this for each of the four cases given above. We observe that when there are no path loss differences (left hand side) the *SNRs* for each user are similar to the overall *SNR*. However when there are path loss differences (right hand side), we see that the *SNRs* for each user are about 5 dB lower than the overall *SNR*. This happens since most of the power is compensating for the path loss difference between users.

In particular when there are path loss differences in the measured channels we make the following observation. We see that in *OSDM*, both users are always operating at *SNR* which is ≈ 5 dB lower than the overall *SNR*. Also, we see that in *TDM*, the *SNR* of user two is ≈ 3 dB higher than the overall *SNR* and the *SNR* of user one is ≈ 7 dB lower than the overall *SNR*. Furthermore, we note that *TDM* operates at higher than the overall *SNR* during $\approx 40\%$ of the transmission time when the overall *SNR* is 30 dB. Since the linear growth in channel capacity of multiple antenna channels occurs at high *SNR*, only *TDM* scheme can exploit this linear growth in channel capacity during a fraction of time (which is significant when the overall *SNR* is high). This results in the superior ergodic and outage rates of the *TDM* scheme with respect to those of the *OSDM* scheme, when the overall *SNR* is high.

(a) no path loss differences & no shadowing

(b) path loss differences & no shadowing

(c) shadowing & no path loss differences

(d) path loss differences & shadowing

Figure K.2: TDM - time allocation per user vs. overall SNR when the channel is affected by path loss differences and or shadowing

(a) no path loss differences & no shadowing

(b) path loss differences & no shadowing

(c) shadowing & no path loss differences

(d) path loss differences & shadowing

Figure K.3: TDM - SNR per user vs. overall SNR when the channel is affected by path loss differences and or shadowing

(a) no path loss differences & no shadowing

(b) path loss differences & no shadowing

(c) shadowing & no path loss differences

(d) path loss differences & shadowing

Figure K.4: *cdfs* of the channel capacities of *OSDM* scheme when the channel is affected by path loss differences and or shadowing

(a) no path loss differences & no shadowing

(b) path loss differences & no shadowing

(c) shadowing & no path loss differences

(d) path loss differences & shadowing

Figure K.5: *OSDM* transmit power allocation per user vs. overall *SNR* when the channel is affected by path loss differences and or shadowing

(a) no path loss differences & no shadowing

(b) path loss differences & no shadowing

(c) shadowing & no path loss differences

(d) path loss differences & shadowing

Figure K.6: *OSDM* - *SNR* per user vs. overall *SNR* when the channel is affected by path loss differences and or shadowing

Bibliography

- Almers, P., Tufvesson, F., and Molisch, A. (2003). Measurement of keyhole effect in a wireless multiple-input multiple-output (*mimo*) channel. *IEEE Communications Letters*, 7(8), 373–375.
- Almers, P., Bonek, E., Burr, A., Debbah, M., Esposti, V., Hofstetter, H., Kyösti, P., Laurenson, D., Matz, G., Molisch, A., Oestges, C., and Özcelik, H. (2006). Survey of channel and radio propagation models for wireless MIMO systems. *EURASIP Journal on Wireless Communications and Networking*, 2007, 1–20.
- Anderson, T. (1984). *An introduction to multivariate statistical analysis*. John Wiley & Sons. Inc., New York, 2 edition.
- Baum, D., Gore, D., Nabar, R., Panchanathan, S., Hari, K., Erceg, V., and Paulraj, A. (2000). Measurement and characterization of broadband MIMO fixed wireless channels at 2.5 GHz. *International Conference in Personal Wireless Communications*, 1, 203–206.
- Bello, P. (1963). Characterization of randomly time-varying linear channels. *IEEE Trans. on Commun. Syst.*, 11, 360–393.
- Biglieri, E., Caire, G., and Giorgio, T. (2001). Limiting performance of block-fading channels with multiple antennas. *IEEE Trans. Inform. Theory*, 47, 1273–1289.

- Blum, R., Winters, J., and Sollenberger, N. (2001). On the capacity of cellular systems with MIMO. *Proc. IEEE Vehicular Technology Conference*, pages 1220–1224.
- Bonek, E. (2006). The MIMO radio channel. *Antenn06 Nordic Antenna Symposium*.
- Bonek, E., Herdin, M., Weichselberger, W., and Özcelik, H. (2003). MIMO - study propagation first. *Proc. IEEE Int. Symp. Signal Processing and Inf. Techn.*, pages 150–153.
- Caire, G. and Shamai, S. (2003). On the achievable throughput of a multi-antenna Gaussian broadcast channel. *IEEE Trans. Information Theory*, **49**, 1691–1706.
- Chizhik, D., Ling, J., P., W., Valenzuela, R., Costa, N., and Huber, K. (2003). Multiple-Input-Multiple-Output Measurements and Modeling in Manhattan. *IEEE Journal on Selected Areas in Communications*, **21**(3), 321–331.
- Chizhik, J., Wolniansky, R., Valenzuela, R., Costa, N., and K., H. (2002). Multiple input multiple output measurements and modeling in manhattan. *Proceedings of IEEE Vehicular Tech. Conf.*
- Chuah, C., Tse, D., Kahn, J., and Valenzuela, R. (2002). Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Transactions on Information Theory*, **48**, 637–650.
- Clerckx, B., Vanhoenacker, D., and Vandendorpe, L. (2003). Mutual coupling effects on the channel capacity and the space-time processing of MIMO communication systems. *ICC International Conference on Communications*, **4**, 2638–2642.
- D’Agostino, R. B. and Stephen, M. A. (1986). *Goodness of Fit Techniques*. Marcel Dekker, New York.
- Efron, B. (1979a). Bootstrap methods. Another look at the Jackknife. *The Annals of Statistics*, **7**, 1–26.

- Efron, B. (1979b). Computers and the theory of statistics: Thinking the unthinkable. *SIAM review*, **4**, 460–480.
- Efron, B. (1993). *Better bootstrap confidence intervals*. Chapman and Hall, 1 edition.
- Efron, B. and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statistical Science*, **1**(1), 54–75.
- Eklund, C., Marks, R., Stanwood, K., and S., W. (2002). IEEE Standard 802.16: A technical overview of the WirelessMAN air interface for broadband wireless access. *IEEE communications magazine*, pages 98–107.
- Foschini, G. (1996). Layered space-time architecture for wireless communication in a fading environment when using multiple-element antennas. *AT&T Bell Labs Tech Memo*.
- Foschini, G. and Gans, J. (1998). On limits of wireless communication in a fading environment when using multiple antennas. *Wireless Personal Commun.*, **6**(3), 311–335.
- Gesbert, D., H., B., and D., G. (2002). Outdoor MIMO wireless channels: Models and performance prediction. *IEEE Trans. Commun.*, **50**, 1926–1934.
- Gnanadesikan, R. (1986). *Methods for statistical data analysis of multivariate observations*. John Wiley & Sons, Inc., New York, second edition.
- Goldsmith, A. (1994). *Design and performance of high-speed communication systems over time-varying radio channels*. Thesis U.C. Berkeley.
- Goldsmith, A. (2003). *Class Reader EE359*. <http://wsl.stanford.edu/andrea/EE359Reader/>.
- Goodman, N., R. (1963). The distribution of the determinant of a complex Wishart distributed matrix. *The Annals of Mathematical Statistics*, **34**(1), 178–180.
- Haykin, S. (1994). *Communication Systems*. John Wiley & Sons. Inc., 3 edition.

- Hendrantoro, G. (2002). Outdoor-to-indoor path loss measurement and modelling. *Internal Technical Report Department of Systems and Computer Engineering*.
- Herdin, M., Özcelik, H., Hofstetter, H., and Bonek, E. (2002). Variation of measured indoor MIMO capacity with receive direction and position at 5.2 GHz. *Electronic Letters*, **38**(21), 1283–1285.
- Herdin, M., Czink, N., Özcelik, H., and Bonek, E. (2005). Correlation matrix distance, a meaningful measure for evaluation of non-stationary MIMO channels. *IEEE Vehicular Technology Conference*, **1**, 136–140.
- IEEE P802.16a, D.-. (2002). Draft amendment to IEEE standard for local and metropolitan area networks. *Part 16: Air interface for fixed wireless access systems. Medium access control modifications and additional physical layers specifications for 2 – 11 GHz*.
- Ivrlač, M. and Nossek, J. (2003). Quantifying diversity and correlation of Rayleigh fading MIMO channels. *Proc. IEEE Int. Symp. Signal Processing and Inf. Techn.*, pages 158–161.
- Ivrlač, M. and Nossek, J. (2005). Diversity and correlation in Rayleigh fading MIMO channels. *IEEE 61st Vehicular Technology Conference*, **1**, 151–155.
- Jafar, S. A. (2003). *Fundamental capacity limits of multiple antenna wireless systems*. Ph. D. Thesis, Stanford University.
- Jindal, N. and Goldsmith, A. (2005). Dirty-paper coding versus TDMA for MIMO broadcast channels. *IEEE Trans. Information Theory*, **51**(5), 1783–1794.
- Kalbfleisch, J. (1985). *Probability and statistical inference*, volume 2. Springer texts in statistics, 2 edition.
- Kay, S. (1998). *Fundamentals of Statistical Signal Processing - Detection theory*, volume 2. Prentice Hall, 1 edition.

- Kermoal, J., schumacher, L., Pedersen, K., Mogensen, P., and Frederiksen, F. (2002). A stochastic MIMO radio channel model with experimental validation. *IEEE Journal on Selected Areas in Communications*, **20**(6), 1211–1225.
- Krishnaiah, P., Lee, J., and Chang, T. (1976). The distributions of the likelihood ratio statistics for tests of certain covariance structures of complex multivariate normal populations. *Biometrika*, **63**(3), 543–549.
- Looney, S. (1995). How to use tests for univariate normality to assess multivariate normality. *The American Statistician*, **49**, 64–70.
- Mazzarese, D. and Krzymień, W. (2003). Throughput maximization and optimal number of active users on the two transmit antenna downlink of a cellular system. *IEEE Pacific RIM Conference on Communications, Computers and Signal Processing*, pages 498–501.
- Medbo, J., Harrysson, F., Asplund, H., and Berg, J. (2004). Measurements and analysis of a MIMO macrocell outdoor-indoor scenario at 1947 mhz. *IEEE Vehicular Technology Conference*, **1**, 261–265.
- Mohseni, M., Zhang, R., and Cioffi, J. (2006). Optimized transmission for fading multiple-access and broadcast channels with multiple antennas. *IEEE Journal of Selected Areas in Communications*, **24**(8), 1627–1539.
- Muirhead, R. (1982). *Aspects of multivariate statistical theory*. John Wiley & Sons. Inc.
- Nabar, R., Bölcskei, H., and Paulraj, A. (2004). Diversity and outage performance in Ricean MIMO channels. *IEEE Trans. Wireless Communications*, **4**(5), 2519–2532.
- Neeser, F. D. and Massey, J. L. (1993). Proper complex random processes with applications to information theory. *IEEE Trans. on Information Theory*, **39**(4), 1293–1302.

- Nguyen, H., Andersen, J., Heddergott, R., and Tschudin, M. (2004). Characterization of the indoor/outdoor to indoor MIMO radio channel at 2.44 ghz. *COST273 meeting, Gothenburg, Sweden*, pages 1–17.
- Oestges, C. and Erceg, V. Paulraj, A. (2003). A physical scattering model for MIMO macro-cellular broadband wireless channels. *IEEE Journal on Selected Areas in Communications*, **21**(5), 721–729.
- Özcelik, H., Herdin, M., Weichselberger, W., Wallace, J., and Bonek, E. (2003a). Deficiencies of Kronecker MIMO radio channel model. *IEEE Electronics Letters*, **39**(6), 1209–1210.
- Özcelik, H., Herdin, M., Prestros, R., and Bonek, E. (2003b). How MIMO capacity is linked with multipath distribution. *International Conference on Electromagnetics in Advanced Applications*, pages 775–778.
- Özcelik, H., Czink, N., and Bonek, E. (2005). What makes a good MIMO channel Model? *Proc. of the 61st Semiannual Vehicular Technology Conference*, pages 1–6.
- Özdemir, M., Arvas, E., and Arslan, H. (2004). Dynamics of spatial correlation and implications on MIMO systems. *IEEE Radio Communications*, pages 14–19.
- Paulraj, A., Nabar, R., and Gore, D. (2003). *Introduction to Space-Time Wireless Communications*. Cambridge.
- Paulraj, A., D., G., R., N., and H., B. (2004). An overview of MIMO communications - a key to gigabit wireless. *Proceedings of the IEEE*, **92**, 198–218.
- Proakis, J. (2001). *Digital Communications*. Mc Graw Hill, 4 edition.
- Rappaport, T. (1996). *Wireless communications principles and practice*. Prentice Hall, 1 edition.

- Rhee, W., Yu, W., and Cioffi, J. (2004). The optimality of beamforming in uplink multiuser wireless systems. *IEEE Transactions Wireless Communications*, **3**(1), 86–96.
- Royston, J. P. (1982). An extension of shapiro and Wilk's W test for normality to large samples. *Applied Statistics*, **31**, 115–124.
- Shapiro, S. and Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika*, **52**, 591–611.
- Small, N. J. H. (1980). Marginal skewness and kurtosis in testing multivariate normality. *Applied Statistics*, **29**(1), 85–87.
- Sousa, E., Jovanović, V., and Daigneault, C. (1994). Delay spread measurements for the digital cellular channel in Toronto. *IEEE Transactions on Vehicular Technology*, **43**(4), 837–847.
- Spencer, Q., Swindlehurst, A., and Haardt, M. (2004). Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *IEEE Transactions on Signal Processing*, **52**(2), 461–471.
- Stuber, G. (2001). *Principles of mobile communication*. Kluwer Academic Publishers, 2 edition.
- Telatar, E. (2003). *Wireless Communications and Mobility*. <http://www.epfl.ch/WCM/>.
- Telatar, I. (1995). Capacity of multi-antenna Gaussian channels. *AT&T Bell Labs Tech Memo.*, pages 1–28.
- Tse, N. and Hanly, S. (1998). Multiaccess fading channels - Part I: Polymatroid structure, optimal resource allocation and throughput capacities. *Transactions on Information Theory*, **44**(7), 2796–2815.

- Vishwanath, S. (2003). *Multi-user multiple antenna systems: theoretical limits and practical strategies*. Thesis Stanford University.
- Vishwanath, S., Jindal, N., and Goldsmith, A. (2003). Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. *IEEE Transactions on Information Theory*, **49**(10), 2658–2668.
- Viswanath, P. and Tse, D. (2003). Sum capacity of the vector Gaussian broadcast channel and uplink downlink duality. *IEEE Transactions on Information Theory*, **49**(8), 1912–1921.
- Weingarten, H., Steinberg, Y., and Shamai, S. (2005). The capacity region of the Gaussian MIMO broadcast channel. *Submitted for publications to IEEE Transactions on Information Theory*.
- Winters, J. (1987). On the capacity of radio communication systems with diversity in a Rayleigh fading environment. *IEEE Journal on Selected Areas in Communications*, **5**(5), 871–878.
- Wyne, S., Almers, P., Eriksson, G., Karedal, J., Tufvesson, F., and Molisch, A. (2004). Outdoor to indoor office MIMO measurements at 5.2 ghz. *IEEE Vehicular Technology Conference*, **1**, 101–105.
- Wyne, S., Molisch, A., Almers, P., Eriksoon, G., Karedal, J., and Tufvesson, F. (2005). Statistical evaluation of outdoor-to-indoor office MIMO measurements at 5.2 ghz. *IEEE Vehicular Technology Conference*, **1**, 146–150.
- Yépez, E., Falconer, D., and Bultitude, R. (2006a). Estimated capacities for a multiple input multiple output system based on outdoor to indoor channel measurements. *IEEE European Conference on Antennas and Propagation*.

- Yépez, E., Falconer, D., and Bultitude, R. (2006b). Estimating spatial correlations of the channel elements of a multiple input multiple output system based on outdoor to indoor channel measurements. *ANTEM/URSI 12th International Symposium on Antenna Technology and Applied Electromagnetics in Canadian Radio Sciences*.
- Yu, W. and Cioffi, J. (2001). Trellis precoding for the broadcast channel. *IEEE Global Telecommunications Conf. (GLOBECOM)*, pages 1344–1338.
- Yu, W. and Rhee, W. (2006). Degrees of freedom in multi-user spatial multiplex systems with multiple antennas. *IEEE Transactions on Communications*.
- Zoubir, A. and Boashash, B. (1998). The Bootstrap and its application in signal processing. *IEEE Signal Processing Magazine*, 7, 56–76.