THESE

présentée pour obtenir le grade de

Docteur de l'Ecole Nationale Supérieure des Télécommunications

Spécialité: Communication et Electronique

Maxime Guillaud

Techniques de transmission et de modélisation de canal pour les systèmes de communications multi-antennes

Thèse soutenue le 8 Juillet 2005, devant le jury composé de :

Pierre COMON	Président
Jean-François HELARD	Rapporteurs
Laurent KOPP	
Marie-Laure BOUCHERET	Examinateurs
Markus RUPP	
Bernard FLEURY	Invité
Dirk T.M. SLOCK	Directeur de thèse

THESIS

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from Ecole Nationale Supérieure des Télécommunications

Specializing: Communication and Electronics

Maxime Guillaud

Transmission and Channel Modeling Techniques for Multiple-Antenna Communication Systems

Defended on July 8, 2005, before the committee composed by:

Pierre COMON	President
Jean-François HELARD	Readers
Laurent KOPP	
Marie-Laure BOUCHERET	Examiners
Markus RUPP	
Bernard FLEURY	Invited examiner
Dirk T.M. SLOCK	Thesis supervisor

Acknowledgements

I wish to thank all the people who got me interested in science, and who pushed me towards research in general. In particular, I thank Prof. Inbar Fijalkow for her role in putting me in contact with research in the first place. I am also extremely grateful to Prof. Markus Rupp for his influence on my decision to start this PhD, and would like to thank him for accepting to be on my thesis committee.

Many thanks go to my advisor Prof. Dirk Slock, for his creativity, enthusiasm, and patience, during our collaboration. I greatly appreciated the freedom that I enjoyed while pursuing my research during those four years.

I wish to express my gratitude to Prof. Jean-François Hélard and Dr. Laurent Kopp for accepting to read the thesis and for their feedback on my work. I also wish to thank Prof. Marie-Laure Boucheret, Prof. Pierre Comon, and Prof. Bernard Fleury for their participation in my committee.

I would also like to extend my warmest thanks to my colleagues and friends at Eurécom: the international atmosphere there made my time working with them truly enjoyable. Among these, I would like to thank Souad Guemghar, Navid Nikaein, Abdelkader Medles, Albert Guillén y Fàbregas, David Mary, Saad Kiani, Marios Kountouris, for their friendship, and to extend a special thank to Mari Kobayashi for her role as the office's entropy source. I also with to thank Mérouane Debbah and David Gesbert for their advice regarding my work and for their friendship.

This paragraph would of course be incomplete if I didn't mention Magali, for her patience and support during the preparation of this thesis and of the manuscript, and whose presence makes my life so much more enjoyable.

Finally, I would like to thank my parents for their love, care and education, which shaped my mind in many ways.

Abstract

This thesis deals with different aspects of multiple-antenna (MIMO) wireless communications. In a first part, we introduce Space-Time-Frequency Spreading (STFS), a space-time code for Orthogonal Frequency Division Multiplexing (OFDM) systems that maximally exploits the channel space and frequency diversity, as well as time diversity. For low spectral efficiencies (e.g. BPSK), Viterbi-based decoding of STFS associated with iterative interference cancellation provides the same performance as BICM without requiring soft decoding.

In a second part, the importance of having channel state knowledge at the transmitter is discussed. A channel reciprocity model is introduced for the case of Time-Division Duplex (TDD) systems, which models the impairments of the radio-frequency components with linear filters. After a collaborative training phase (relative calibration), this model enables the transmitter to infer the downlink channel impulse response from the uplink channel estimates, thus lifting the requirement for continuous feedback. The frequency-selective reciprocity model was experimentally validated for SISO channels.

Finally, the problem of modeling the temporal evolution of MIMO frequency-selective channels is addressed. A pathwise model is introduced, and we propose to use a blind method to decompose the time-varying channel realizations into, for each path, a set of constants representing the physical characteristics of the environment, and a time-varying, structured process (such as a Doppler series) that can be easily tracked or predicted. The performance of this method is evaluated by simulations, using both synthetic and experimental data.

Résumé

Dans cette thèse, différents aspects des communication sans fils multi-antennes (MIMO) sont abordés. Dans un premier temps, nous introduisons l'Etalement Temps-Espace-Fréquence (STFS), un code destiné aux systèmes à multiplexage fréquentiel orthogonal (OFDM), qui maximise l'exploitation de la diversité spatiale, fréquentielle et temporelle. Pour des efficacités spectrales faibles (par exemple utilisant des BPSK), le décodage du code STFS par l'algorithme de Viterbi, associé à une annulation d'interférence itérative atteint les mêmes performances qu'un code BICM.

Dans un second temps, nous considérons l'importance pour l'émetteur de connaître l'état du canal. Un modèle de réciprocité est proposé pour les systèmes à duplexage temporel (TDD), dans lequel les imperfections des composants radio-fréquence sont représentées par des filtres linéaires. A l'issue d'une phase d'apprentissage coopératif (calibration relative), ce modèle permet à l'émetteur de déduire l'état du canal descendant à partir de l'estimation de l'état du canal montant, sans nécéssiter de retour continu d'information. Ce modèle a été validé expérimentalement dans le cas d'un canal SISO.

Enfin, la modélisation de l'évolution temporelle des canaux MIMO sélectifs en fréquence est envisagée. Un modèle par trajets est utilisé, et une méthode d'identification aveugle est proposée, qui permet de décomposer les variations du canal, pour chaque trajet identifié, en un ensemble de constantes représentant les caractéristiques physiques de l'environnement, et un processus variable dans le temps, structuré (tel que l'effet Doppler), et pouvant être facilement suivi ou prédit. Les performances de cette méthode sont analysées sur des signaux de synthèse et mesurés.

Contents

List of Figures

List of Tables

Acronyms

Here are the main acronyms used in this document. The meaning of an acronym is usually indicated once, when it first occurs in the text. The english acronyms are also used for the french summary.

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BICM	Bit-Interleaved Coded Modulation
CDF	Cumulative Distribution Function
CP-OFDM	Cyclic Prefix Orthogonal Frequency Division Multiplexing
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DFT	Discrete Fourier Transform
DoA/DoD	Direction of Arrival/Departure
DS	Direct-Spread
FIR	Finite Impulse Response
IC	Interference Cancellation
ICI	Inter-Carrier Interference
iff	if and only if
i.i.d.	independent and identically distributed
ISI	Inter-Symbol Interference
LCP	Linear Constellation Precoding
LLR	Log Likelihood Ratio
LMMSE	Linear Minimum Mean-Square Error
LSTBC	Linear Space-Time Block Code
LTI	Linear Time-Invariant
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output

ML	Maximum Likelihood
MRC	Maximum Ratio Combining
NSTC	Natural Space-Time Code
OFDM	Orthogonal Frequency Division Multiplexing
PEP	Pairwise Error Probability
PIC	Parallel Interference Cancellation
RF	Radio-Frequency
r.v.	Random Variable
Rx	Receiver
SER	Symbol Error Rate
SIMO	Single-Input Multiple-Output
SIR	Signal-to-Interference Ratio
SISO	Single-Input Single-Output
sIsO	Soft-Input Soft-Output
SNR	Signal-to-Noise Ratio
SOVA	Soft-Output Viterbi Algorithm
S/P	Serial-to-Parallel Conversion
s.t.	such that
STC	Space-Time Code
STFC	Space-Time-Frequency Code
STFS	Space-Time-Frequency Spreading
STLS	Structured Total Least-Squares
SVD	Singular Value Decomposition
TSTC	Threaded Space-Time Code
Tx	Transmitter
UMTS	Universal Mobile Telecommunications Services
WLAN	Wireless Local Area Network
w.l.o.g.	without loss of generality
w.r.t.	with respect to
WSS	Wide-sense Stationary

Notations

$\mathrm{E}_x[\cdot]$	expectation operator taken over variable x
$(\cdot)^*$	complex conjugate operator
$(\cdot)^T$	transpose operator
$(\cdot)^H$	Hermitian transpose operator (transpose + complex conjugate)
$\odot$	componentwise product
$\otimes$	Kronecker product
$\{\cdot\}_+$	causal part of a z-function
$[\cdot]_{i,j}$	the $(i,j)^{th}$ element of a matrix
$*$	convolution operator
$\sim$	distributed as
$\mathcal{N}(a,\sigma^2)$	normal (Gaussian) distribution of mean a and variance σ^2
$\mathcal{CN}(a,\sigma^2)$	complex normal distribution of mean a and variance σ^2
$\|\|\cdot\|\|_F$	Frobenius norm
$\mathrm{diag}(\underline{\mathbf{v}})$	square matrix with the coefficients of vector $\underline{\mathbf{v}}$ on its diagonal
$\mathbf{1}_n$	vector containing n ones
N_t	number of transmit antennas
N_r	number of receive antennas
$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
$\mathbb{Z}$	set of integers
$[\![m,n]\!]$	set of integers from m to n, *i.e.* $[m,n]\cap\mathbb{Z}$

When there is no ambiguity, time-domain and frequency-domain representations of linear filters are used interchangeably. For instance, $s(t)$ and $s(f)$ denote respectively the impulse-response and the frequency response of the same filter s.

Chapter 1

Introduction and thesis outline

The issues associated with reliable communication over wireless channels have dramatically evolved in the recent years, due to a shift in the field of application of wireless communication systems. Although the history of wireless communications is over a century old, the introduction of consumer-grade portable communication equipment, such as mobile phones, or computer Local Area Networks (LANs) is fairly recent. The huge market demand for these applications has oriented research towards improved spectral efficiency (Spectral efficiency, measured in bits per second per Hertz, or bits per channel access, is a measure of the amount of information that can be transmitted using a given amount of energy). Current examples of this trend can be found in the context of Wireless Local Area Network (WLAN) applications, such as IEEE 802.11a/802.11g, or ETSI HiperLAN/2. The continuous growth of computer network traffic has pushed the design of modern WLAN protocols towards higher throughputs.

Besides increasing spectral efficiency, the use of multiple antennas has been heralded as an almost-free solution to throughput problems, since the channel capacity scales linearly with the number of antennas for a given total transmit power, whereas increasing the transmit power only yields a logarithmic capacity improvement. Multiple antenna, a.k.a. Multiple-Input Multiple-Output (MIMO) systems, have been widely studied in the

last decade, after the initial interest was initiated at Bell Labs by the theoretical results of Winters and Telatar on mutual information [1, 2], quickly followed by the first practical MIMO coding scheme, proposed by Foschini [3].

The counterpart to the increased spectral efficiency is that it uncovers new challenges that were previously hidden by the suboptimal use of the channel resource:

exploiting diversity: exploiting channel diversity provided by the increased bandwidth and the introduction of multiple antennas

equalization of fast-fading channels: tracking and equalizing the time-varying channels encountered in mobile applications has become more difficult as the amount of data needed to accurately represent the channel state has increased

This thesis is comprised of two parts, following a brief introduction to the issues of communications over a noisy wireless channel in Chapter 2.

In the first part of this thesis, we will address the specific issues of Orthogonal Frequency Division Multiplexing (OFDM), a transmission technique consisting in decomposing a wideband channel into a set of discrete, narrowband channels. Besides opening the way to multiplexing, OFDM has interesting properties: the well-known fact that it turns a frequency-selective channel into a set of flat-fading channel was already noticed by Shannon in his 1949 paper [4].

However, the need to deal with fast channel fading and to exploit frequency diversity both reintroduce the need to consider all OFDM subbands jointly, which runs counter to the principle of dealing separately with the data streams on each subcarrier, thus making the advantage of OFDM over direct-spread (DS) techniques less obvious. However, the evolution of current OFDM-based standards mandates to find solutions to these issues.

Chapter 3 addresses the problem of equalization of OFDM transmission in the high-Doppler spread case. We outline a duality between time-domain and frequency-domain equalization, and show that the Doppler spread in the frequency domain plays the same role as delay-spread in the time-domain. This work was published in [5].

In Chapter 4, we explore the linear precoding technique as an alternative to binary codes as a way to exploit channel diversity. We introduce the new

STFS code for OFDM-based MIMO systems, and the matching iterative decoding technique, and show that STFS achieves good space, time and frequency diversity exploitation within a reasonable decoding complexity. This work was published in [6, 7, 8], and used in the RNRT Antipode deliverables [9, 10, 11, 12, 13].

In the second part of this thesis, we deal with the issues of channel representation, tracking and estimation, for MIMO systems in the general case of frequency-selective channels. First, in Chapter 5, we propose an information-theoretic analysis of the importance for the transmitter of knowing the channel. We outline the situations in which knowledge of the channel by the transmitter can really improve the performance of a communications system, and study the effects of imperfect (noisy) channel knowledge and imperfect calibration.

In Chapter 6, we introduce a new model for channel reciprocity in TDD systems, that lets the transmitter gain knowledge of the downlink channel properties from uplink channel estimates. We term this method *relative calibration*, and present experimental evidence that the proposed reciprocity model holds for real systems. This work was published in [14], and is the subject of a patent [15].

Finally, in Chapter 7, we propose to use a pathwise model to represent the time-variations of MIMO frequency-selective channels, and develop a a blind method to decompose the channel variations into a low number of spatial and temporal parameters associated to each path. This work was published in [16, 17, 18].

Chapter 2

Digital communications over noisy MIMO fading channels

In this chapter we introduce the general framework of wireless communications, i.e. coding to communicate over a noisy channel, and lay the notations for the multipath, discrete time channels that will be considered throughout this thesis. The model is extended to MIMO AWGN and fading channels, for both frequency-flat and frequency selective cases, and basic capacity formulas corresponding to those models are recalled.

2.1 Introduction

Reliable communication of digital information over noisy channels, such as those encountered in wireless applications, involves the use of a channel encoder and its corresponding decoder. Those are usually designed with a particular channel model in mind. The relevant parameters were defined by Shannon as early as 1948 [19], for the case of infinitely long signals. Those parameters are:

The code rate R, measured in bits per second, is the amount of information that can be transmitted for a given energy expense.

The channel capacity C, in bits per second, is the maximum rate of any code that can reliably (with probability 1) transmit information over this channel.

The normalized versions of the rate and the capacity (in bits per channel access) are often used to compare codes. The infinite coding length and the complexity of Maximum-Likelihood (ML) decoding implied by information theory are not practical for real-life implementations, and we can only try to minimize the probability of erroneous decoding. However, as we shall see, the notion of mutual information for block-fading channels will provide an good insight on the properties of certain coding methods.

2.2 SISO channel model

2.2.1 Continuous time model

Let us consider the general case of the transmission of a complex signal over a fast-varying Single-Input Single-Output (SISO) channel, *i.e.* from one transmit to one receive antenna. The signal is defined in discrete-time by the complex time series $\{x_k\}_{k\in\mathbb{Z}}$, and is transmitted at the symbol rate T_s. The transmitted signal can be represented (using the distribution notation) as

$$x(t) \triangleq \sum_{l=-\infty}^{+\infty} x_l.\delta_{lT_s}(t), \tag{2.1}$$

where $\delta_{t_0}(t)$ is Dirac's delta function. In order for the transmitted signal to remain band-limited, a pulse-shaping filter is used at the transmitter. Denoting $p_t(\tau)$ its impulse response, the baseband signal that actually gets transmitted is $s \triangleq p_t * x$, or

$$s(t) = \sum_{l=-\infty}^{+\infty} x_l p_t(t - lT_s). \tag{2.2}$$

The electromagnetic signal path between the transmit (Tx) antenna and the receive (Rx) antenna is modeled by a linear filter. In general, there is no reason to assume that this filter has constant characteristics, since it depends on the physical structure of the propagation channel (length of the propagation paths, reflections, etc...). However, in practice, it is desirable to model the channel as a linear filter, possibly with time-varying characteristics. Therefore, we denote $h(t, \tau)$ the channel impulse response at time t, where τ is the lag. Hence, $h(t, \cdot)$ is the channel impulse response as experienced by the signal received at time t.

The received signal r results from the superimposition of the transmitted signal convolved with the channel impulse response h, and a noise v. It is subsequently low-pass filtered, in order to reduce the out-of-band noise. Let p_r denote this filter. This yields

$$y \triangleq p_r * r = p_r * h * p_t * x + p_r * v. \tag{2.3}$$

Throughout this thesis, we will assume that the pulse-shaping filter p_t and the corresponding receive filter p_r are chosen such that

- their cascade $p \triangleq p_r * p_t$ does not create inter-symbol interference when $y(t)$ is sampled at rate T_s: $p(nT_s) = 0$, $\forall n \in \mathbb{Z}, n \neq 0$. This property is known as the *Nyquist criterion* [20],
- the spectral properties of the in-band noise are not modified by p_r, *i.e.* $|p_r(f)| = 1, \forall f \in [-\frac{1}{2T_s}, \frac{1}{2T_s}]$.

This can be achieved by using square-root raised-cosine filters for both p_t and p_r.

The continuous-time channel $h(t, \tau)$ is commonly modeled as a tapped delay line, *i.e.* a finite number P of attenuated replicas of the transmitted

signal, resulting from multiple propagation paths between the Tx and Rx antennas:

$$h(t,\tau) \triangleq \sum_{p=1}^{P} a_p(t)\delta_{t_p(t)}(\tau), \tag{2.4}$$

where $a_p(t)$ and $t_p(t)$ are respectively the (time-varying) amplitude and the lag corresponding to each signal path. This model yields

$$y(t) = \sum_{l\in\mathbb{Z}} \sum_{p=1}^{P} x_l \; a_p(t) \; p(t - lT_s - t_p(t)) + (p_r * v)(t). \tag{2.5}$$

2.2.2 Discrete-time model

Sampling the received signal at the symbol rate T_s ($y_k \triangleq y(t_0 + kT_s)$, using the epoch t_0) yields

$$y_k = \sum_{l\in\mathbb{Z}} \sum_{p=1}^{P} x_l \; a_p(t_0 + kT_s) \; p(t_0 + (k-l)T_s - t_p(t_0 + kT_s)) + n_k, \tag{2.6}$$

where $n_k \triangleq (p_r * v)(t_0 + kT_s)$. Since we want to relate the discrete-time sequences y_k and s_l, it is convenient to express the equivalent discrete-time channel impulse response

$$h'_{n,m} \triangleq \sum_{p=1}^{P} a_p(t_0 + nT_s)p(t_0 + mT_s - t_p(t_0 + nT_s)), \tag{2.7}$$

which lets us rewrite y_k in a convolutive way as

$$y_k = \sum_{l\in\mathbb{Z}} x_l h'_{k,k-l} + n_k. \tag{2.8}$$

This channel is sometimes called an Inter-Symbol Interference (ISI) channel, since the signal received at any time contains contributions from several interfering symbols.

The apparent simplicity of eq. (2.8) hides a fairly intricate channel model: the properties of the time-varying channel as defined in (2.7) can not be easily determined. Hence, some simplifying assumptions can be made, in order to make the model easier to deal with.

Finite delay-spread In general, p has an infinite-length impulse response. However, since most of its energy is gathered in a fairly short time span, it is commonly assumed that $p(t)$ has finite support. As a consequence, h' also has finite support in the lag domain. In the sequel, we will assume that the channel impulse response is causal, and that there exists an upper bound L on the channel delay spread, such that

$$\forall n \in \mathbb{Z},\ h'_{n,m} \neq 0 \Rightarrow 0 \leq m \leq L. \tag{2.9}$$

In this case, eq. (2.8) can be rewritten as a finite sum

$$y_k = \sum_{l=0}^{L} h'_{k,l} x_{k-l} + n_k. \tag{2.10}$$

Time-invariant channel If the channel variations are slow, or are not relevant to the considered problem, $h'_{n,m}$ can be considered independent from n. In this case only, the channel can be modeled using the classical linear time-invariant (LTI) model. This yields the classical Additive White Gaussian Noise (AWGN) model.

Independent channel coefficients It seems reasonable to assume that the fading coefficients $a_p(t)$ have independent fading properties. However, as can be seen from eq. (2.7), if the relative tap delays $t_0 - t_p(t_0 + nTs)$ are not integer multiples of T_s, the Nyquist filter p introduces some correlation between the elements of h'. A simplifying assumption, motivated by the fact that most of the energy in the pulse lies in a short time span, is that the support of $p(t)$ is shorter than T_s. In this case, each coefficient a_p affects at most one element of h'.

2.3 MIMO channel model

We consider a point-to-point digital transmission between two stations equipped with multiple Tx and/or Rx antennas, as represented in Fig. 2.1. Let N_t (resp. N_r) denote the number of antennas at the transmitter (resp. the receiver).

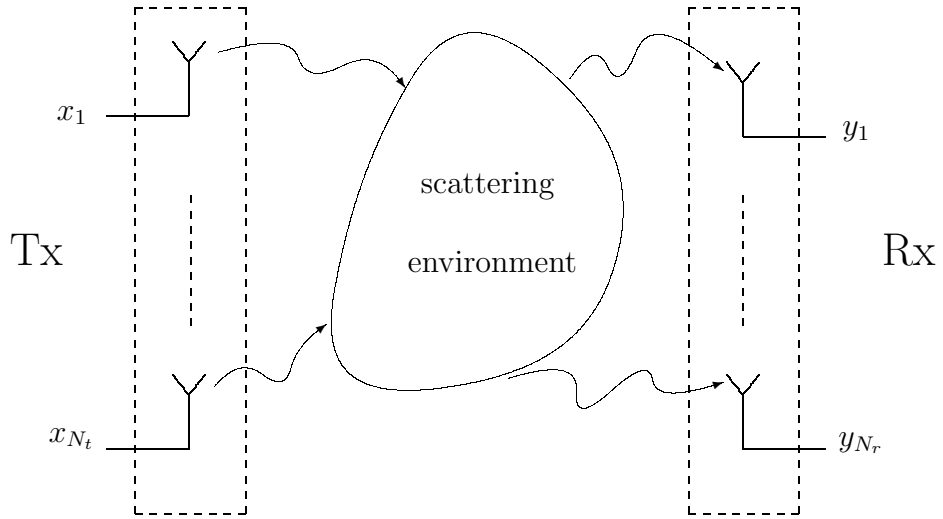


Figure 2.1: MIMO physical channel

The SISO channel model described previously applies to each Tx-Rx antenna pair, hence each Rx antenna senses the superposition of the signals sent from all the Tx antennas, convoluted with the corresponding SISO channels, together with some additive noise. Denoting $x_n(l)$ the baseband symbols transmitted from antenna n ($n = 1 \dots N_t$), and $h'_{m,n}(k,l), l = 0 \dots L$ the discrete-time channel impulse response from Tx antenna n to Rx antenna m at time kT_s, the signal received by the mth Rx antenna is

$$y_m(k) = \sum_{n=1}^{N_t} \sum_{l=0}^{L} h'_{m,n}(k,l) x_n(k-l) + n_m(k), \tag{2.11}$$

where $n_m(k)$ is the noise at the mth Rx antenna. Using the vector representation,

$$\underline{\mathbf{y}}_k \triangleq \begin{bmatrix} y_1(k) \\ \vdots \\ y_{N_r}(k) \end{bmatrix}, \quad \underline{\mathbf{x}}_k \triangleq \begin{bmatrix} x_1(k) \\ \vdots \\ x_{N_t}(k) \end{bmatrix}, \quad \text{and} \quad \underline{\mathbf{n}}_k \triangleq \begin{bmatrix} n_1(k) \\ \vdots \\ n_{N_r}(k) \end{bmatrix}, \tag{2.12}$$

and denoting $[\mathbf{H}(k,l)]_{(m,n)} \triangleq h'_{m,n}(k,l)$, the whole MIMO system is described by

$$\underline{\mathbf{y}}_k = \sum_{l=0}^{L} \mathbf{H}(k,l) \underline{\mathbf{x}}_{k-l} + \underline{\mathbf{n}}_k. \tag{2.13}$$

The case where the delay spread of the channel is negligible w.r.t. the symbol rate ($L = 0$) is called *flat-fading*, as opposed to the *dispersive* (also called *frequency-selective*) case where $L > 0$. The particular situation of a constant flat-fading MIMO channel has been extensively studied. In this case, (2.13) becomes

$$\underline{\mathbf{y}}_k = \mathbf{H}\underline{\mathbf{x}}_k + \underline{\mathbf{n}}_k. \tag{2.14}$$

2.3.1 MIMO capacity and channel knowledge

Telatar [2] has shown that for such a frequency-flat MIMO channel, with $\underline{\mathbf{n}}_k$ i.i.d. with independent components of variance N_0, the mutual information between $\underline{\mathbf{x}}_k$ and $\underline{\mathbf{y}}_k$ is maximized by a Gaussian circularly-symmetric input distribution: $\underline{\mathbf{x}} \sim \mathcal{CN}(0, \mathbf{D})$, and that the expression in this case is

$$\mathcal{I}\left(\underline{\mathbf{x}}; \underline{\mathbf{y}}, \mathbf{H}\right) = \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{D}\mathbf{H}^H\right). \tag{2.15}$$

The Shannon capacity is the supremum of the mutual information over all input distributions:

$$C \triangleq \sup_{\mathbf{D}} \mathcal{I}\left(\underline{\mathbf{x}}; \underline{\mathbf{y}}, \mathbf{H}\right). \tag{2.16}$$

The optimal input covariance can be obtained by the "waterfilling" method [21]. However, this requires that $\mathbf{H}$ (or some information about it) be known by the transmitter. In many cases, this knowledge is not available to the transmitter. In this case, one can still optimize the covariance for a given channel distribution. A classical solution is to transmit uncorrelated signals of equal variance from all antennas ($\mathbf{D} = \mathbf{I}$), since this setting maximizes the expectation of the mutual information in the case of a channel with i.i.d. random coefficients with equal variance. The mutual information achieved by this input distribution is

$$C_{\mathbf{I}} \triangleq \mathcal{I}\left(\underline{\mathbf{x}}; \underline{\mathbf{y}}, \mathbf{H}\right) = \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{H}^H\right) \leq C. \tag{2.17}$$

This highlights an important distinction in the mode of operation of a communications system: the case where the transmitter adapts the input distribution to the instantaneous channel is known as **Channel State Information at the Transmitter** (CSIT). If this information is only available

at the receiver, we will use the term **Channel State Information at the Receiver** (CSIR).

Note that in practice, $\mathbf{H}$ is never perfectly known, neither at the Tx nor at the Rx side, and must be estimated. This is the subject of Chapter 5.

2.4 Channel fading and reciprocity models

So far, we have assumed that the channel can be different for each transmitted symbol, except for the capacity formulas where the channel is considered constant. Practical communication systems operate in between these two extremes: the channel varies, but the symbol period is short enough for the channel not to change substantially over the duration of a symbol. This can be exploited through various methods. For instance, differential modulation [20] is adapted to fast-evolving channels where the coherence time is limited to a few symbol periods. If the channel variation is slower, it can be estimated either through blind or training methods (CSIR situation). The *block-fading* model (also called *quasi-static*) fits this situation. It asserts that the channel remains completely constant for a given time (coherence time), and therefore a block of successive received symbols can be decoded under the assumption that they experienced the same channel conditions. The Rayleigh fading model is an important example of fading model with independent realizations. For a given delay intensity profile $\underline{p}$, the individual coefficients are modeled as independent Gaussian random variables: $\underline{h}'_n \sim \mathcal{CN}\left(0, \mathrm{diag}(\underline{p})\right)$, where $\underline{h}'_n \triangleq [h'_{n,0} \dots h'_{n,L-1}]^T$.

If the symbol rate is very high, or if the physical setting of the channel evolves rapidly, such as the situation of a mobile in a car or a train, a more accurate tracking of the channel evolution might be desirable: exploiting the correlation between successive channel estimates becomes necessary in this case. For instance, the linear phase rotation created by the Doppler effect for common vehicular speeds translates into a fast channel evolution. Fortunately, this evolution is very predictable, and can be easily tracked if the channel model includes temporal evolution. The mutual information (2.15) associated to a fading channel is a random variable, and its Shannon capacity is zero. In this case, we define the outage probability associated to rate R as the probability that the mutual information achieved over the instantaneous

channel realization is lower than R.

The case of CSIT introduces the extra constraint that, in addition to the tracking problem, one has to let the transmitter know about the channel, since the channel estimation is made by the receiver. This can be done either by regularly sending channel estimates from the receiver to the transmitter through a feedback link, or by exploiting the fact that, for duplex systems operating on the same frequency on the uplink and the downlink, the wireless channel behaves similarly in both directions (reciprocity).

A large part of this thesis is dedicated to modeling and exploiting various channel properties in order to better exploit its structure. Chapter 3 deals with the Doppler effect and the corresponding equalization methods in the case of an OFDM transmission. The specific case of CSIT is studied in Chapters 5 and 6: Chapter 5 deals with the impact of imperfect CSIT on the reliability of the transmission, whereas in Chapter 6 we propose a method to model and exploit the reciprocity property of a Time-Division Duplex (TDD) channel. Chapter 7 addresses channel modeling and prediction by exploiting the temporal and spatial correlations in the MIMO case.

Part I

Diversity exploitation in MIMO OFDM

Chapter 3

Equalization of OFDM transmission in high Doppler spread environment

In this chapter, we address the problem of OFDM transmission over a SISO time-varying, frequency-selective channel with high Doppler spread. We show that, when using CP-OFDM, fast (inside one OFDM symbol) channel variations can be decomposed over a base of sinusoid functions sampling the Doppler spectrum at subcarrier frequencies. This leads to a parsimonious parameterization of the time-varying channel impulse response. We show that equalization of the delay spread of a time-varying channel in time domain, and the equalization of the Doppler spread of a frequency selective channel in frequency domain are equivalent. Using this duality, equalization in the frequency domain can benefit from all known time domain equalization methods.

3.1 Introduction

The problem of inter-carrier interference (ICI) in OFDM transmission has been largely ignored in the early days of OFDM. ICI in OFDM appears when the channel impulse response varies significantly over the duration of one OFDM symbol. Recently, the need to consider fast-varying channels for OFDM transmission arose. The influence of uncompensated ICI was investigated by Li in [22], where bounds on the ICI power are derived from the Doppler spread of the channel. Attempts to combat the effects of ICI have been presented, such as Gorokhov and Linnartz's proposal to use the Taylor expansion of the Doppler effect for equalization [23].

Since the channel varies continuously, and since it can not be estimated at every instant in order to limit the amount of training information, the receiver has to make assumptions on the maximum variation speed of the channel coefficients (*i.e.* Doppler spread) and to rely on interpolation. This is used for instance by Stamoulis, Diggavi and Al-Dhahir in [24], where the channel state is linearly interpolated in time between samples over an OFDM symbol. They also proposed an associated time-varying filtering of high complexity to make the channel (almost) time-invariant over an OFDM symbol. Recently, Sayeed *et al.* proposed in [25] to represent channel variations using a set of fixed basis functions, built a CDMA system relying on this "canonical" coordinate system based on Doppler and multipath.

In this chapter, we show that these canonical coordinates for temporal channel variations are particularly suitable for equalization of Cyclic Prefix Orthogonal Frequency Division Multiplexing (CP-OFDM) transmissions, since it leads to a description of Doppler spreading as FIR filtering in frequency domain. The number of taps of this filter is proportional to the Doppler spread.

Results presented in this chapter were published in [5].

3.2 Channel model

Let us consider the CP-OFDM transmission of a complex signal over a fast-varying channel. We consider the transmission of the the discrete-time SISO

channel (2.10). Let $\underline{\mathbf{c}} = [c_0 \cdots c_{N-1}]^T$ represent the complex constellation symbols to be transmitted. The time-domain equivalent is obtained via the N-point inverse Discrete Fourier Transform (DFT)

$$\underline{\mathbf{x}} \triangleq \mathrm{F}^{-1}\underline{\mathbf{c}} \tag{3.1}$$

In order to avoid interference between consecutive OFDM symbols, we assume that a cyclic prefix [26] of length P ($P \geq L-1$) is used, *i.e.* the last P samples of $\underline{\mathbf{s}}$ are prepended in the time-domain to the OFDM symbol itself before transmission. The cyclic prefix insertion is represented by

$$\mathrm{C} \triangleq \begin{bmatrix} 0_{P\times N-P} & \mathrm{I}_P \\ & \mathrm{I}_N \end{bmatrix} \tag{3.2}$$

It yields the transmitted signal

$$\underline{\mathbf{s}}' \triangleq \mathrm{C}\,\underline{\mathbf{s}}. \tag{3.3}$$

For the sake of simplicity, let us assume that the first value in $\underline{\mathbf{s}}$ is transmitted at time zero. Therefore, the first value of the cyclic prefix is transmitted at time $-P$. Let us represent the convolutive ISI channel experienced by $\underline{\mathbf{s}}'$ as

$$\mathrm{H}' \triangleq \begin{bmatrix} h_{-P,0} & 0 & \dots & 0 \\ \vdots & h_{1-P,0} & \ddots & \vdots \\ h_{L-1-P,L-1} & \vdots & \ddots & 0 \\ 0 & h_{L-P,L-1} & & \\ & & \ddots & h_{N-1,0} \\ \vdots & \ddots & & \vdots \\ 0 & \dots & 0 & h_{N+L-2,L-1} \end{bmatrix}. \tag{3.4}$$

This yields the received signal

$$\underline{\mathbf{r}}' \triangleq \mathrm{H}'\underline{\mathbf{s}}' + \underline{\mathbf{v}}', \tag{3.5}$$

where $\underline{\mathbf{v}}'$ is some additive noise. Then, the cyclic prefix removal operation consists in discarding the first P values of the received signal, which in general contain interference from the previous OFDM symbol. The last $L-1$ samples in $\underline{\mathbf{r}}'$ are themselves interfering with the following OFDM symbol, and should be ignored as well. Therefore, in the sequel we will only consider

$$\underline{\mathbf{r}} = \mathrm{D}\underline{\mathbf{r}}' \tag{3.6}$$

where

$$\mathrm{D} \triangleq \left[\begin{array}{ccc} 0_{N\times P} & \mathrm{I}_N & 0_{N\times L-1} \end{array} \right]. \tag{3.7}$$

Finally, the frequency-domain equivalent of $\underline{\mathbf{r}}$ is obtained by DFT

$$\underline{\mathbf{u}} \triangleq \mathrm{F}\underline{\mathbf{r}} = \mathrm{FDH'CF}^{-1}\underline{\mathbf{c}} + \underline{\mathbf{w}}, \tag{3.8}$$

$\underline{\mathbf{w}} = \mathrm{FD}\underline{\mathbf{v}}'$. Note that the cyclic prefix addition and removal does not modify the noise properties, since $\mathrm{FD}(\mathrm{FD})^H = \mathbf{I}_N$.

3.3 Canonical channel representation

Let us focus on the properties of the equivalent channel between $\underline{\mathbf{c}}$ and $\underline{\mathbf{u}}$. In particular, the effect of the pre- and post-multiplication by D and C respectively, is to create an equivalent $N \times N$ channel matrix $\mathrm{H} = \mathrm{DH'C}$ with the following structure:

$$\mathrm{H} = \left[\begin{array}{ccccc} h_{0,0} & & h_{0,L-1} & \cdots & h_{0,1} \\ \vdots & h_{1,0} & & \ddots & \vdots \\ \vdots & & & & h_{L-2,L-1} \\ h_{L-1,L-1} & \vdots & & & \\ & h_{L,L-1} & \ddots & & \\ & & & & \\ & & \ddots & & \\ & & & & h_{N-1,0} \end{array} \right]. \tag{3.9}$$

Let us consider the variation speed of the channel. This can be related to the spectrum of $h(t,\tau)$ with respect to the time variable t (*i.e.* the Doppler spectrum). Let us assume that $h(f,\tau)$ has a finite spectrum support $[-B_D,\ B_D]$ for every τ.

In a quasi-static environment ($NB_DT_s \ll 1$), H is a circulant matrix, i.e. $\forall(p,l), h_{p,l} = h_{0,l}$. It is well known that the consequence of this is that the frequency-domain expression of the channel is a diagonal matrix: FHF^{-1} has non-zero values on its diagonal only, and their amplitude represents the channel gain on each frequency subband. In this case, equalization becomes a trivial operation, and this characteristic was one of the keys of the success of OFDM.

Contrarily, we focus on the case where the channel coefficients noticeably vary in time during one OFDM symbol. In this case, H is not circulant anymore. Let us decompose the variations of each $h_{p,l}$ over one OFDM symbol over a base of F_s/N-spaced sinusoid functions. (This corresponds to the frequency spacing of the OFDM subcarriers). Since we need to represent the varying channel coefficients over a time period of NT_s, a sampling frequency in frequency domain of F_s/N allows to satisfy Nyquist. In order to cover all the Doppler spectrum, we define $M = \lceil NB_D\ T_s \rceil$ and decompose

$$h_{p,l} = \sum_{m=-M}^{M} a_{l,m} \mathrm{e}^{j\frac{2\pi mp}{N}}, \qquad \begin{array}{l} p = 0 \dots N-1 \\ l = 0 \dots L-1 \end{array} \tag{3.10}$$

This base can be represented in the matrix space of H by $\{\mathrm{Q}^m,\ m = -M \dots M\}$, where

$$\mathrm{Q} \triangleq \mathrm{diag}\left(1,\ \mathrm{e}^{j\frac{2\pi}{N}},\ \dots,\ \mathrm{e}^{j\frac{2\pi(N-1)}{N}}\right). \tag{3.11}$$

Using the $N \times N$ circulant shift matrix

$$\mathrm{J} \triangleq \begin{pmatrix} 0 & & & 1 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}, \tag{3.12}$$

we can rewrite the H as

$$\mathrm{H} = \sum_{l=0}^{L-1} \sum_{m=-M}^{M} a_{l,m}\ \mathrm{J}^l\ \mathrm{Q}^m\ . \tag{3.13}$$

3.3.1 Time-Frequency Duality, applications to equalization

Let us point out that the DFT has interesting properties on Q and J:

$$\mathrm{FQF}^{-1} = \mathrm{J}, \quad \text{and} \quad \mathrm{FJF}^{-1} = \mathrm{Q}^* \tag{3.14}$$

(complex conjugate). Which yields

$$\mathrm{A} \triangleq \mathrm{FHF}^{-1} = \sum_{l=0}^{L-1} \sum_{m=-M}^{M} a_{l,m}\ \mathrm{Q}^{l*}\ \mathrm{J}^m \tag{3.15}$$

Comparing equations (3.13) and (3.15), we see a duality between the time-domain and the frequency-domain representations of the channel. Both A and H are band matrices, with respectively L and $2M+1$ non-zero diagonals. This duality points to a possible optimization of the equalization process: among the two varieties of interference (ISI and ICI), either one can be suppressed through the use of the Fourier transform and a scalar multiplication, whereas the other one requires the use of a conventional equalizer. In eqs. (3.13) and (3.15), the exponent of J represents the actual equalization effort. The exponent of Q has no influence on the equalization complexity, since it represents N independent, scalar weighting operations. Therefore, the decision between these two alternatives depend on the spread of the interference:

- if $(2M+1) < L$, the channel is preferably considered as a $(2M+1)$-tap frequency-varying cyclic FIR filter in the frequency domain
- if $(2M+1) > L$, it is preferable to equalize the L-tap time-varying FIR filter in the time-domain.

Any conventional (time-domain) equalization algorithm can be used with this method. Furthermore, the correlation between successive OFDM symbols can be exploited by considering the set of parameters describing the evolution inside one OFDM symbol as a random variable:

$$\mathrm{A} \triangleq \begin{bmatrix} a_{0,-M} & \dots & a_{0,M} \\ \vdots & \ddots & \vdots \\ a_{L-1,-M} & \dots & a_{L-1,M} \end{bmatrix}. \tag{3.16}$$

The correlation of the successive values of A ($\mathrm{A}_n, n \in \mathbb{Z}$) can be exploited (*e.g.* through Wiener filtering), in order to improve their estimation .

3.4 Conclusion

We proposed a new modeling scheme for time-varying, frequency-selective SISO channels with high Doppler spread, that proves particularly convenient when used in conjunction with CP-OFDM modulation. We showed that this model enables the receiver to exploit the duality between Doppler spread in OFDM systems and delay spread in classical systems, by trading ISI for ICI and *vice versa*, and that virtually any equalization algorithm existing in the time domain can be transposed into the frequency domain.

Chapter 4

Exploiting Diversity Through Space-Time-Frequency Spreading

In this chapter, we introduce a coding scheme for MIMO-OFDM that exploits space, frequency and time diversity through the concatenation of a binary convolutive code and a full-rate linear precoder. We optimize the design of the linear precoder w.r.t. various properties, including diversity, coding gain, and a mutual information criterion. We evaluate the performance of an iterative PIC decoder adapted to the structure of our code through simulation, and show how the linear precoder improves the performance of the overall system.

4.1 Introduction

The use of wideband OFDM modulations for MIMO systems, in addition to the obvious increase of throughput that it enables, also promises some diversity advantage over narrowband flat-fading channels, through the use of multiple frequency subbands. The initial trend in digital communications engineering was to divide the channel into multiple disjoints subchannels, and to multiplex the data over those subchannels. The OFDM modulation technique is a natural way of turning a dispersive channel into multiple parallel flat-fading channels by transmitting different symbols on different frequency bands. This was a sensible choice as long as the processing power available for equalization remained expensive: without the necessary processing power, diversity (other than receive diversity) is merely a source of self-interference. However, technological advances in equalization, combined with a renewed interest for the sources and mechanisms of diversity stemming from Foschini's BLAST proposal [3], prompted the research community to reconsider this choice. Diversity at the transmission side is nowadays considered positively, especially for wireless communications, where its "channel hardening" effect [27] is desirable.

In the light of this analysis, the use of multiple antennas for Wireless Local Area Networks (WLANs) is an obvious choice to adapt current WLAN standards to the ever-growing bandwidth demand. This chapter presents a transmission scheme that was developed in the framework of a project by the French National Telecommunications Research Forum (Réseau National de Recherche en Télécommunications – RNRT) set to extend the HiperLAN/2 standard [28] (the European counterpart of IEEE 802.11a/g) to using multiple antennas. The proposed coding scheme relies on Linear Constellation Precoding (LCP) on top of an OFDM modulation to exploit the channel diversity, including frequency diversity (the signal bandwidth used in HiperLAN/2 is 20MHz), space diversity (through the use of multiple Tx and Rx antennas), and possibly time diversity (if some delay is acceptable). Therefore, constraints such as tractable decoding complexity were taken into account. For instance, the rate-diversity trade-off (presented in Section 4.2) is clearly biased in favour of diversity: the proposed scheme achieves a throughput that grows linearly with the number of Tx antennas N_t (this property is closely tied to the notion of full-rate precoding, described in Section 4.3). The remaining degrees of freedom of the channel are dedicated to improving

diversity, thereby reducing the energy required to transmit a given amount of data.

This chapter is organized as follows: the notions of diversity and rate are introduced in Section 4.2. The proposed coding scheme is exposed in Section 4.3, and the precoder is shown to be mutual information-lossless, to exploit all the available diversity, and to maximize the coding gain. Section 4.4 analyzes the proposed iterative interference cancellation (IC) scheme. Simulation results are presented in Section 4.5.

Results presented in this chapter were published in [6, 7, 8].

4.2 Coding for MIMO OFDM: diversity and rate

The concept of diversity originates in SIMO systems: for a non-dispersive channel, the received signal is written

$$\underline{\mathbf{y}} = \underline{\mathbf{h}}x + \underline{\mathbf{n}} \tag{4.1}$$

where x is the transmitted scalar symbol, and the channel is represented by

$$\underline{\mathbf{h}} \triangleq \begin{bmatrix} h_1 \\ \vdots \\ h_{N_r} \end{bmatrix}. \tag{4.2}$$

Under the assumption that $\underline{\mathbf{h}}$ is perfectly known to the receiver, maximum-ratio combining (MRC) of the N_r received signals yields

$$z = \frac{1}{|\underline{\mathbf{h}}^T\underline{\mathbf{h}}|}\underline{\mathbf{h}}^T\underline{\mathbf{y}} = x + \frac{1}{N_t}\mathbb{1}_{N_r}^T\underline{\mathbf{n}} \tag{4.3}$$

Let us consider the AWGN case, where $\underline{\mathbf{h}} = \frac{1}{\sqrt{N_r}}\mathbf{1}_{N_r}$, and the noise on all channels is Gaussian i.i.d., *i.e.* $\underline{\mathbf{n}} \sim \mathcal{CN}(0, \sigma_n^2\mathbf{I})$. Assuming $|x| = 1$, the Signal-to-Noise Ratio (SNR) of a SISO system operating in the same conditions is $\rho \triangleq \frac{1}{\sigma_n^2}$. The SNR at the output of the MRC is

$$\rho' \triangleq \frac{\left|\underline{\mathbf{h}}^T\underline{\mathbf{h}}\right|^2}{\mathrm{E}_{\underline{\mathbf{n}}}\left[\left|\underline{\mathbf{h}}^T\underline{\mathbf{n}}\right|^2\right]} = \frac{N_r}{\sigma_n^2}. \tag{4.4}$$

The effective SNR gain in the case of receive diversity is therefore $\rho'/\rho = N_r$. In the more involved case of coded transmission over a fading channel, the channel is a random variable $\{\mathbf{H}\}$, and therefore so is the instantaneous capacity $C(\mathbf{H})$. It can be shown [29, 30, 31] that for a given code of rate R, and for a given SNR, diversity decreases the channel outage probability $P\left(C(\mathbf{H}) < R\right)$. This is the "channel hardening" effect described by Hochwald *et al.* in [27]. From a slightly different point of view, in the limit of infinite diversity, any fading channel can be transformed into an AWGN channel [32].

Among the different forms of diversity, let us differentiate space, time, and delay diversity.

Space diversity is obtained by utilizing several antennas. A rough assumption is that if the multipath environment is sufficiently rich, an antenna separation of the order of the wavelength is sufficient to create wireless channels with statistically uncorrelated impulse response realizations.

Time diversity is based on the observation that, for moving settings, the channel changes over time. However, exploiting time diversity can require very long code lengths.

Delay diversity consists in exploiting the taps in the channel impulse response, since their gains, determined by different signal paths over the air, fade independently. It has been shown in [33] that frequency diversity in OFDM-coded systems is the counterpart of multipath diversity for wideband systems. Time-diversity corresponds to the variable t in eq. (2.4), whereas multipath (or frequency) diversity corresponds to τ.

Channel codes that exploit the space and time diversity of the MIMO wireless channel are known as *Space-Time Codes* (STCs). A large number of early works on MIMO communications have focused on the Space-Time Coding problem, and many extensions to the OFDM case were proposed later (they are known as Space-Time-Frequency Codes, STFCs). The V-BLAST [3] STC, which consists in the transmission of independent symbols on all antennas in parallel, was instrumental in the realization that the degrees of liberty in the channel could be used not only to decrease the probability of error in the transmitted symbols, but also to transmit more data in parallel.

These two notions are respectively formalized as *diversity gain* and *multiplexing gain*. The various forms (space, time, multipath) of diversity are essentially different sources of a common resource: if one considers the channel representation as a random variable, the amount of independent sources of randomness involved in the description of the channel equals the dimensionality of the vector space spanned by the channel realizations. Zheng and Tse later showed [34] that any modulation scheme essentially taps this "pool" of channel diversity and use it either towards channel hardening, or towards multiplexing, the trade-off between these two uses being precisely quantified. They define the multiplexing gain r as

$$r \triangleq \lim_{\mathsf{SNR}\to+\infty} \frac{R}{\log \mathsf{SNR}}, \tag{4.5}$$

and the diversity gain is the asymptotic slope of the (log-log) error probability curve, formally

$$d \triangleq -\lim_{\mathsf{SNR}\to+\infty} \frac{\log P_e(\mathsf{SNR})}{\log \mathsf{SNR}}. \tag{4.6}$$

The optimal trade-off curve for a flat-fading channel is given by

$$d^*(k) = (N_t - k)(N_r - k), \quad k = 0 \dots \min\{N_t, N_r\}. \tag{4.7}$$

Medles [35, 36] has shown that a similar expression, taking into account the delay diversity, exists for frequency-selective channels. In this case, assuming that all L taps in the impulse response have non-zero variance, the optimal trade-off becomes

$$d^*(k) = (Lq - k)(p - k), \quad k = 0 \dots p. \tag{4.8}$$

where $p \triangleq \min\{N_t, N_r\}$ and $q \triangleq \max\{N_t, N_r\}$.

In this chapter, we introduce a family of codes designed for MIMO OFDM channels, based on the concatenation of a binary channel code and LCP, that we denote by Space-Time-Frequency Spreading (STFS). STFS was developed to provide the maximum diversity available from the channel, while remaining within reasonable complexity bounds. For this reason, we only explore maximum-diversity solutions with non SNR-adaptive codes. (they have a constant rate R, therefore the multiplexing gain is $r = 0$, and we can seek to exploit the maximum diversity gain $d^*(0) = N_t N_r L$).

4.2.1 Space-Time-Frequency Coding

Linear Space-Time Block Codes (LSTBCs) designates a class of channel codes where the consecutive signals transmitted by the antennas are defined block-by-block by a linear transformation of a set of constellation symbols. Among them, orthogonal codes are remarkable in that they only require a simple Maximum Ratio Combining (MRC) operation at the receiver to achieve ML decoding. Alamouti's elegant solution [37] to the 2 Tx antennas case is an example of orthogonal code. Tarokh *et al.* have shown in [38] that single-rate (1 constellation symbol per channel access) orthogonal codes only exist for $N_t = 2$. For a greater number of transmit antennas, simpler coding-decoding will always be done at the expense of the code rate.

In the case of OFDM modulations, the channel is split into multiple flat-fading subbands. Trellis codes adapted to this case were proposed [39], as well as multicarrier extensions of the Alamouti technique [40]. However, simple transposition of the Alamouti scheme into the frequency space fails to exploit the full diversity [41]. In order to leverage the frequency diversity of the channel, it is necessary to code jointly over the frequency domain. However, the number N of subbands in OFDM systems is typically higher than the actual channel delay diversity [42], which means that the number of coefficients needed to represent the OFDM channel state ($N_t N_r N$) is higher than the actual diversity available ($N_t N_r L$). This is because the channel coefficients in the OFDM representation are not independent.

Our method uses groups of frequencies that form a partition of the spectrum. Inside each group, the diversity is exploited by linear precoding. We present a stream-based coding scheme, along the lines of the threaded scheme proposed by El Gamal in [43]. In this approach, each stream is separately encoded by a binary code, mapped onto a complex constellation, and then spread over the OFDM tones so that it is not self-interfering. Iterative decoding with streams-based interference cancellation is proposed. The subcarrier allocation that we propose is similar to the one in [44], although our method goes further into exploiting the space diversity using LCP.

4.3 Encoding Scheme Description

Let us now describe the proposed coding scheme, based on the concatenation of a linear precoder and a binary code, that exploits all the available spatial and frequency diversity with reasonable decoding complexity constraints.

We designed the encoding scheme using the following assumptions:

- Space and frequency diversities are exploited through LCP: constellation symbols are grouped into vectors that are linearly coded with a square matrix Q (more details on the definition of Q will follow), and interspersed over the available antennas and frequencies.
- An outer binary code (such as a convolutive code) provides for extra coding gain
- The potentially huge decoding complexity of the linear code is alleviated by splitting the data into several groups, or *streams*. This eases the requirements on the decoder, since this independence enables iterative streams-based interference cancellation.
- In order to avoid error propagation in the iterative interference cancellation, streams shall not be self-interfering.
- Each stream must take advantage of all the available diversity, *i.e.* the size of the precoder Q is at least the diversity characteristic of the channel.

This is achieved by a frequency (or tone) allocation scheme that maps different streams to all Tx antennas over each frequency subband, thus ensuring that the transmitted signal is spatially white. For this reason, the number of streams must be greater or equal to N_t. Since using more streams increases the complexity without added benefit, we will work under the assumption that there are N_t streams.

Channel model In this chapter, we assume a block-fading channel, one block being equivalent to an OFDM symbol. Therefore, the ICI described in Chapter 3 is not considered here, and the OFDM modulation turns the channel into a set of parallel MIMO AWGN flat-fading channels. Let us

denote the channel experienced by OFDM symbol j by a $N_r \times N_t$ matrix of complex coefficients for each frequency subband $p \in [\![1, P]\!]$,

$$\mathrm{H_{j,p}} = \begin{bmatrix} h_{j,1,1,p} & \dots & h_{j,1,N_t,p} \\ \vdots & \ddots & \vdots \\ h_{j,N_r,1,p} & \dots & h_{j,N_r,N_t,p} \end{bmatrix}. \tag{4.9}$$

Let $\mathrm{S}_{j,m} = [s_{j,1,m}, \dots, s_{j,P,m}]^T$ denote the the frequency-domain representation of the OFDM symbol transmitted over antenna m during OFDM symbol j. The input/output relationship of the channel in each subband $p \in [\![1, N]\!]$ of OFDM symbol j is

$$\underline{\mathbf{y}}_{j,p} = \mathrm{H_{j,p}} \begin{bmatrix} s_{j,p,1} \\ \vdots \\ s_{j,p,N_t} \end{bmatrix} + \underline{\mathbf{k}}_{j,p} \tag{4.10}$$

where $\underline{\mathbf{k}}_{j,p} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$. Furthermore, due to the orthonormality of the DFT, the noise can be assumed independent among subbands.

Let us consider J channel realizations, which we will assume independent, and define the block-diagonal channel matrices $\mathbf{H_j}$ and $\mathbf{H}$, corresponding respectively to the jth OFDM symbol, and to all J considered OFDM symbols:

$$\mathbf{H_j} \triangleq \begin{bmatrix} \mathrm{H_{j,0}} & & 0 \\ & \ddots & \\ 0 & & \mathrm{H_{j,p}} \end{bmatrix}, \quad \mathbf{H} \triangleq \begin{bmatrix} \mathrm{H_0} & & 0 \\ & \ddots & \\ 0 & & \mathrm{H_{J-1}} \end{bmatrix}. \tag{4.11}$$

We assume that the channel coefficients are spatially and temporally uncorrelated (temporal decorrelation can always be achieved by interleaving with a sufficient delay, whereas spatial decorrelation is a function of antenna spacing). Channel coefficients can be correlated between frequencies for a single Tx-Rx antenna pair. Therefore, all non-zero coefficients in the block-diagonal matrix $\mathbf{H}$ are independent. These assumptions can be summarized as

$$\forall (p_1, p_2), \quad (t_1, m_1, n_1) \neq (t_2, m_2, n_2) \Rightarrow \mathrm{E}\left[h_{t_1,n_1,m_1,p_1} {h_{t_2,n_2,m_2,p_2}}^*\right] = 0 \tag{4.12}$$

We also assume that the channel frequency diversity is at least L. In other words, frequencies taken N/L tones apart or more are assumed totally uncorrelated. (This measure is the frequency-domain equivalent of the channel

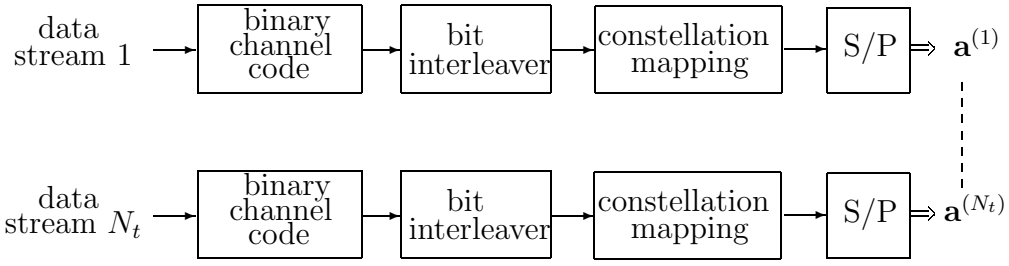


Figure 4.1: Initial stream separation

delay spread - see [33]).

$$\forall(t,m,n), \quad |p_1 - p_2| \geq L \Rightarrow \mathrm{E}\left[h_{t,n,m,p_1} {h_{t,n,m,p_2}}^*\right] = 0. \tag{4.13}$$

The noise samples are assumed to be white complex Gaussian, independent and identically distributed over all frequencies.

4.3.1 Streams principle and tone assignment

The incoming data is split into N_t streams that are treated separately, as represented in Fig. 4.1. Each stream is encoded using a binary (convolutional or block) channel code, bit interleaved, mapped onto complex symbols, and serial-to-parallel (S/P) converted to form $N_t L \times 1$ vectors $\underline{\mathbf{a}}_t^{(k)}, \quad k \in [\![1, N_t]\!]$. Inside each stream k, those are linearly precoded through multiplication by a square matrix Q:

$$\underline{\mathbf{x}}_t^{(k)} = \mathrm{Q}\underline{\mathbf{a}}_t^{(k)} \tag{4.14}$$

Since Q is square, no redundancy is added at the linear precoding stage: we denote this a *full-rate* precoder.

Every constellation symbol is spread by matrix Q over a set of $N_t L$ values, that must be mapped onto particular tones and Tx antennas to form OFDM symbols. They are interspersed in time, space and frequency to ensure that the fading coefficients are as little correlated as possible. To this end, we used the following criterions:

- all antennas must be used evenly by each symbol
- all frequencies must be used evenly by each symbol

Due to the correlation between adjacent frequencies over the same antenna pair, the spectrum need only be sampled L times on each Tx antenna for every symbol. Conversely, since there is no spatial correlation, it makes sense for every symbol to use all N_t Tx antennas.
Furthermore, the absence of self-interference constraint dictates that symbols from the same stream can not be transmitted over the same frequency from different Tx antennas.

This leads us to the design of a $JN_tN \times JN_tN$ permutation of all the constellation symbols transmitted over J OFDM symbols: We assume that N is an integer multiple of N_tL: $N = GN_tL$. Over this period of time, the output of the linear precoders defined in (4.14) for all N_t streams is

$$\underline{\mathbf{x}}_{jG+g}^{(k)}, \quad \begin{array}{l} k \in [\![1, \ldots, N_t]\!], \\ g \in [\![0, \ldots, G-1]\!], \\ j \in [\![0, \ldots, J-1]\!]. \end{array} \tag{4.15}$$

Since each of those contains N_tL values, a total of ${N_t}^2JGL = JN_tN$ complex values are output. In our simulations, we used the permutation defined as follows:

$$\begin{gathered} \forall (l, i, f) \in [\![0, L-1]\!] \times [\![1, N_t]\!] \times [\![1, N]\!], \\ i + k - 1 + lGN_t + gN_t = eN + f \Rightarrow s_{j+e,f,i} = x_{jG+g,lN_t+i}^{(k)}, \end{gathered} \tag{4.16}$$

with $e = 1$ iff $i+k-1+lGN_t+gN_t \geq N$ and $j \neq J-1$ (when $\underline{\mathbf{x}}_{jG+g}^{(k)}$ is split over two consecutive OFDM symbols), and $e = -(J-1)$ iff $i+k-1+lGN_t+gN_t \geq N$ and $j = J-1$ (wrapping). An example of this tone allocation (not showing the border effects) is pictured on Figure 4.2, for the case $N_t = 4$, $L = 3$, $J = 2$.

The overall operation is a $JN_tN \times JN_tN$ permutation, and can be written as a matrix $\mathbf{W}$ (with $\mathbf{W}^H\mathbf{W} = I_{JN_tN}$). Let $\mathbf{W}_k$ denote the matrix representing the spreading of the values from stream k, *i.e.* columns $(k-1)JN+1 \ldots kJN$ of $\mathbf{W}$. Let us gather all the $\underline{\mathbf{a}}_t^{(k)}$ transmitted over J OFDM symbols, and the corresponding received signals, in

$$\mathbf{A^{(k)}} \triangleq \begin{bmatrix} \underline{\mathbf{a}}_1^{(k)} \\ \vdots \\ \underline{\mathbf{a}}_{JG}^{(k)} \end{bmatrix}, \quad \underline{\mathbf{a}} \triangleq \begin{bmatrix} \mathbf{A^{(1)}} \\ \vdots \\ \mathbf{A^{(N_t)}} \end{bmatrix} \tag{4.17}$$

and

$$\mathbf{Y_j} \triangleq \begin{bmatrix} \underline{\mathbf{y}}_{j,1} \\ \vdots \\ \underline{\mathbf{y}}_{j,N} \end{bmatrix}, \quad \mathbf{Y} \triangleq \begin{bmatrix} \mathbf{Y_0} \\ \vdots \\ \mathbf{Y_{J-1}} \end{bmatrix}. \tag{4.18}$$

Using these notations, and with

$$\mathbf{Q} = \mathbf{I}_{N_t GJ} \otimes \mathrm{Q} \tag{4.19}$$

and the noise $\mathbf{K} \triangleq [\underline{\mathbf{k}}_{1,1} \dots \underline{\mathbf{k}}_{J,N}]^T$, the signal received over J consecutive OFDM symbols is

$$\mathbf{Y} = \mathbf{HWQA} + \mathbf{K}. \tag{4.20}$$

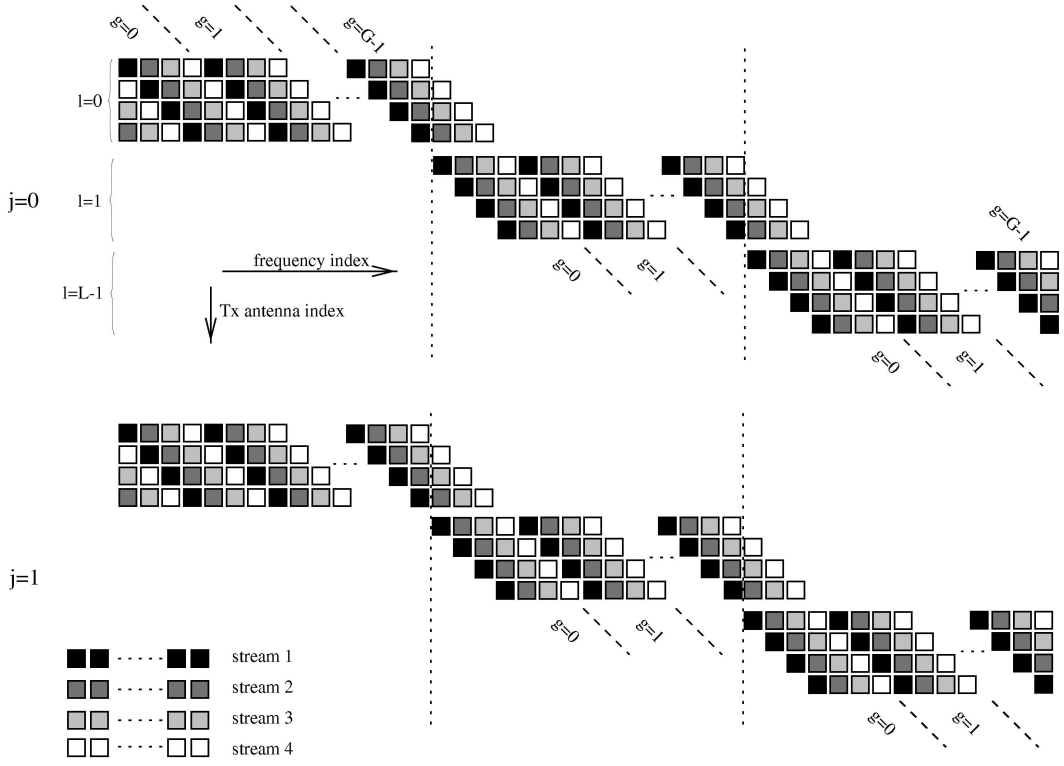


Figure 4.2: Streams-based tone assignment

4.3.2 Choice of the linear precoding matrix

In the coding scheme exposed so far, the precoding matrix Q has still not been precisely defined. We now derive some necessary conditions on Q, and try to optimize it w.r.t. the noise and channel fading statistics, using consecutively an information-theoretic criterion, and a Pairwise Error Probability (PEP) criterion that maximizes the diversity and the coding gain.

Mutual information criterion

We use an information-theoretic criterion to restrict our search to information-lossless precoding matrices. For this analysis, let us assume that the transmitted symbols $\underline{\mathbf{a}}_t^{(k)}$ are circularly symmetric, Gaussian random variables, of variance σ_a^2. Although this is only an approximation of the (usually discrete) constellations used in digital communications, outage capacity appears to give a reasonable estimate of the achievable Packet Error Rate (PER) of a Bit-Interleaved Coded Modulation (BICM) code [12, 13].

Assuming perfect channel knowledge, the mutual information between $\mathbf{A}$ and $\mathbf{Y}$ is

$$\mathcal{I}(\mathbf{A};\mathbf{Y}) = \mathcal{H}(\mathbf{Y}) - \mathcal{H}(\mathbf{Y}|\mathbf{A}) = \mathcal{H}(\mathbf{Y}) - \mathcal{H}(\mathbf{K}) \tag{4.21}$$

Using equation (4.20), and assuming that the noise samples in $\mathbf{K}$ and the symbols in $\mathbf{A}$ are independent circularly symmetric, Gaussian random variables, the mutual information becomes [2]:

$$\begin{aligned}
\mathcal{I}(\mathbf{A};\mathbf{Y}) =& \log\det\left(\pi e\,\mathrm{E}_{\mathbf{Y}}\left[\mathbf{Y}\mathbf{Y}^{\mathrm{H}}\right]\right) - \log\det\left(\pi e\,\mathrm{E}_{\mathbf{K}}\left[\mathbf{K}\mathbf{K}^{\mathrm{H}}\right]\right) && (4.22)\\
=& -\log\det\left(\pi e\,\mathbf{I}_{JN_rN}\right) + \log\det\left(\pi e\left(\mathbf{I}_{JN_rN}\right.\right. && (4.23)\\
&\left.\left. +\rho\mathbf{H}\mathbf{W}\left(\mathbf{I}_{N_tGJ}\otimes \mathrm{Q}\right)\left(\mathbf{I}_{N_tGJ}\otimes \mathrm{Q}\right)^H\mathbf{W}^H\mathbf{H}^H\right)\right)\\
=& \log\det\left(\left(\,\mathbf{I}_{JN_rN} + \rho\mathbf{H}\mathbf{W}\left(\mathbf{I}_{N_tGJ}\otimes \mathrm{Q}\mathrm{Q}^H\right)\mathbf{W}^H\mathbf{H}^H\right)\right) && (4.24)
\end{aligned}$$

where $\rho \triangleq \frac{\sigma_a^2}{\sigma_n^2}$ is the receive SNR. Since the channel capacity is known to be

$$\mathrm{C} = \log\det\left(\mathbf{I}_{JN_rP} + \rho\mathbf{H}\mathbf{H}^H\right), \tag{4.25}$$

and since $\mathbf{W}\mathbf{W}^H = \mathbf{I}_{\mathbf{JN_tP}}$ by definition of a permutation, comparing (4.24) and (4.25) shows that choosing Q among unitary matrices (such that $\mathrm{Q}\mathrm{Q}^H = \mathbf{I}_{\mathbf{LN_t}}$) ensures that there is no capacity loss at the linear precoding stage ($\mathcal{I}(\mathbf{A};\mathbf{Y}) = \mathrm{C}$).

Diversity criterion

Let us consider a single error event at the linear processing stage, focusing on one particular symbol ($\underline{\mathbf{a}}_{jG+g}^{(k)}$, with j, k and g fixed): the decoded vector $\mathbf{a}'^{(k)}_{jG+g}$ differs from the transmitted one $\underline{\mathbf{a}}_{jG+g}^{(k)}$ only in the u-th coefficient, $u \in [\![1, N_t L]\!]$

$$\underline{\mathbf{e}} \triangleq \underline{\mathbf{a}}_{jG+g}^{(k)} - \underline{\mathbf{a}}'^{(k)}_{jG+g} = (0, \dots 0, e_u, 0, \dots 0)^T \tag{4.26}$$

where $e_u \neq 0$ is any difference between two complex symbols of the constellation.

With j, k and g fixed, equation (4.16) shows that the values in $\underline{\mathbf{x}}_{jG+g}^{(k)}$ are affected to $N_t L$ different tones, on one and only one Tx antenna for every tone involved. Let $\{p_1 \dots p_{N_t L}\}$, $\{i_1 \dots i_{N_t L}\}$ and $\{j_1 \dots j_{N_t L}\}$ denote respectively the tone number, antenna number and OFDM symbol number corresponding to the $N_t L$ elements in $\underline{\mathbf{x}}_{jG+g}^{(k)}$, and let us form a vector that gathers the LN_tN_r channel coefficients actually involved in its transmission:

$$\underline{\mathbf{c}}_{j,k,g} \triangleq \left(h_{j_1,1,i_1,p_1} \dots h_{j_1,N_r,i_1,p_1}, \dots, h_{j_{LN_t},1,i_{LN_t},p_{LN_t}} \dots h_{j_{LN_t},N_r,i_{LN_t},p_{LN_t}}\right)^T. \tag{4.27}$$

Similarly, at the receiver,

$$\underline{\mathbf{u}}_{jG+g}^{(k)} \triangleq \left(y_{j,p1,1} \dots y_{j,p1,N_r}, \dots, y_{j,p_{LN_t},1} \dots y_{j,p_{LN_t},N_r}\right)^T. \tag{4.28}$$

The received signals corresponding to the transmission of $\underline{\mathbf{a}}_{jG+g}^{(k)}$ can be written as

$$\underline{\mathbf{u}}_{jG+g}^{(k)} = \left(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{a}}_{jG+g}^{(k)}) \otimes \mathbf{I}_{N_r}\right) \underline{\mathbf{c}}_{j,k,g}, \tag{4.29}$$

and therefore

$$\underline{\mathbf{u}}_{jG+g}^{(k)} - \underline{\mathbf{u}}'^{(k)}_{jG+g} = \left(\mathrm{diag}\left(\mathrm{Q}(\underline{\mathbf{a}}_{jG+g}^{(k)} - \underline{\mathbf{a}}'^{(k)}_{jG+g})\right) \otimes \mathbf{I}_{N_r}\right) \underline{\mathbf{c}}_{j,k,g} \tag{4.30}$$

which lets us write the the squared Euclidean distance between the received samples as

$$\mathrm{d}^2\left(\underline{\mathbf{u}}_{jG+g}^{(k)}, \underline{\mathbf{u}}'^{(k)}_{jG+g}\right) \triangleq \left(\underline{\mathbf{u}}_{jG+g}^{(k)} - \underline{\mathbf{u}}'^{(k)}_{jG+g}\right)^H \left(\underline{\mathbf{u}}_{jG+g}^{(k)} - \underline{\mathbf{u}}'^{(k)}_{jG+g}\right) \tag{4.31}$$

$$= \underline{\mathbf{c}}_{j,k,g}^H \left(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})^H \otimes \mathbf{I}_{N_r}^H\right) \left(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}}) \otimes \mathbf{I}_{N_r}\right) \underline{\mathbf{c}}_{j,k,g} \tag{4.32}$$

$$= \underline{\mathbf{c}}_{j,k,g}^H \Big(\left(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})^H \mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})\right) \otimes \mathbf{I}_{N_r}\Big) \underline{\mathbf{c}}_{j,k,g} \tag{4.33}$$

The Euclidean distance between symbols is instrumental in bounding the symbol error probability, since

$$P\left(\underline{\mathbf{a}}_k \rightarrow \underline{\mathbf{a}}'_k\right) \leq \exp\left(-\mathrm{d}^2\left(\underline{\mathbf{u}}^{(k)}_{jG+g}, \underline{\mathbf{u}}'^{(k)}_{jG+g}\right)\rho\right). \tag{4.34}$$

Tarokh has shown [45, 46] that if the channel coefficients $\underline{\mathbf{c}}_{j,k,g}$ are complex, zero-mean, independent Gaussian random variables (*i.e.* the channel is Rayleigh-fading), the symbol error probability is bounded by

$$P\left(\underline{\mathbf{a}}_k \rightarrow \underline{\mathbf{a}}'_k\right) \leq \left(\prod_{i=1}^{r}\lambda_i\right)^{-N_r}\left(E_s/4N_0\right)^{-rN_r} \tag{4.35}$$

where $\lambda_1 \ldots \lambda_r$ are the non-zero eigenvalues of $\left(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})^H\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})\right)$. Note that each $\lambda_i, i = 1 \ldots r$ is an eigenvalue of $\left(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})^H\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})\right) \otimes \mathbf{I}_{N_r}$ with multiplicity N_r (see equation (4.33)), since receive diversity is always present, irrelevant of the coding scheme. This explains the constant multiplier N_r in the exponent of (4.35).

Therefore, the rank of $(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})^H\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}}))$ determines the diversity advantage of this coding scheme. Since this is a diagonal matrix, it is full-rank ($r = N_tL$) iff none of its diagonal elements (the squared amplitudes of the elements of $\mathrm{Q}\underline{\mathbf{e}}$) is zero. For a given symbol error at position u ($\underline{\mathbf{e}} = (0, \ldots 0, e_u, 0, \ldots 0)^T$), this is equivalent to

$$\forall i \in [\![1, N_tL]\!],\ q_{i,u} \triangleq [\mathrm{Q}]_{i,u} \neq 0. \tag{4.36}$$

Since this must be true for all possible symbol errors, and therefore all possible $\underline{\mathbf{e}}$, this yields

$$\forall (i,u) \in [\![1, N_tL]\!]^2,\ q_{i,u} \neq 0. \tag{4.37}$$

This is the first condition: *In order to achieve full diversity,* Q *must contain no zero.* In the sequel, we will assume that this condition is fulfilled.

Coding gain criterion

Once the maximum transmit diversity N_tL is achieved, the coding gain (as defined by Tarokh, this denotes the product in the right-hand side of equation (4.35)) must be maximized. Therefore, we seek to maximize

$$\min_{u,e_u} \det\left(\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})^H\mathrm{diag}(\mathrm{Q}\underline{\mathbf{e}})\right), \tag{4.38}$$

or, equivalently if we assume that the errors e_u have constant modulus, maximize

$$\min_{u\in[\![1,N_tL]\!]} \prod_{i=1}^{N_tL} q_{i,u}q_{i,u}^* \tag{4.39}$$

under the energy constraint

$$\forall n, \quad \sum_{i=1}^{N_tL} q_{i,n}q_{i,n}^* = 1. \tag{4.40}$$

The Lagrange multipliers method yields

$$\forall (n,i), \quad |q_{i,n}| = \frac{1}{\sqrt{N_tL}}. \tag{4.41}$$

A satisfactory solution with respect to all three criterions (maximizing mutual information, diversity, and coding gain) is the Vandermonde matrix

$$\mathrm{Q} = \frac{1}{\sqrt{N_tL}} \begin{bmatrix} 1 & \theta_1 & \dots & \theta_1^{N_tL-1} \\ 1 & \theta_2 & \dots & \theta_2^{N_tL-1} \\ \vdots & \vdots & & \vdots \\ 1 & \theta_{N_tL} & \dots & \theta_N^{N_tL-1} \end{bmatrix} \tag{4.42}$$

where $\theta_k = e^{j\frac{\pi}{N_tL}(1+2k)}$, $k \in [\![1, N_tL]\!]$. This matrix was used in the simulations presented hereinafter.

4.4 Decoding

The streams-based structure of the code mandates the use of iterative decoding to cancel inter-streams interference. This can take several forms: either subtract the interference corresponding to the signals estimated at the previous stage, or use soft information throughout the process. In the first case, the interference is linearly subtracted to the received signal. The constellation symbols are subsequently separated by an LMMSE filter, and demapped independently, as exposed in Section 4.4.1. The alternative method, using soft-information at all stages of the iterative process, consists in computing the posterior probabilities of the constellation bits, given *a priori* information on some of the bits, and the received symbol. Due to its complexity, this solution has not been fully implemented, but is briefly analyzed in Section 4.4.2.

4.4.1 Iterative interference cancellation decoding

Since several streams are superimposed in our transmission scheme, successful decoding involves interference cancellation. In this section, we describe a Parallel Interference Cancellation (PIC) iterative scheme. The decoding process is done in two steps, similarly to the encoding operations: the linear precoding and spreading operations are undone using linear minimum mean-square error (LMMSE) estimation, whereas the binary convolutive code is decoded by a Viterbi decoder.

Recall the joint linear transformation of the signal by the linear precoder and the channel (equation (4.20)). At the first iteration of the decoder, the LMMSE estimate of the transmitted signal $\mathbf{A}$ is

$$\widehat{\mathbf{A}}_1 = \mathcal{R}_{\mathbf{AY}} \mathcal{R}_{\mathbf{YY}}^{-1} \mathbf{Y} \tag{4.43}$$

where

$$\mathcal{R}_{\mathbf{AY}} \triangleq \mathrm{E}_{\mathbf{K}} \left[\mathbf{A}\mathbf{Y}^H \right] = \mathbf{Q}^H \mathbf{W}^H \mathbf{H}^H \tag{4.44}$$

and

$$\mathcal{R}_{\mathbf{YY}} \triangleq \mathrm{E}_{\mathbf{K}} \left[\mathbf{Y}\mathbf{Y}^H \right] = \mathbf{H}\mathbf{W}\mathbf{Q}\mathbf{Q}^H \mathbf{W}^H \mathbf{H}^H + \sigma_{\mathbf{K}}^2 \mathbf{I} = \mathbf{H}\mathbf{H}^H + \sigma_{\mathbf{K}}^2 \mathbf{I}. \tag{4.45}$$

Since $\mathbf{H}$ is block-diagonal, the biased version of the LMMSE estimator can be decomposed into one $N_t \times N_r$ LMMSE per tone.

The streams are separated ($\widehat{\mathbf{A}}_1 = \left[\widehat{\mathbf{A}}_1^{(1)\,T} \dots \widehat{\mathbf{A}}_1^{(N_t)\,T} \right]$), de-mapped, de-interleaved, and fed into N_t Viterbi decoders, which estimate the original uncoded data $\hat{\mathrm{d}}^{(k,0)}$ separately for each stream k (see Figure 4.3). The inter-stream interference is canceled through an iterative process: at iteration $i > 1$, and for stream k, the estimated interference caused by other streams is subtracted from the received signal:

$$\mathbf{Y}_{\mathrm{ICi}}^{(\mathrm{k})} = \mathbf{Y} - \sum_{k' \neq k} \mathbf{H}\mathbf{W}_{k'} \left(\mathbf{I}_{GJ} \otimes \mathrm{Q} \right) \widehat{\widehat{\mathbf{A}}}_{i-1}^{(k')}, \tag{4.46}$$

where $\widehat{\widehat{\mathbf{A}}}_{i-1}^{(k')}$ is obtained by re-mapping the binary data decoded at the previous iteration $(i-1)$ for stream k'.

After the first iteration, the LMMSE filter is modified to take into account the cancellation of the interference: for stream k,

$$\widehat{\mathbf{A}}_i^{(k)} = \left(\mathbf{I}_{GJ} \otimes \mathrm{Q} \right)^H \mathbf{W}_k^H \mathbf{H}^H \left(\mathbf{H}\mathbf{W}_k \mathbf{W}_k^H \mathbf{H}^H + \sigma_{\mathbf{K}}^2 \mathbf{I} \right)^{-1} \mathbf{Y}_{\mathrm{ICi}}^{(\mathrm{k})}. \tag{4.47}$$

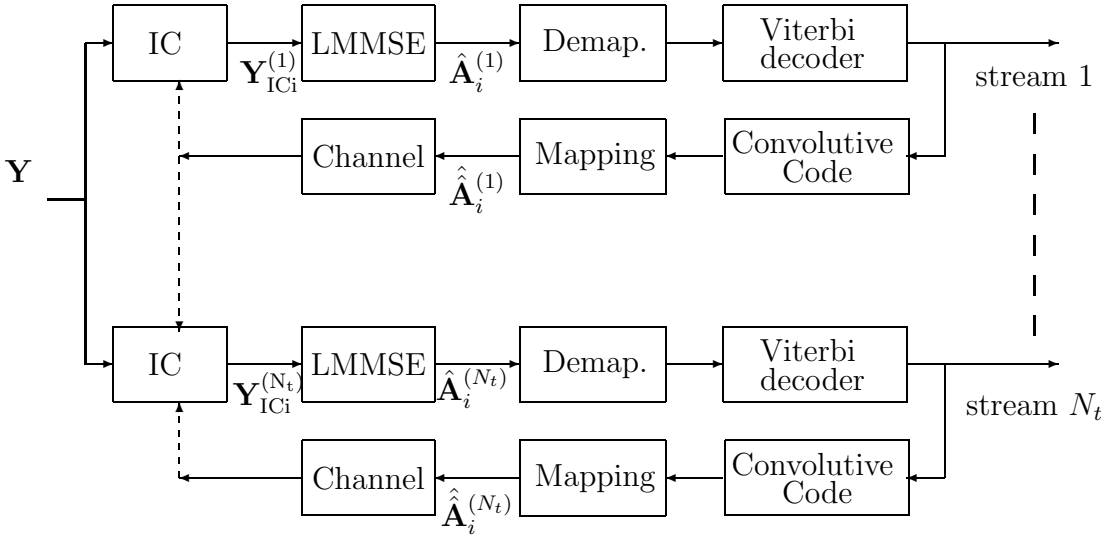


Figure 4.3: Iterative interference cancellation decoding

4.4.2 sIsO decoding advantages and limitations

The structure of the proposed code leaves many decoding possibilities: since the constellation mapping and the LCP stage can be regarded as a form of channel coding, the code itself is a concatenation of two codes. Therefore, iterative decoding is possible, and can take several forms, due to the particular code structure.

The iterative interference cancellation scheme described in Section 4.4.1 uses hard-decision: the output of the Viterbi decoder are bits, no additional information about the decision reliability is provided. However, as shown by Müller *et al.* in [47], making hard decisions before the end of the decoding process is in general suboptimal. Information about the reliability of the decision on each bit, or *soft information*, can be obtained by using a soft-output decoder, based on a Sum-Product algorithm [48]. Obviously, applying the turbo decoding principle [49] requires that both codes in the concatenated scheme can be sIsO decoded [50]. This is not an issue for the (outer) convolutive code, which can be sIsO-decoded by a Bahl-Cocke-Jelinek-Raviv (BCJR) decoder [51], which is a particular instance of the Sum-Product algorithm applicable to trellis codes. The inner code, however, takes the form of a noisy linear combination of constellation symbols. Soft

information about the original constellation symbols can be gathered if the noise level is known. For instance, improvements to the Viterbi algorithm have been proposed, such as the soft-output Viterbi algorithm (SOVA) [52]. However, SOVA does not exploit prior information, and therefore is not suitable for an iterative scheme. sIsO decoding of a noisy linear combination of symbols, such as the one considered here, can be achieved by performing a joint sIsO demapping over all the interfering symbols. Although this does not constitute interference cancellation *per se*, it achieves a similar effect in that prior information about the interfering signals is taken into account. Such a sIsO demapper is described in Annex 4.A.

Therefore, although sIsO decoding can improve the performance of the proposed Space-Time-Frequency Spreading scheme, preserving the soft information throughout the interference cancellation stage (even through soft-demapping) requires a rather accurate knowledge of the noise level to be expected for each symbol, which in our case must be varied according to the linear (LMMSE) decoding stage. Various experiments with soft decoders have shown that an average (over a complete frame) noise figure is not good enough for sIsO decoding.

4.5 Simulation results

All the simulations were carried out using the unpunctured rate $\frac{1}{2}$, 64-states convolutional code from the HiperLAN/2 standard [53]. The LCP stage operates on a BPSK constellation.

# of antennas ($N_t \times N_r$)	2×2	2×2	4×4	4×4
Constellation	BPSK	QPSK	BPSK	QPSK
# of OFDM symbols	8	4	4	2
Spectral efficiency	1 b/s/Hz	2 b/s/Hz	2 b/s/Hz	4 b/s/Hz

Table 4.1: Summary of the simulation parameters used in Section 4.5.

The number of uncoded data bits per stream per frame is kept constant (at 384 bits) across all the simulations, by adapting the number of OFDM symbols used in each frame. Since each frame is spread over several OFDM symbols, we considered both a quasi-static channel (it does not change during

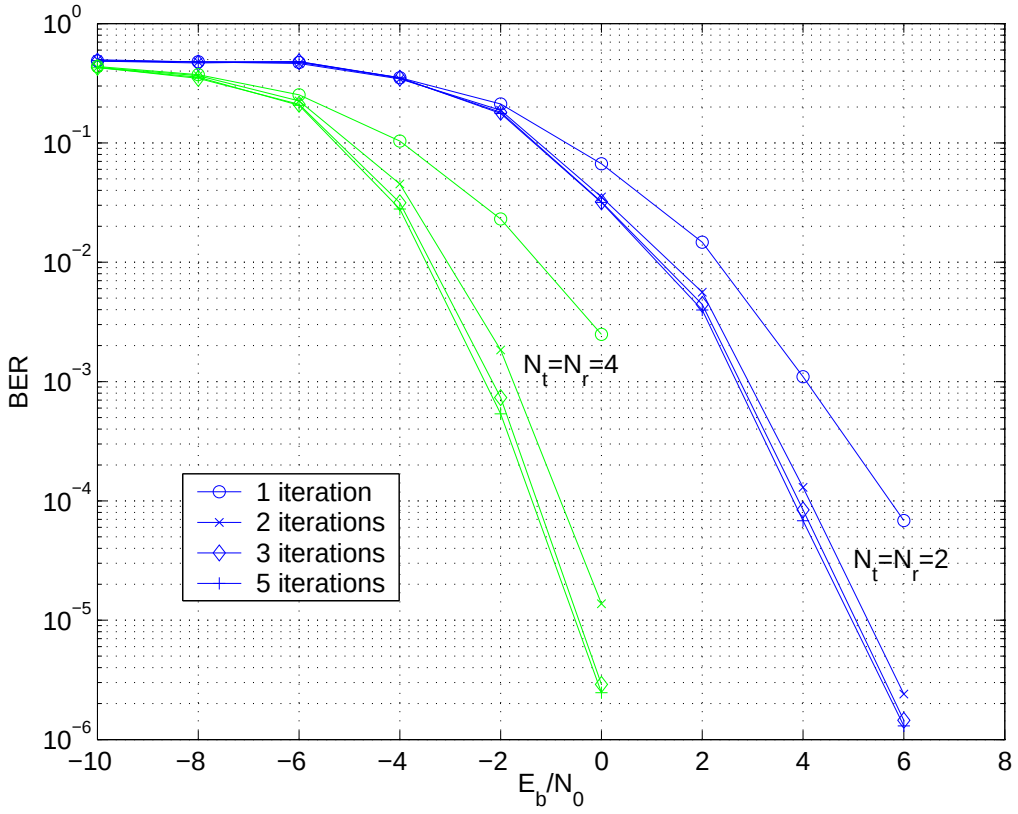


Figure 4.4: Influence of the number of decoding iterations on the BER with BPSK constellations.

a whole frame), and a fast-fading channel (in this case, the various OFDM symbols in each frame are transmitted over independent channel realizations). Table 4.1 summarizes the various simulation parameters.

4.5.1 Influence of the number of decoder iterations

The number of necessary iterations to reach a steady state has been investigated, and the results are presented in Figures 4.4 and 4.5, respectively for modulations using BPSK and QPSK constellations, for the case of quasi-static 2×2 and 4×4 channels. The total channel diversity is $4N_tN_r$ (each impulse response is comprised of four i.i.d. Gaussian random coefficients, and are independent across antennas). The size of the linear precoder is $4N_t$. The

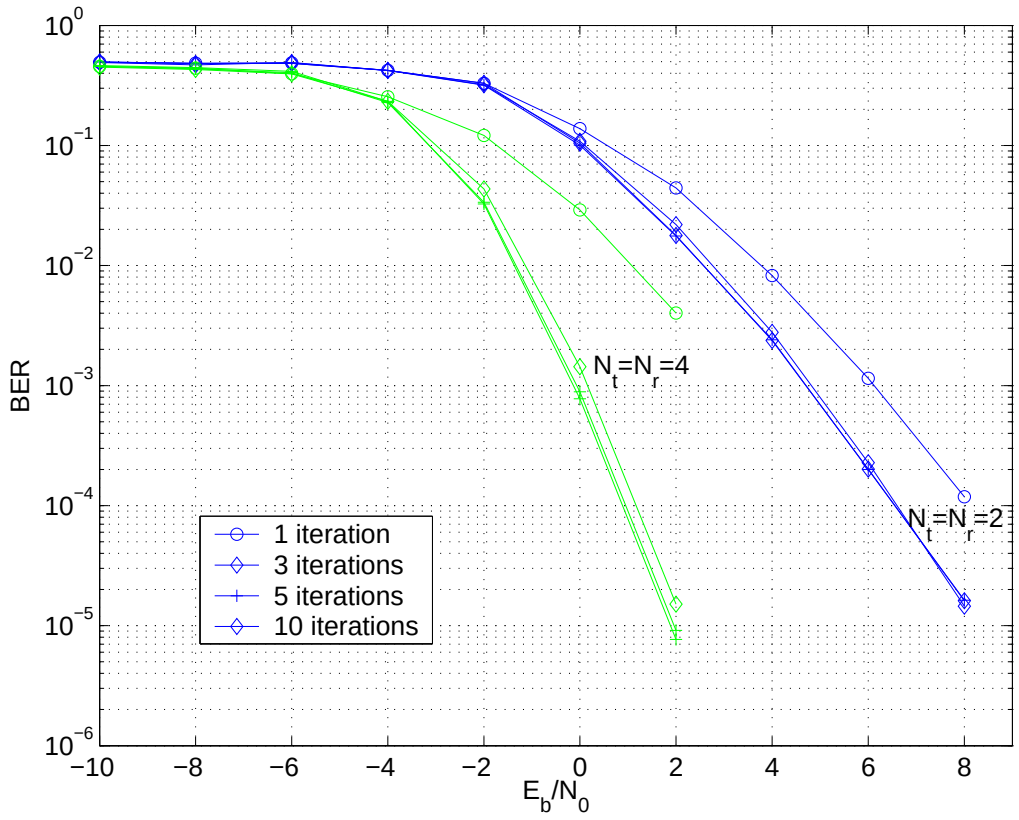


Figure 4.5: Influence of the number of decoding iterations on the BER with QPSK constellations.

BER at the output of the Viterbi decoder is plotted after various numbers of iterations of the decoding process (note that no interference cancellation is performed at iteration 1).

Using a BPSK modulation (Figure 4.4), for both the 2×2 and 4×4 channels, the first interference cancellation step yields a significant drop in BER. The decoder reaches its steady state at the third iteration, and further executions yield no performance gain, the curves corresponding to the third and fifth iterations being almost indistinguishable. Conversely, with QPSK symbols, more iterations are necessary to reach an acceptable error rate: Figure 4.5 shows that around 5 iterations are necessary to reach the steady-state, again both for the 2×2 and 4×4 channels.

4.5.2 Comparison to other coding schemes

Figure 4.6 compares the performance of the proposed streams approach to El Gamal's Threaded Space-Time Code (TSTC) approach [43] over the previously described channel. The TSTC approach is similar to STFS in that various, separately encoded streams of data are transmitted from different Tx antennas, the streams being rotated across the antennas. It is equivalent to the STFS method except that Q is replaced by an identity matrix. The curves show that SFTS consistently outperforms the TSTC approach by about 1dB for frame error rates of 10^{-3}, for both the quasi-static and fast-fading channels.

Figures 4.7 and 4.8 compare the STFS approach to a Natural Space-Time Code (NSTC) [54] Bit-Interleaved Coded Modulation (BICM) using a 64-state code of rate $\frac{1}{2}$ over a quasistatic BRAN A [55] channel. Contrarily to STFS and TSTC, the NSTC approach uses only a single binary encoder. The coded bits are interleaved and separated into N_t groups, which are independently mapped and transmitted from the N_t Tx antennas. For BPSK constellations (Figure 4.7), STFS and NSTC-BICM have very similar performance, as evidenced by the almost identical curves. With QPSK constellations (Figure 4.8), the BICM approach remains more efficient, with a 1dB advantage in terms of E_b/N_0 w.r.t. STFS.

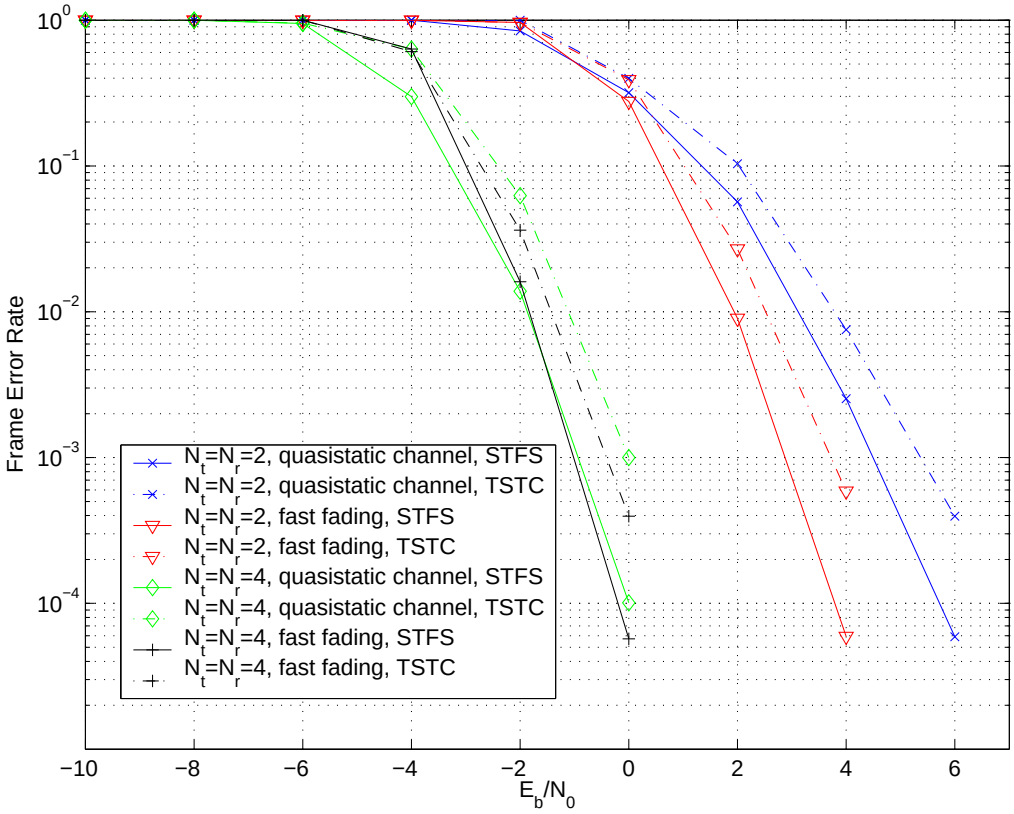


Figure 4.6: Comparison of STFS with TSTC, for BPSK modulations.

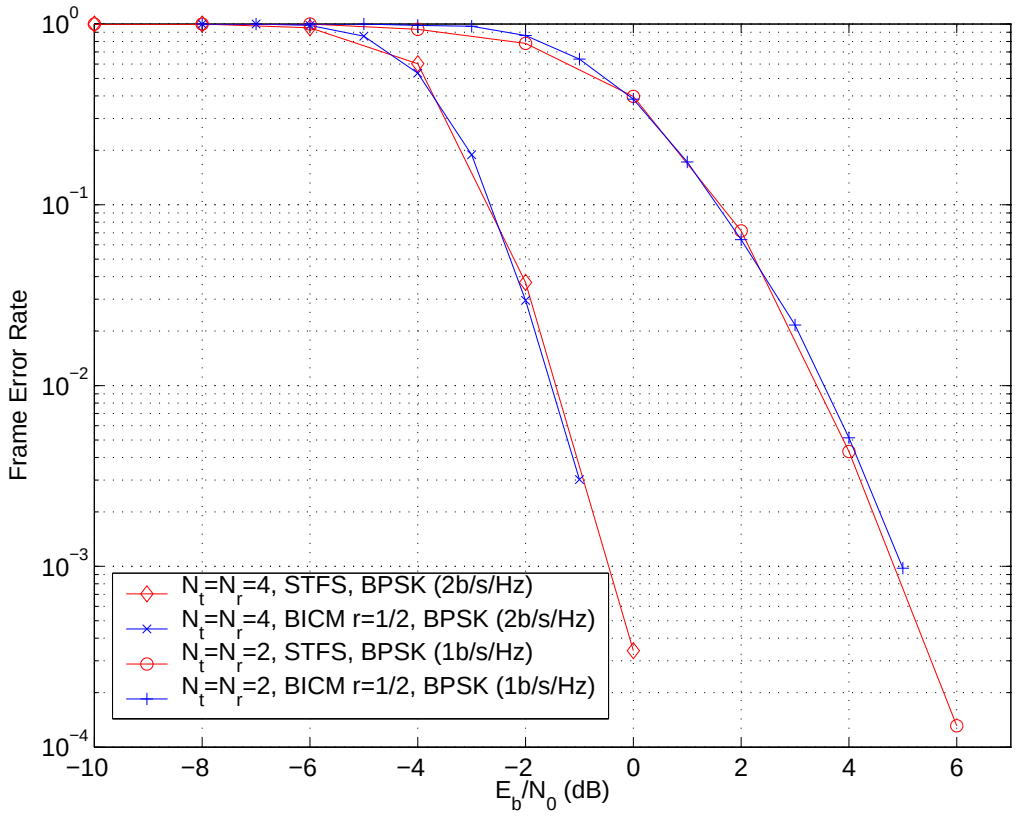


Figure 4.7: Comparison of STFS with BICM over a BRAN A quasistatic channel, for BPSK modulations.

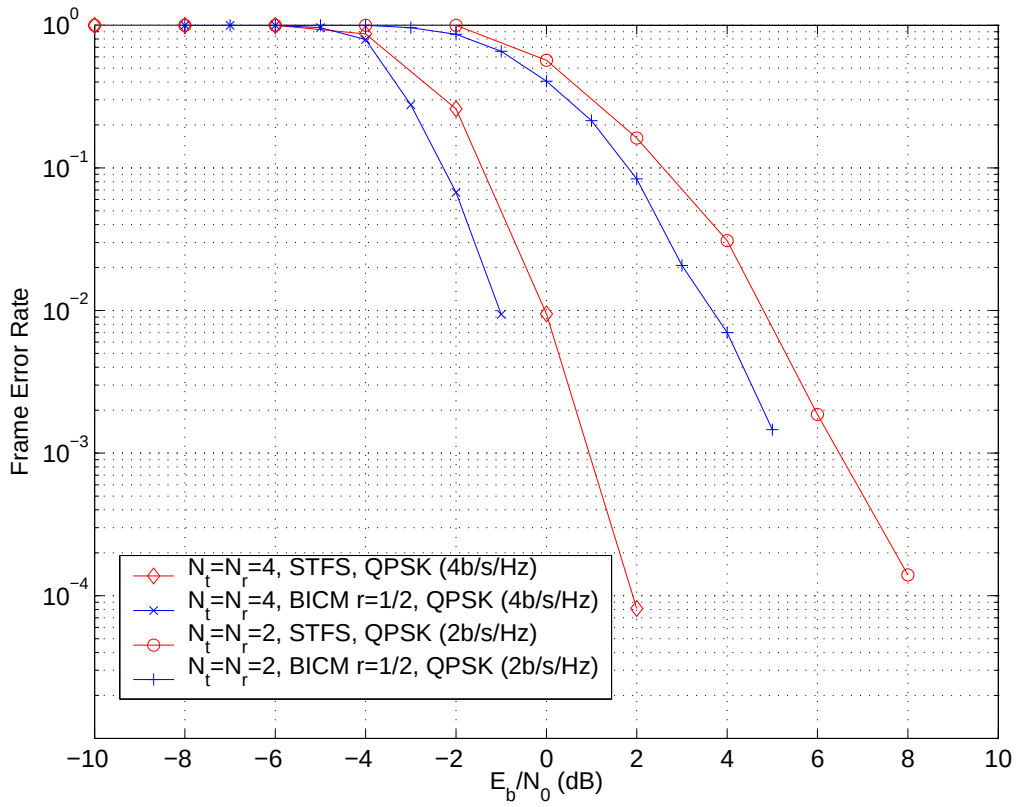


Figure 4.8: Comparison of STFS with BICM over a BRAN A quasistatic channel, for QPSK modulations.

4.6 Conclusion

In this chapter, we proposed a Space-Time-Frequency Spreading coding scheme for MIMO OFDM systems, optimized to exploit the maximum space, frequency and possibly time diversity present in the channel. This code is easily adapted to an arbitrary number of Tx antennas. It was shown to be particularly suitable for iterative PIC decoding, and, in the case of BPSK modulation, to achieve the same performance as the BICM approach, using only Viterbi decoders instead of BCJR.

4.A Appendix: Soft-Demapper

In order to expose the principle of the soft demapper, let us consider the transmission over an AWGN channel of a complex vector constellation symbol $\underline{\mathbf{a}}$ from the set of constellation symbols $\mathcal{C}$. Assume that the cardinality of $\mathcal{C}$ is 2^C for some integer C, and let $b_1 \dots b_C$ denote the binary representation of the transmitted symbol. The received signal is

$$\underline{\mathbf{y}} = \mathbf{H}\underline{\mathbf{a}} + \underline{\mathbf{n}}, \tag{4.48}$$

where $\underline{\mathbf{n}}$ is i.i.d complex random Gaussian noise of variance $\sigma^2_{\underline{\mathbf{n}}}$, and $\mathbf{H}$ is a matrix representing the channel fading characteristics. Since the noise is unknown, the received symbol $\underline{\mathbf{y}}$ is related to the transmitted value through its probability distribution

$$\mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}}) = \frac{1}{\sigma_{\underline{\mathbf{n}}}\sqrt{2\pi}} \exp^{-\frac{|\underline{\mathbf{y}}-\mathbf{H}\underline{\mathbf{a}}|^2}{2\sigma^2_{\underline{\mathbf{n}}}}} \tag{4.49}$$

(the channel $\mathbf{H}$ is assumed to be perfectly known, therefore we do not include it in the probability notations). The probability that the transmitted bit b_i is zero, given the received signal $\underline{\mathbf{y}}$, is obtained by marginalizing over the other bits:

$$\mathrm{P}(b_i = 0|\underline{\mathbf{y}}) = \sum_{\underline{\mathbf{a}}/B_i[\underline{\mathbf{a}}]=0} \mathrm{P}(\underline{\mathbf{a}}|\underline{\mathbf{y}}), \tag{4.50}$$

where $B_k[\underline{\mathbf{a}}]$ denotes the kth bit of the binary representation of $\underline{\mathbf{a}}$. Using Bayes' Law,

$$\mathrm{P}(b_i = 0|\underline{\mathbf{y}}) = \sum_{\underline{\mathbf{a}}/B_i[\underline{\mathbf{a}}]=0} \frac{\mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}})\mathrm{P}(\underline{\mathbf{a}})}{\mathrm{P}(\underline{\mathbf{y}})}. \tag{4.51}$$

For binary signals, the probability is commonly replaced by the equivalent log-likelihood ratio (LLR)

$$\mathrm{LLR}(b_i|\underline{\mathbf{y}}) \triangleq \log\left(\frac{\mathrm{P}(b_i = 1|\underline{\mathbf{y}})}{\mathrm{P}(b_i = 0|\underline{\mathbf{y}})}\right) \tag{4.52}$$

$$= \log\left(\frac{\sum_{\underline{\mathbf{a}}/B_i[\underline{\mathbf{a}}]=1} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}})\mathrm{P}(\underline{\mathbf{a}})}{\sum_{\underline{\mathbf{a}}/B_i[\underline{\mathbf{a}}]=0} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}})\mathrm{P}(\underline{\mathbf{a}})}\right) \tag{4.53}$$

Note that the *a priori* probability to observe the symbol $\underline{\mathbf{a}}$ appears in eq. (4.53). When dealing only with binary codes, no prior information on the transmitted symbol is available at the demapping stage, and therefore all symbols

are assumed equally likely ($\mathrm{P}(\underline{\mathbf{a}}) = 1/2^C$). However, in the case of linear precoding with several interfering streams, inter-stream interference cancellation is necessary. In this case, if other streams have already been decoded, some prior information about the interfering signals (*i.e* bits $b_j, j \neq i$) is available, and can be used at the demapping stage: assuming that the b_is can be considered independent, let us rewrite the *a priori* probability of $\underline{\mathbf{a}}$ as

$$\mathrm{P}_a(\underline{\mathbf{a}}) = \prod_{k=1\ldots C} \mathrm{P}_a(b_k = B_k[\underline{\mathbf{a}}]), \tag{4.54}$$

The independence lets us marginalize over the $b_j, j \neq i$, which yields

$$\begin{aligned}\mathrm{LLR}(b_i|\underline{\mathbf{y}}) = &\log\left(\frac{\mathrm{P}_a(b_i=1)}{\mathrm{P}_a(b_i=0)}\right) \\ &+ \log\left(\frac{\sum_{\underline{\mathbf{a}}/b_i[\underline{\mathbf{a}}]=1} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}}) \prod_{\substack{k=1\ldots C,\\ k\neq i}} \mathrm{P}_a(b_k = B_k[\underline{\mathbf{a}}])}{\sum_{\underline{\mathbf{a}}/b_i[\underline{\mathbf{a}}]=0} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}}) \prod_{\substack{k=1\ldots C,\\ k\neq i}} \mathrm{P}_a(b_k = B_k[\underline{\mathbf{a}}])}\right)\end{aligned} \tag{4.55}$$

Letting $\mathrm{LLR}_a(b_i) \triangleq \log\left(\frac{\mathrm{P}_a(b_i=1)}{\mathrm{P}_a(b_i=0)}\right)$ denote the *a priori* LLR of b_i, and since $\mathrm{P}_a(b_k = 1) = \frac{\exp(\mathrm{LLR}_a(b_k))}{1+\exp(\mathrm{LLR}_a(b_k))}$ and $\mathrm{P}_a(b_k = 0) = \frac{1}{1+\exp(\mathrm{LLR}_a(b_k))}$, equation (4.55) becomes

$$\begin{aligned}&\mathrm{LLR}(b_i|\underline{\mathbf{y}}) = \\ &\qquad \mathrm{LLR}_a(b_i) + \log\left(\frac{\sum_{\underline{\mathbf{a}}/b_i[\underline{\mathbf{a}}]=1} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}}) \frac{\prod_{\substack{k=1\ldots C,\\ k\neq i,\\ B_k[\underline{\mathbf{a}}]=1}} \exp(\mathrm{LLR}_a(b_k))}{\prod_{\substack{k=1\ldots C,\\ k\neq i}} 1+\exp(\mathrm{LLR}_a(b_k))}}{\sum_{\underline{\mathbf{a}}/b_i[\underline{\mathbf{a}}]=0} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}}) \frac{\prod_{\substack{k=1\ldots C,\\ k\neq i,\\ B_k[\underline{\mathbf{a}}]=1}} \exp(\mathrm{LLR}_a(b_k))}{\prod_{\substack{k=1\ldots C,\\ k\neq i}} 1+\exp(\mathrm{LLR}_a(b_k))}}\right)\end{aligned} \tag{4.56}$$

$$= \mathrm{LLR}_a(b_i) + \log\left(\frac{\sum_{\underline{\mathbf{a}}/b_i[\underline{\mathbf{a}}]=1} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}}) \prod_{\substack{k=1\ldots C,\\ k\neq i,\\ B_k[\underline{\mathbf{a}}]=1}} \exp(\mathrm{LLR}_a(b_k))}{\sum_{\underline{\mathbf{a}}/b_i[\underline{\mathbf{a}}]=0} \mathrm{P}(\underline{\mathbf{y}}|\underline{\mathbf{a}}) \prod_{\substack{k=1\ldots C,\\ k\neq i,\\ B_k[\underline{\mathbf{a}}]=1}} \exp(\mathrm{LLR}_a(b_k))}\right) \tag{4.57}$$

This describes a sIsO demapper, whereby prior information about the interference can be exploited. Note that the logarithm does not contain any

reference to $\mathrm{LLR}_a(b_i)$: this part is denoted the extrinsic *a posteriori* information. A similar soft-demapper was proposed by ten Brink in [56]. However, we use it in place of interference canceler, whereas in [56] it is used to improve the decoding of a single convolutive code.

The complexity of soft-demapping makes it a prohibitively complex IC scheme (note that the sums in (4.57) enumerate all possible constellation symbols, and therefore have the same complexity as an ML decoder – although the approximation proposed by Vikalo and Hassibi in [57], along the lines of the Sphere Decoder [58], can somewhat reduce this complexity). However, note that the ML decoder is only optimal in the case of no available prior information, whereas a soft-demapping scheme can exploit the prior knowledge provided by iterative decoding.

Part II

MIMO Channel modeling and CSIT

Chapter 5

CSIT and calibration issues in MIMO systems

In this chapter, we consider the use of perfect and degraded channel state information by the transmitter, when the (mutual information-maximizing) eigenwaterfilling method is used. We investigate the possible discrepancies between the actual propagation channel and the available CSIT, in particular the phase and amplitude perturbations due to the RF components, as well as the issues associated with their absolute calibration.

5.1 Introduction

Knowledge of the channel state by the transmitter of a communications system has been demonstrated to be beneficial to wireless communications. Various methods have been proposed to shape the transmitted signal according to the channel state information (CSI) available at the transmitter [21], for both single-antenna and multiple-antenna settings. A wide variety of methods are available for this, ranging from simple power control to frequency-domain or eigendomain signal shaping [59] to time-reversal methods [60] aiming at reducing the channel delay spread. All these methods utilize CSI, and getting reliable CSI at the transmitter where it is needed has proved to be an issue.

The methods available to gather CSI at the transmitter (CSIT) can be classified into two categories, relying either on feedback or on reciprocity. In the feedback scheme, the channel is estimated on one side of the link, and is fed back to the other side, sharing the bandwidth resource with the data. For simple channels, this method has been successfully applied, power control being one particular instance of a feedback scheme. However, for more complex channels such as MIMO frequency-selective channels, full CSI is constituted of multiple impulse responses corresponding to each transmit (Tx)-receive (Rx) antenna pair, and thus the amount of bandwidth consumed by the feedback link can become prohibitive, in particular for channels with relatively small coherence time such as those encountered in wireless communications, where the CSI has to be updated often. A wide class of proposed solutions to the problem of large feedback throughput is to use a suboptimal shaping at the transmitter, that only makes use of part of the CSI. Covariance feedback [61], SNR-based Tx antenna selection [62], and Partial Phase Combining [63] for instance, are different instances of partial CSIT schemes. Of course, none of the partial CSIT schemes achieve the performance of full CSIT. Furthermore, quantization and latency issues in feedback schemes must be well understood (see *e.g.* [64]).

The other avenue for the transmitter to gain knowledge of the channel state is to make use of reciprocity. The reciprocity principle is based on the property that electromagnetic waves traveling in both directions will undergo the same physical perturbations (*i.e.* reflection, refraction, diffraction, etc...). Therefore, in Time-Division Duplex (TDD) systems where the link

operates on the same frequency band in both directions, the impulse response of the channel observed between any two antennas should be the same regardless of the direction. Application of the reciprocity principle lifts the requirement for a continuous feedback of the channel estimates while still allowing to make use of CSIT in order to optimize the transmission. Specific issues related to channel estimation through reciprocity will be covered in detail in Chapter 6.

In this chapter, we study the issues of inaccurate channel knowledge at the transmitter and improper calibration in the exploitation of channel information by the transmitter. The mutual information criterion was studied, since it is particularly suitable to separate the effects of transmitter and receiver calibration. Although the mutual information metric does not take channel coding into account, and therefore might seem far away from the actual system performance evaluation, it should be noted that, in the case of quasi-static fading channels, the channel outage probability provides an upper bound to the frame error rate achievable by any code. In some simulated cases [13], the spectral efficiency of a coded system operating at a FER of 10^{-2} has been seen to reach 80% of the 10^{-2} outage capacity of the channel.

5.2 Eigenwaterfilling with perfect CSIT

Let us consider the classical flat-fading AWGN MIMO channel introduced in Section 2.3. It is represented by

$$\underline{\mathbf{y}}_k = \mathbf{H}\underline{\mathbf{x}}_k + \underline{\mathbf{n}}_k. \tag{5.1}$$

With Gaussian input, the mutual information can be written as a function of the channel input covariance $\mathbf{D} \triangleq \mathrm{E}_x\left[\underline{\mathbf{x}}\underline{\mathbf{x}}^H\right]$ (the Gaussian circularly symmetric input distribution itself maximizes the mutual information for a given $\mathbf{D}$):

$$\mathcal{I}\left(\underline{\mathbf{x}};\underline{\mathbf{y}}\right) = \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{D}\mathbf{H}^H\right). \tag{5.2}$$

If $\mathbf{H}$ is perfectly known by the transmitter, the following so-called *eigenwaterfilling* method yields the optimal $\mathbf{D}(\mathbf{H})$: let $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^H$ denote the singular value decomposition (SVD) of the channel matrix. $\mathbf{U}$ (resp. $\mathbf{V}$) is a unitary $N_r \times N_r$ (resp. $N_t \times N_t$) matrix, containing the left (resp. right)

singular vectors of $\mathbf{H}$. $\mathbf{S}$ is an $N_r \times N_t$ diagonal matrix containing the singular values, which we will denote by s_i, in non-increasing order ($s_i \geq s_{i+1}$). Let K denote the number of non-zero singular values of $\mathbf{H}$.

The optimal input covariance matrix is given by $\mathbf{D}(\mathbf{H}) = \mathbf{V}\mathbf{\Gamma}\mathbf{V}^H$ where $\mathbf{\Gamma} = \mathrm{diag}(\gamma_1, \ldots, \gamma_{N_t})$ has diagonal elements given by

$$\gamma_i = \begin{cases} \left[\xi - \frac{N_0}{s_i^2}\right]_+ & \text{if} \quad i = 1 \ldots K \\ 0 & \text{if} \quad i > K \end{cases} \tag{5.3}$$

where ξ is a Lagrange multiplier set to satisfy the power constraint, i.e., it is the solution of the equation

$$\mathrm{tr}\,(\mathbf{D}) = \sum_{i=1}^{K} \gamma_i = \gamma. \tag{5.4}$$

This is called an eigenwaterfilling method since the squared singular values s_i^2 constitute the eigenvalues of $\mathbf{H}\mathbf{H}^H$. The mutual information achieved by this scheme is therefore

$$\begin{aligned} C_{\mathrm{WTF}} &\triangleq \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{D}(\mathbf{H})\mathbf{H}^H\right) && (5.5) \\ &= \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{S}\mathbf{\Gamma}\mathbf{S}^H\right) && (5.6) \\ &= \sum_{i=1}^{i_0} \log\left(\frac{s_i^2\xi}{N_0}\right) && (5.7) \end{aligned}$$

where we first used the identity $\det(\mathrm{AB}) = \det(\mathrm{BA})$ with $\mathrm{A} = \mathrm{U}$ and $\mathrm{B} = \mathrm{U}^H + \frac{1}{N_0}\mathbf{S}\mathbf{V}^H\mathbf{V}\mathbf{\Gamma}\mathbf{V}^H\mathbf{V}\mathbf{S}^H\mathbf{U}^H$, and the fact that U is unitary, and the diagonal structure of the resulting matrix inside the determinant. The number of used eigenmodes i_0 is determined by

$$i_0 = \arg\min_{j=1\ldots K} \frac{1}{j}\left(\gamma + N_0 \sum_{i=1}^{j} \frac{1}{s_i^2}\right). \tag{5.8}$$

(see proof in Appendix 5.A).

The eigenwaterfilling method can be generalized to the frequency-selective case, using the principle of OFDM transmission: assuming that the frequency-selective channel is comprised of a finite number of resolvable paths, it can be split into N subbands that become individually frequency-flat when $N \to +\infty$. For large but finite values of N, the waterfilling algorithm can be applied jointly over the singular values corresponding to all the subbands, in order to determine an input signal distribution matching the channel characteristics both in space and frequency (indeed, a waterfilling approach was suggested by Shannon as soon as 1949 [4] to distribute the power over the frequencies of a SISO frequency-flat channel having a non-flat noise spectrum).

In general, the joint distribution of the singular values over all frequency bands is not equal to the product of the (per-subband) marginal distributions. However, for the sake of simplicity, we will consider only frequency-flat channels in the sequel, where a qualitative analysis of various (perfect and imperfect) CSIT situations is proposed. First, we show in the next section that the advantage of CSIT over the CSIR-only situation vanishes as high SNR. Then, in Section 5.3, we evaluate the impact of imperfect channel knowledge and imperfect calibration on eigenbeamforming.

5.2.1 Asymptotic (high SNR) behaviour

First, let us rewrite the mutual information in the CSIR case – see equation (2.17) – in terms of the channel singular values:

$$
\begin{aligned}
C_{\mathbf{I}} &= \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{U}\mathbf{S}\mathbf{V}^H\mathbf{V}\mathbf{S}^H\mathbf{U}^H\right) && (5.9)\\
&= \log\det\left(\mathbf{U}^H\mathbf{U} + \frac{1}{N_0}\mathbf{S}\mathbf{S}^H\right) && (5.10)\\
&= \sum_{i=1}^{K}\log\left(1 + \frac{s_i^2}{N_0}\right). && (5.11)
\end{aligned}
$$

It is obvious from equation (5.8) that for asymptotically small N_0, $i_0 = K$. Note also that $\gamma = \mathrm{tr}\,(\mathbf{I}) = N_t$, in order to impose the same power constraint in both cases. The difference between the waterfilling capacity

and the capacity without channel knowledge is

$$C_{\mathrm{WTF}} - C_{\mathbf{I}} = \sum_{i=1}^{K} \log\left(\frac{s_i^2 \xi_K}{N_0}\right) - \log\left(1 + \frac{s_i^2}{N_0}\right) \tag{5.12}$$

$$= \sum_{i=1}^{K} \log\left(\frac{s_i^2 \xi_K}{N_0 + s_i^2}\right). \tag{5.13}$$

Since $\lim_{N_0 \to 0} \xi_K = \frac{\gamma}{K}$, we obtain

$$\lim_{N_0 \to 0} C_{\mathrm{WTF}} - C_{\mathbf{I}} = \sum_{i=1}^{K} \log\left(\frac{\gamma}{K}\right) \tag{5.14}$$

$$= K \log\left(\frac{N_t}{K}\right). \tag{5.15}$$

Note that for the commonly used model where the components of $\mathbf{H}$ are random Gaussian i.i.d. variables, $K = \min(N_t, N_r)$ with probability 1. Therefore, we need to distinguish two cases:

- if $N_r \geq N_t$, $K = N_t$, and $\lim_{N_0 \to 0} C_{\mathrm{WTF}} - C_{\mathbf{I}} = 0$.

- if $N_r < N_t$, $K = N_r$, and $\lim_{N_0 \to 0} C_{\mathrm{WTF}} - C_{\mathbf{I}} = N_r \log\left(\frac{N_t}{N_r}\right)$. Intuitively, this gain corresponds to the amount of energy saved by the transmitter not transmitting in the subspace that the receiver does not "see".

Furthermore, since the increase in mutual information due to CSIT (5.15) tends to a constant at high SNR, while $\lim_{N_0 \to 0} C_{\mathbf{I}} = +\infty$, the use of CSIT represents a vanishing mutual information gain at high SNR.

This phenomenon is evidenced by the simulation results of Figure 5.1. The Cumulative Distribution Function (CDF) of the mutual information of a $(N_t = 4, N_r = 5)$ channel has been evaluated through Monte-Carlo simulations for various SNRs, for both the cases of perfect CSIT (using the optimal eigenwaterfilling), and no CSIT. At -10dB SNR, CSIT provides an almost twofold increase in mutual information over spatially white transmission, with the 10^{-2} outage capacity going from .26 (no CSIT) to .46 (perfect CSIT).

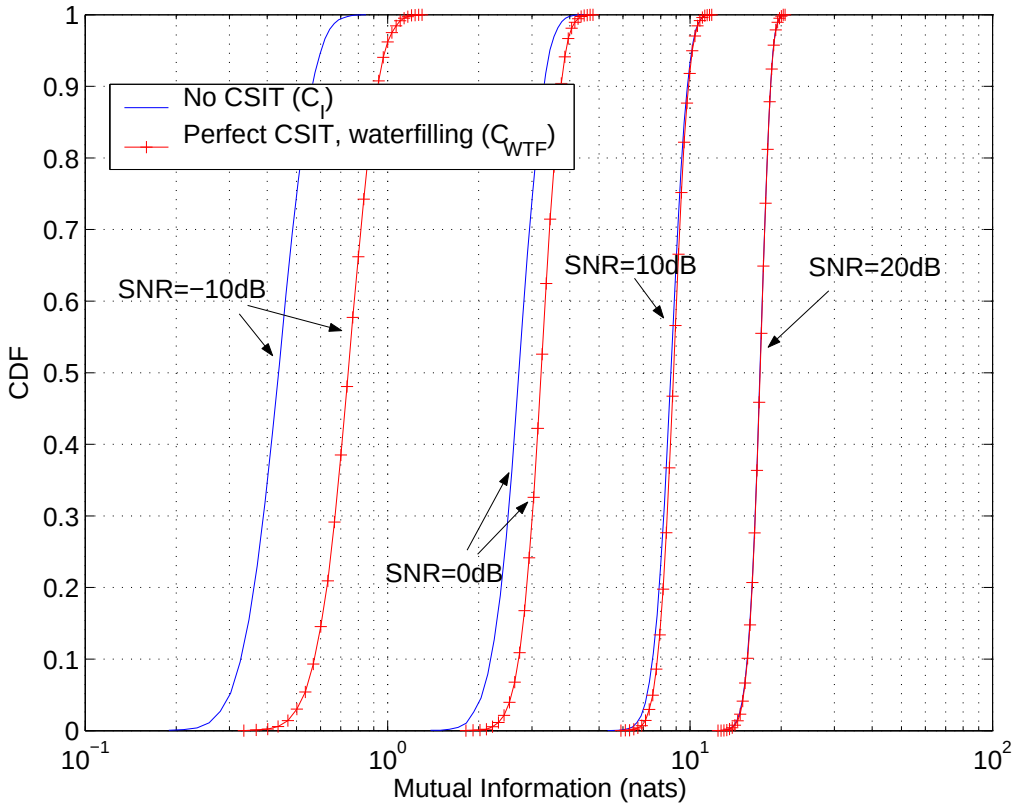


Figure 5.1: Mutual information CDF for perfect CSIT and no CSIT situations at various SNRs.

5.3 Imperfect CSIT

In this section, we tackle the issue of imperfect (noisy) CSIT, and the possible perturbation of CSIT by calibration problems. Although the imperfect CSIR has been treated by Médard, in terms of the reduction of mutual information in [65], and by Hassibi and Hochwald in [66] for the case of a separate training and data transmission phase, a symmetric analysis of the case of imperfect CSIT is missing. Although no analytical results are known to date, we try in this section to draw some conclusions based on simulations results.

The impairments to CSIT can be broadly classified into two categories, namely estimation noise and calibration issues, which need to be addressed separately.

5.3.1 Additive channel estimation noise

In most practical cases, the channel matrix is not known perfectly, but rather must be estimated. Dedicated pilot signals consume part of the total energy allowed to the transmission, and thus must be kept to a minimum. Also, in the specific case of CSIT through channel feedback, where the channel estimate is transmitted back from the receiver (where it can be reliably estimated) to the transmitter (where it can be put to some use), and the discretization that this feedback link implies, also adds to the uncertainty on $\mathbf{H}$.

Therefore, let us study the influence of using a noisy channel estimate $\widehat{\mathbf{H}}$ instead of the true $\mathbf{H}$ in the waterfilling algorithm of Section 5.2. Specifically, we model the coefficients of the error $\mathbf{E} \triangleq \widehat{\mathbf{H}} - \mathbf{H}$ as Gaussian i.i.d. random variables of the same variance σ_e^2. The mutual information of the system adapting the transmit covariance to $\hat{\mathbf{H}}$ whereas the actual channel is $\mathbf{H}$ is

$$C_n \triangleq \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{D}(\widehat{\mathbf{H}})\mathbf{H}^H\right), \tag{5.16}$$

The waterfilling algorithm now operates on the singular values of $\widehat{\mathbf{H}}$. Let us write its SVD as $\hat{\mathbf{H}} = \mathbf{U}'\mathbf{S}'\mathbf{V}'^H$. Assuming that the transmitter applies the eigenwaterfilling method to the noisy channel estimate, $\mathbf{D}(\widehat{\mathbf{H}}) = \mathbf{V}'\mathbf{\Gamma}'\mathbf{V}'^H$ where $\mathbf{\Gamma}'$ has diagonal elements obtained by waterfilling over the diagonal elements of $\mathbf{S}'$. Unfortunately, this perturbation is hard to analyze, since the

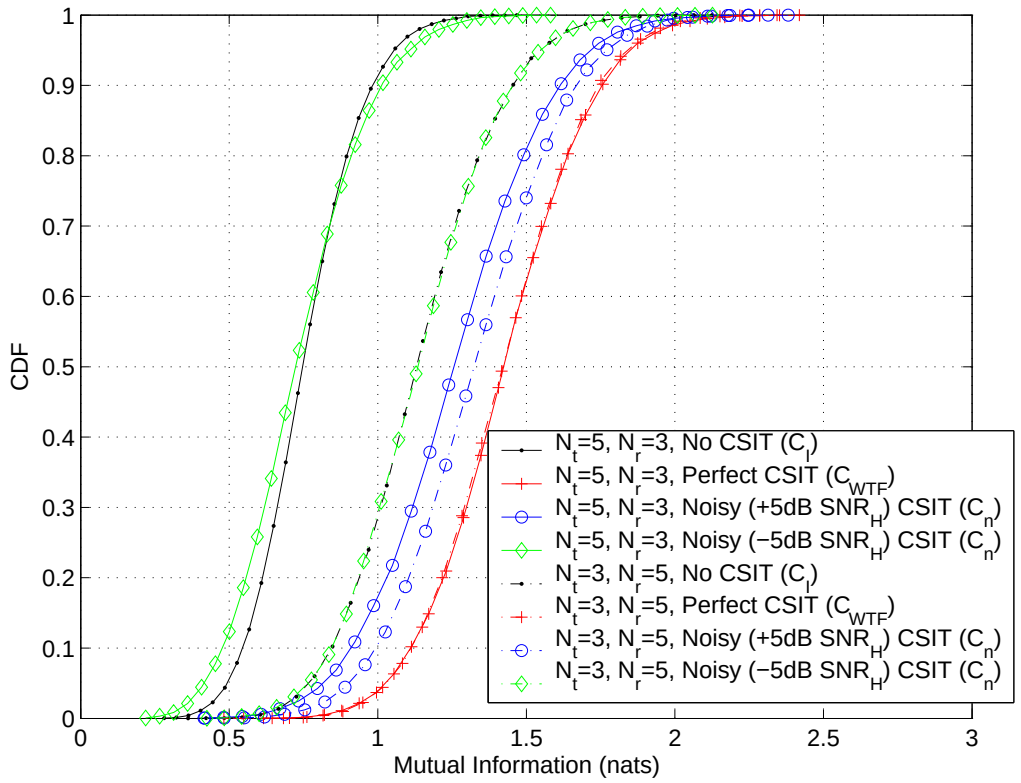


Figure 5.2: Mutual information CDF for perfect, noisy and no CSIT situations, SNR = -5dB, channel estimate SNR = 5dB.

mutual information

$$C_n = \log\det\left(\mathbf{I}+\frac{1}{N_0}\mathbf{U}\mathbf{S}\mathbf{V}^H\mathbf{V}'\mathbf{\Gamma}'\mathbf{V}'^H\mathbf{V}\mathbf{S}^H\mathbf{U}^H\right) \quad (5.17)$$

$$= \log\det\left(\mathbf{I}+\frac{1}{N_0}\mathbf{S}\mathbf{V}^H\mathbf{V}'\mathbf{\Gamma}'\mathbf{V}'^H\mathbf{V}\mathbf{S}^H\right) \quad (5.18)$$

is perturbed in two places: first, the singular subspaces are not exactly known ($\mathbf{V}^H\mathbf{V}' \neq \mathbf{I}$), and the perturbation on the singular values of $\mathbf{S}$ (since $\mathbf{S}' \neq \mathbf{S}$) passes through the nonlinear effects of the waterfilling method.

Therefore, we resorted to simulation to evaluate the impact of additive noise on CSIT. The mutual information CDF curves of Figure 5.2 have been obtained for 5×3 and 3×5 antennas systems, for channel estimate SNRs

(defined as $\mathsf{SNR}_{\mathbf{H}} \triangleq \frac{\mathrm{E}_{\mathbf{H}}\left[||\mathbf{H}||_F^2\right]}{\sigma_e^2 N_t N_r}$, $||\cdot||_F$ denoting the Frobenius norm) of +5dB and -5dB. The signal SNR ($\frac{1}{N_0}$) is set to 5dB for all curves. For the 3×5 system, increasing the channel estimation noise level clearly decreases the mutual information, down to the point where there is no difference between CSIT and the absence thereof. In the 5×3 case, the influence of additive channel estimation noise is even worse, the achieved mutual information in the case of very noisy channel estimate ($\mathsf{SNR}_{\mathbf{H}} = -5$dB) is lower than what is achieved with a spatially white transmission.

5.3.2 Convolutive channel impairments

Convolutive impairments can result from discrepancies in the behaviour of the electronic circuitry associated with different antennas. For instance, signals that are transmitted in-phase in the signal domain can yield out-of-phase electromagnetic signals, because the RF components forming the radio ways introduce slightly different delays. This behaviour is extremely difficult to avoid in practice, since the error lies basically in the tolerance of the components. Nevertheless, assuming that it can be accurately measured, it can be compensated. This constitutes an absolute calibration. Note that in Chapter 6, we delve in more detail into the calibration issue.

Let us now study the effects on mutual information of these convolutive perturbations, in the case where they are not (or not properly) compensated. Let $H(\tau)$ denote the $N_r \times N_r$ MIMO impulse response of the channel (each one of the $N_t N_r$ elements is the impulse response of the channel between a pair of Tx-Rx antennas), and $\widehat{H}(\tau)$ its estimate at the transmitter. The convolutive model states that $\widehat{H}(\tau)$ is obtained from $H(\tau)$ by pre- and post-convolving it with the impulse response of the RF Rx and Tx chains, respectively $D_R(\tau)$ and $D_T(\tau)$, *i.e.*

$$\widehat{H}(\tau) \triangleq D_R(\tau) * H(\tau) * D_T(\tau) \tag{5.19}$$

(the convolution takes place in the lag domain, for simplicity we consider only a LTI channel and LTI filters here). $D_T(\tau)$ and $D_R(\tau)$ are respectively $N_t \times N_t$ and $N_r \times N_r$ MIMO LTI filters. The diagonal coefficients represent the gain of the RF chain for each way, and the off-diagonal coefficients model the crosstalk between the radio ways. Ideally, $D_T(\tau) = D_T(\tau) = \delta_0(\tau)\mathbf{I}$, *i.e.* there is no crosstalk and the gains are constant over all the frequency band

and across antennas. In this case, $\widehat{H}(\tau) = H(\tau)$.

Again, w.l.o.g., let us study the frequency-flat case, and replace the convolutions by matrix products: on each frequency band, Let us assume that the channel can be described by

$$\widehat{\mathbf{H}} = \mathbf{D}_R \mathbf{H} \mathbf{D}_T, \tag{5.20}$$

where $\mathbf{H}$, $\widehat{\mathbf{H}}$, $\mathbf{D}_T$ and $\mathbf{D}_R$ are matrices of scalar values. Let us now study two models for the r.v.'s $\{\mathbf{D}_T\}$ and $\{\mathbf{D}_R\}$.

Phase uncertainty

In this model, we assume that the crosstalk is negligible, and that the gain is the same across all ways. Only the phase is random:

$$\mathbf{D}_T \triangleq \mathrm{diag}(e^{j\phi_1^T}, \ldots, e^{j\phi_{N_t}^T}), \quad \text{and} \quad \mathbf{D}_R \triangleq \mathrm{diag}(e^{j\phi_1^R}, \ldots, e^{j\phi_{N_r}^R}). \tag{5.21}$$

We choose to model the angles ϕ_i^T and ϕ_j^R as i.i.d., either uniformly distributed over $[0, 2\pi[$, or as a deviation from zero ($\{\phi_i\} \sim \mathcal{N}\left(0, \sigma_\phi^2\right)$). The first case is more likely to occur naturally, whereas the zero-centered one can be thought of as modeling an imperfect calibration device. Despite the apparent symmetry of these perturbations, their effects are dissimilar: we shall see this again by analyzing the mutual information achievable when the transmitter has the knowledge of $\widehat{\mathbf{H}}$. Again, let $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^H$ denote the SVD of the true channel matrix. Since per equation (5.21), $\mathbf{D}_T$ and $\mathbf{D}_R$ are obviously unitary, so are $\mathbf{D}_R\mathbf{U}$ and $\mathbf{D}_T^H\mathbf{V}$. Therefore, the SVD of the estimated channel, as computed by the transmitter, is $\widehat{\mathbf{H}} = (\mathbf{D}_R\mathbf{U})\mathbf{S}(\mathbf{V}^H\mathbf{D}_T)$. Therefore, eigenbeamforming yields the Tx covariance

$$\mathbf{D}(\widehat{\mathbf{H}}) = \mathbf{D}_T^H \mathbf{V} \mathbf{\Gamma} \mathbf{V}^H \mathbf{D}_T, \tag{5.22}$$

with $\mathbf{\Gamma}$ as defined in Section 5.2. The mutual information achieved by using this precoding is therefore

$$C_n = \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{D}_T^H\mathbf{V}\mathbf{\Gamma}\mathbf{V}^H\mathbf{D}_T\mathbf{H}^H\right) \tag{5.23}$$

$$= \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{S}\mathbf{V}^H\mathbf{D}_T^H\mathbf{V}\mathbf{\Gamma}\mathbf{V}^H\mathbf{D}_T\mathbf{V}\mathbf{S}^H\right) \tag{5.24}$$

Compare this value to eq. (5.6): in general, due to the introduction of $\mathbf{D}_T$, $\mathbf{V}^H\mathbf{D}_T\mathbf{V}$ and $\mathbf{V}^H\mathbf{D}_T^H\mathbf{V}$ are not identity matrices anymore. Note that $\mathbf{D}_R$ has no influence on C_n. Intuitively, this can be explained by considering that the whole eigenbeamforming process relies upon the fact that the radio waves coming from different Tx antennas will interfere to yield the received signal, and steers the beam so as to generate constructive interference. Therefore, $\mathbf{D}_T$ modifies the interference pattern, and affects the performance of the eigenbeamforming method. Conversely, $\mathbf{D}_R$ modifies the phase of the received signal *after* the radio waves superposition, and therefore the interference pattern remains unaffected.

Phase and amplitude uncertainty

A more realistic model than the phase-only perturbation introduced in the previous section is a model where the ways have slightly different gains: let $\alpha_1^T \dots \alpha_{N_t}^T$ and $\alpha_1^R \dots \alpha_{N_r}^R$ denote respectively the real, positive gains associated to each antenna and its RF path. Our model becomes

$$\mathbf{D}_T \triangleq \mathrm{diag}(\alpha_1^T e^{j\phi_1^T}, \dots, \alpha_{N_t}^T e^{j\phi_{N_t}^T}), \text{ and } \mathbf{D}_R \triangleq \mathrm{diag}(\alpha_1^R e^{j\phi_1^R}, \dots, \alpha_{N_r}^R e^{j\phi_{N_r}^R}). \tag{5.25}$$

where $\{\alpha_i^T\} \sim \mathcal{N}(1, \sigma_\alpha^2)$. A similar model including a per-antenna gain and phase uncertainty per Rx antenna was used in [67], where an algorithm to estimate $\mathbf{D}_R$ is proposed.

Again, it is expected that gain discrepancies on the Tx and Rx sides will not produce similar effects: scaling the power of the transmitted signals by arbitrary (per-antenna) coefficients will likely harm the constructive interference process. An arbitrary scaling of the received signal will be compensated by equalization at the receiver. Assuming that all the received noise comes from interfering radio signals and not from thermal noise, the SNR would remain the same.

Simulations corresponding to the two previous models are presented in Figure 5.3, for a $N_t = 5$ and $N_r = 3$ system operating at a SNR of -5dB. Alongside the already considered cases of no and perfect CSIT, four different perturbations were applied to the knowledge of $\mathbf{H}$:

1. Uniform angle perturbation: ϕ_i^T and ϕ_j^R are i.i.d. uniformly distributed over $[0, 2\pi[$. This corresponds to the curve labeled C_{n1} on Figure 5.3.

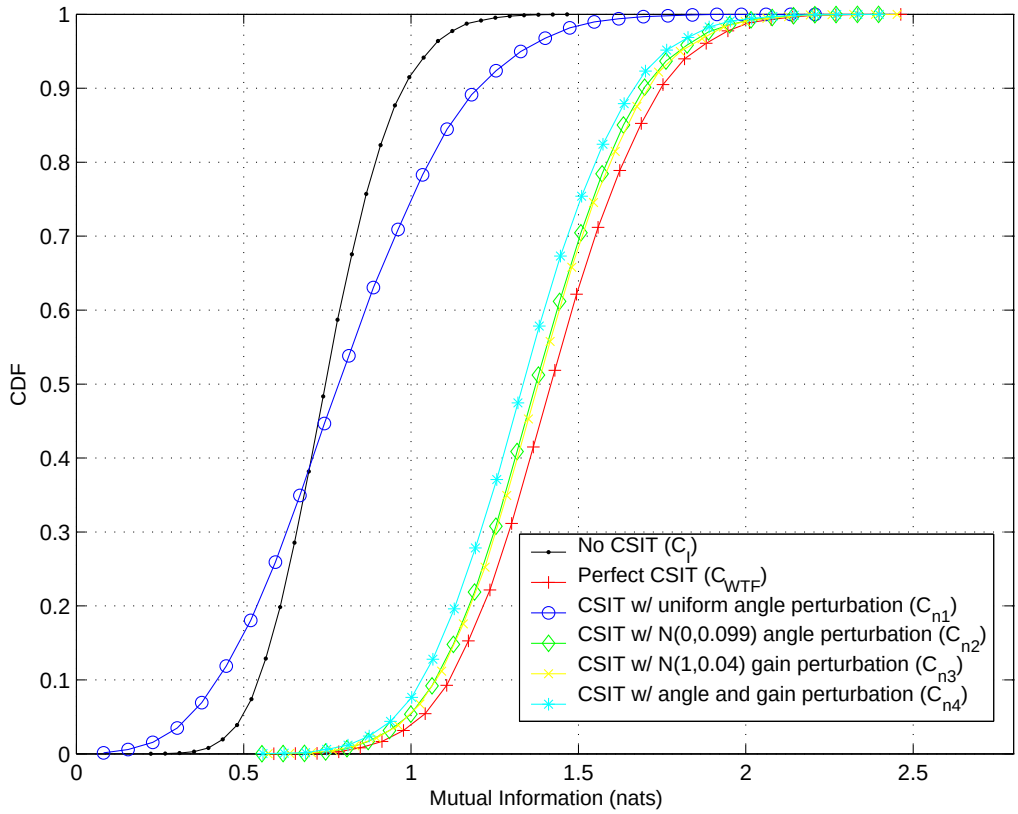


Figure 5.3: Mutual information CDF of a 5×3 channel for various multiplicative CSIT perturbations, SNR = -5dB.

2. Gaussian angle perturbation: ϕ_i^T and ϕ_j^R are i.i.d. distributed as $\mathcal{N}(0, \sigma_\phi = \pi/10)$. This corresponds to the curve labeled C_{n2}.

3. Gaussian gain perturbation: $\alpha_1^T \dots \alpha_{N_t}^T$ and $\alpha_1^R \dots \alpha_{N_r}^R$ are i.i.d. distributed as $\mathcal{N}(1, \sigma_\alpha = 0.2)$. This is denoted by C_{n3}.

4. Gaussian angle and gain perturbation: angle $\mathcal{N}(0, \sigma_\phi = \pi/10)$ and gain $\mathcal{N}(1, \sigma_\alpha = 0.2)$ perturbation. This is denoted by C_{n4}.

It is obvious from Figure 5.3 that the worst situation is the one with uniform angle perturbation (C_{n1}). In particular, the outage capacity for low outage probabilities is dramatically degraded, even w.r.t. the no-CSIT case: the 1% outage C_{n1} capacity lies at 0.21 nats, whereas the 1% outage $C_{\mathbf{I}}$ is 0.42 nats. For comparison, they are 0.90 for C_{WTF} and 0.82 nats for C_{n4}.

The reasonable angle and amplitude perturbations of cases 2, 3 and 4 do not seem to detrimental to the use of CSIT through eigenwaterfilling. We can conclude that although the use of this method in an uncalibrated system is bound to fail, this method is usable if some amount of calibration can bring the discrepancies between Tx antennas back within a reasonable range.

5.4 Conclusion

We showed that transmitter calibration is a real issue if full channel knowledge at the transmitter must be exploited (thereby mandating the use of eigenwaterfilling methods), since discrepancies between the Tx RF chains can modify the right singular subspaces of the channel. We studied the possible perturbations (in angle and in amplitude), and their effect on the maximum mutual information achievable over a given channel.

Finding bounds on the mutual information variations due to such perturbations remains an interesting open problem, and combining the rich theories of singular subspace perturbations [68] and random matrices [69] can probably yield interesting solutions to this question.

Another open issue lies with degraded CSIT. It arises in particular when a feedback link is used to transmit CSIR back to the transmitter (to be used to shape the downlink signal): this can be regarded as a source coding problem, where the channel estimate at the receiver constitutes the source, and the uplink channel constitutes a channel. Although some attempts have

been made to address this issue (see *e.g.* [70] where the instantaneous Tx covariance is fed back through its eigenvalues, or [71] where only the strongest eigendirection is used), the situation is made difficult by the fact that various partial CSIT schemes exist (ranging from simple power control to covariance feedback [61] to the full CSI feedback considered in this chapter), and that there is no uniform distortion metric on CSIT to unify them.

5.A Appendix: Proof for equation (5.8)

Let us assume that i_0 and ξ are the solutions that we seek to determine, hence they verify equations (5.3) and (5.4). Let

$$\xi_j = \frac{1}{j}\left(\gamma + N_0 \sum_{i=1}^{j} \frac{1}{s_i^2}\right), \quad \text{for} \quad j = 1 \dots K. \tag{5.26}$$

It is obvious from equation (5.4) and the definition of i_0 that $\xi = \xi_{i_0}$. We prove equation (5.8) by showing that $j \neq i_0 \Rightarrow \xi_j \geq \xi_{i_0}$.

Combining (5.3) and (5.4) yields

$$\sum_{i=1}^{i_0} \xi - \frac{N_0}{s_i^2} = \gamma \tag{5.27}$$

and

$$\sum_{i=1}^{j} \xi_j - \frac{N_0}{s_i^2} = \gamma. \tag{5.28}$$

- if $j > i_0$, subtracting (5.27) from (5.28) yields

$$i_0(\xi - \xi_j) = \sum_{i_0+1}^{j} \left(\xi_j - \frac{N_0}{s_i^2}\right) \tag{5.29}$$

  Since all the terms in the sum are negative by definition of i_0, this yields $\xi_j \geq \xi$.

- similarly, if $j < i_0$,

$$i_0(\xi_j - \xi) = \sum_{j+1}^{i_0} \left(\xi_j - \frac{N_0}{s_i^2}\right) \tag{5.30}$$

  all the terms in the sum being positive, we conclude that $\xi_j \geq \xi$.

Chapter 6

Practical exploitation of channel reciprocity

A relative calibration method for a wireless TDD link is presented, which, after a calibration phase involving feedback, lets the transmitter acquire knowledge of the downlink channel state from the uplink channel estimates, through proper modeling and estimation of the RF circuitry impulse responses. Contrarily to previous methods, relative calibration does not require specific calibration hardware. Experimental results based on bidirectional channel measurements are presented, and confirm the validity of the proposed linear reciprocity model, and of the calibration approach.

6.1 Introduction

One of the ways for the transmitter to gain knowledge of the channel state, besides using feedback as explained in Chapter 5, is to make use of reciprocity. The reciprocity principle is based on the property that electromagnetic waves traveling in both directions will undergo the same physical perturbations (*i.e.* reflection, refraction, diffraction, etc...). Therefore, in Time-Division Duplex (TDD) systems where the link operates on the same frequency band in both directions, the impulse response of the channel observed between any two antennas should be the same regardless of the direction. Application of the reciprocity principle lifts the requirement for a continuous feedback of the channel estimates while still allowing to make use of CSIT in order to optimize the transmission.

Despite these appealing characteristics, and the fact that the electromagnetic foundations of the reciprocity principle, due to H. A. Lorentz, have been known since 1896 and extensively explored (see for instance [72] and references therein), applications in the field of wireless communications have been scarce. This is due to the general understanding that the non-symmetric characteristics of the radio-frequency (RF) electronic circuitry would break the reciprocity property. Various solutions to this issue have been recently proposed. One is to calibrate each transmitter and receiver, *i.e.* to let them learn and compensate for the characteristics of their own circuitry [73]. We refer to this method as absolute calibration. This method has been is use in the radar community for a long time, since absolute calibration is necessary to determine the direction of arrival of an electromagnetic wave. It requires an external reference source with tight requirements, and is therefore expensive to implement in the context of a communications system. Another method [74] aims at ensuring the reciprocity of the electronic circuitry through a specially crafted transceiver where the same op-amp is used for both transmitting and receiving, thus lifting the requirement of calibration at the expense of design complexity.

Contrarily to these methods relying on hardware solutions, we propose a signal-space calibration (or *relative* calibration) method. It relies on a calibration phase to establish the relationship between the channel as measured in both directions. We deem this a relative calibration since it takes place entirely in signal space, and no external reference source, nor any other hard-

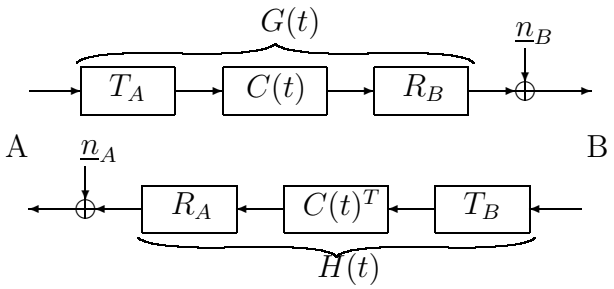


Figure 6.1: Reciprocity model for a MIMO FDD frequency-selective channel

ware, are necessary. The calibration phase requires feedback. Once it has been performed, we show that it is possible to infer the state of the downlink channel (A to B) from the (estimated) state of the uplink channel (B to A).

In this chapter, we first introduce a linear reciprocity model, and the concept of relative calibration. In Section 6.4.1, we show how to apply relative calibration to the particular case of a narrowband flat-fading channel. In Section 6.4, the particular case of time-dispersive Single-Input Single-Output (SISO) channel calibration is studied. These results are applied in Section 6.5 to measurements gathered using a prototype Universal Mobile Telecommunications Services (UMTS) TDD link. The validity of the proposed reciprocity model is confirmed by the measurements.

Results presented in this chapter were published in [14]. The relative calibration process is the subject of patent [15].

6.2 System model

We consider a bidirectional point-to-point TDD MIMO link, between two stations denoted by A and B, using respectively M and N antennas. We model the channel as seen during the baseband processing as the cascade of three linear filters, and some additive white Gaussian noise (AWGN), as represented in Fig. 6.1.

The upper part of the diagram represents a transmission from A to B, whereas the lower part represents a transmission from B to A. T_A denotes the M-input M-output equivalent filtering operation of the transmit circuitry of

A, $C(t)$ is a $N \times M$ matrix containing the impulse responses (one per Tx-Rx antenna pair) of the electromagnetic channel at time t, and R_B is the N-input N-output equivalent filter modeling the receive circuitry of B. Symmetrically, T_B and R_A denote the equivalent filters corresponding respectively to the transmit circuitry of B and the receive circuitry of A. Note that the characteristics of the circuitry do not depend on t, since their variation is usually much slower than the channel variation.

In a TDD setting, assuming that the transmissions in both directions take place within a time frame shorter than the channel coherence time, the reciprocity of the electromagnetic channel guarantees that the impulse responses between each antenna pair is the same in both directions, therefore the filter $C(t, \tau)$ is common to both directions (it needs to be transposed to respect the order of the antennas). Let us denote by

$$G(t, \tau) = R_B(\tau) * C(t, \tau) * T_A(\tau) \tag{6.1}$$

the compound impulse response (τ is the index in the lag domain) of the (noiseless) channel from A to B, and by

$$H(t, \tau) = R_A(\tau) * C(t, \tau)^T * T_B(\tau) \tag{6.2}$$

its counterpart when the transmission takes place from B to A. For the sake of simplicity, the noise $\underline{n}_A$, $\underline{n}_B$ is supposed to be injected after the cascade of the three filters, although it really appears between the electromagnetic channel and the receive circuitry (*i.e.* between the second and third linear filter in Fig. 6.1). Knowledge of $G(t, \tau)$ is easily available to station B, using any classical channel estimation method, and similarly station A can estimate $H(t, \tau)$.

Note that $T_A(f)$, $T_B(f)$, $R_A(f)$ and $R_B(f)$ are all square matrices, and we will work under the assumption that they have no singularities for any f in the considered frequency band. This should be a reasonable requirement, since the design target for the circuitry is usually to have unit diagonal gains over the desired band, and as little crosstalk as possible. A strictly diagonal structure can be assumed if little or no crosstalk is present between antenna channels in the circuitry, thus further simplifying the model.

6.3 Relative calibration

As stated before, although it is common practice for station B to estimate G in order to perform coherent detection of the received signal, the knowledge of G is desirable to A, since it would enable the use of CSIT-exploiting methods. Instead of relying on continuous feedback of CSI from B to A, we link the channel estimates in both directions by eliminating C in eqs. (6.1) and (6.2). In the frequency domain, this yields

$$G(t,f) = R_B(f)T_B(f)^{-T}H(t,f)^T R_A(f)^{-T}T_A(f). \tag{6.3}$$

Note that T_A, T_B, R_A and R_B are generally not known individually, since this would constitute absolute calibration on both sides of the link. Nevertheless, it can be seen from eq. (6.3) that only $P_A(f) = R_A(f)^{-T}T_A(f)$ and $P_B(f) = R_B(f)T_B(f)^{-T}$ are necessary in order to infer $G(t)$ from $H(t)$:

$$G(t,f) = P_B(f)H(t,f)^T P_A(f). \tag{6.4}$$

Estimating $P_A(f)$ and $P_B(f)$ constitutes a relative calibration between A and B, and can be realized entirely through the use of classical channel estimation and feedback techniques, as will be shown in the sequel.

6.4 Reciprocity parameters estimation

In this section, we present a method to estimate the reciprocity parameters $P_A(\tau)$ and $P_B(\tau)$ from one or several pairs of (simultaneous, uplink and downlink) channel measurements.

6.4.1 Narrowband MIMO Flat-Fading Case

Let us first address the relative calibration problem in the case of a narrowband transmission over a flat-fading channel. The flat-fading assumption guarantees that $C(t,f)$ is constant (independent of f) over the considered frequency band, whereas by narrowband we imply that the behaviour of the electronic circuitry can also be assumed constant over the same band. Therefore, let us represent the channel from A to B at time t_k by the $N \times M$ matrix $\mathbf{G}_k = G(t_k, 0)$, and similarly in the reverse direction $\mathbf{H}_k = H(t_k, 0)$. $P_A(f)$

and $P_B(f)$ respectively become $\mathbf{P}_A$ and $\mathbf{P}_B$. Using these notations, eq. (6.4) becomes a simple matrix product

$$\mathbf{G}_k = \mathbf{P}_B \mathbf{H}_k^T \mathbf{P}_A. \tag{6.5}$$

We will show how to estimate $\mathbf{P}_A$ and $\mathbf{P}_B$ (thereby performing the desired relative calibration), or equivalently for $\mathbf{P}'_A = \mathbf{P}_A^{-1}$ and $\mathbf{P}_B$ given one or several estimates $\widehat{\mathbf{G}}_k = \mathbf{G}_k + \widetilde{\mathbf{G}}_k$ and $\widehat{\mathbf{H}}_k = \mathbf{H}_k + \widetilde{\mathbf{H}}_k$, $k = 1 \dots K$, of the channel in both directions ($\widetilde{\mathbf{G}}_k$ and $\widetilde{\mathbf{H}}_k$ represent the estimation noise). These estimates can be obtained through a classical channel feedback operation, during the calibration phase. The only requirement is that $\widehat{\mathbf{G}}_k$ and $\widehat{\mathbf{H}}_k$ be estimated approximately at the same instant, in order for the reciprocity relationship to hold.

Let us first consider one single channel realization (in both directions) $(\mathbf{G}_k, \mathbf{H}_k)$ corresponding to instant k. Let $\underline{\mathbf{h}}_k^{(1)} \dots \underline{\mathbf{h}}_k^{(M)}$ denote the rows of $\mathbf{H}_k$, and define

$$\underline{\mathbf{p}}'_A \triangleq \text{vect}\left(\mathbf{P}'_A\right), \quad \underline{\mathbf{p}}_B \triangleq \text{vect}(\mathbf{P}_B^T), \tag{6.6}$$

as well as

$$h\left(\mathbf{H}_k\right) \triangleq \begin{pmatrix} \mathbf{I}_N \otimes \underline{\mathbf{h}}_k^{(1)} \\ \vdots \\ \mathbf{I}_N \otimes \underline{\mathbf{h}}_k^{(M)} \end{pmatrix} \quad \text{and} \quad g\left(\mathbf{G}_k\right) \triangleq \mathbf{I}_M \otimes \mathbf{G}_k, \tag{6.7}$$

where $h\left(\mathbf{H}_k\right)$ is a $MN \times N^2$ matrix, and $g\left(\mathbf{G}_k\right)$ is a $MN \times M^2$ block-diagonal matrix. With these notations, eq. (6.5) is equivalent to

$$g\left(\mathbf{G}_k\right) \underline{\mathbf{p}}'_{\mathbf{A}} - h\left(\mathbf{H}_k\right) \underline{\mathbf{p}}_{\mathbf{B}} = \underline{0}, \tag{6.8}$$

Let us now define $\mathbf{E}_k \triangleq \left[g\left(\mathbf{G}_k\right) \ -h\left(\mathbf{H}_k\right)\right]$ and $\underline{\mathbf{p}}_0 \triangleq \left[\underline{\mathbf{p}}'^{\,T}_{\mathbf{A}} \underline{\mathbf{p}}^{\,T}_{\mathbf{B}}\right]^T$. According to eq. (6.8), $\underline{\mathbf{p}}_0$ is a non-zero solution of

$$\mathbf{E}_k \underline{\mathbf{p}} = \underline{0}_{MN \times 1}. \tag{6.9}$$

The solutions to (6.9) lie in the right null subspace of $\mathbf{E}_k$. Since we are trying to determine $M^2 + N^2$ unknowns, the MN linear constraints implied by eq. (6.9) are not sufficient to solve the problem. Fortunately, this uncertainty can be lifted if multiple channel measurements are available: since eq. (6.9)

is true for any k, $\underline{\mathbf{p}}_0$ is also in the right null subspace of $\mathbf{E} \triangleq \left[\mathbf{E}_1^T \dots \mathbf{E}_K^T\right]^T$. If the successive channel realizations are linearly independent (this can be achieved by waiting for the channel to evolve between successive estimations), $\underline{\mathbf{p}}_0$ is fully determined (up to a scalar coefficient) when the right null subspace of $\mathbf{E}$ is of dimension 1. The scaling coefficient disappears in eq. (6.5) since it scales $\mathbf{P}_A = {\mathbf{P}'_A}^{-1}$ and $\mathbf{P}_B$ in reciprocal ways.

Since the channel realizations are not known perfectly, but rather through the estimates $(\widehat{\mathbf{G}}_k, \widehat{\mathbf{H}}_k), k = 1 \dots K$, this is a particular case of the Total Least-Squares problem [75]. Letting

$$\widehat{\mathbf{E}} \triangleq \begin{bmatrix} g(\widehat{\mathbf{G}}_1) & -h(\widehat{\mathbf{H}}_1) \\ \vdots & \vdots \\ g(\widehat{\mathbf{G}}_K) & -h(\widehat{\mathbf{H}}_K) \end{bmatrix}, \tag{6.10}$$

and assuming that the noise coefficients in $\widetilde{\mathbf{G}}_k$ and $\widetilde{\mathbf{H}}_k$ are i.i.d. and all have the same variance, the total least squares solution $\widehat{\underline{\mathbf{p}}}_0$ is given (again, up to an arbitrary scalar) by the right singular vector associated to the smallest singular value of $\widehat{\mathbf{E}}$.

Sparse reciprocity matrices

It is likely that the reciprocity matrices encountered in most real systems will be close to diagonal, or that non-zero off-diagonal coefficients resulting *e.g.* from antenna coupling will affect only neighboring antennas (this would yield band matrices). If $\mathbf{P_A}$ and $\mathbf{P_B}$ are experimentally verified to be sparse matrices, only the relevant (non-zero) coefficients need be estimated. Thus the estimation process needs to be adapted by removing coefficients known to be zero from $\underline{\mathbf{p}}'_A$ and $\underline{\mathbf{p}}_B$, as well as the corresponding columns of $g(\mathbf{G}_k)$ and $h(\mathbf{H}_k)$.

As an illustration, let us assume that $\mathbf{P}_A$ and $\mathbf{P}_B$ are diagonal matrices, and denote $\underline{\mathbf{d}}'_A$ and $\underline{\mathbf{d}}_B$ the diagonal coefficients of $\mathbf{P}'_A = \mathbf{P}_A^{-1}$ and $\mathbf{P}_B$ respectively. Furthermore, let $\underline{\mathbf{g}}_k^{(1)} \dots \underline{\mathbf{g}}_k^{(N)}$ denote the rows of $\mathbf{G}_k$, and $|\mathbf{h}_k^{(1)} \dots |\mathbf{h}_k^{(N)}$

the columns of $\mathbf{H}_k$. Then, letting

$$g'(\mathbf{G}_k) \triangleq \begin{bmatrix} \mathrm{diag}(\underline{\mathbf{g}}_k^{(1)}) \\ \vdots \\ \mathrm{diag}(\underline{\mathbf{g}}_k^{(N)}) \end{bmatrix} \quad \text{and} \quad h'(\mathbf{H}_k) \triangleq \begin{bmatrix} |\mathbf{h}_k^{(1)} & & 0 \\ & \ddots & \\ 0 & & |\mathbf{h}_k^{(N)} \end{bmatrix}, \quad (6.11)$$

and $\underline{\mathbf{p}}' \triangleq \left[\underline{\mathbf{d}}_A'^{\,T}\, \underline{\mathbf{d}}_B{}^T\right]^T$, equation (6.9) becomes

$$\mathbf{E}_\mathbf{k}'\underline{\mathbf{p}}' = [g'(\mathbf{G_k}) \;\; -h'(\mathbf{H_k})]\,\underline{\mathbf{p}}' = \underline{0}_{MN\times 1}. \quad (6.12)$$

The solution can be found using the method outlined in the previous section, with the added benefit that since the number of unknowns is reduced ($\underline{\mathbf{p}}'$ has only $M+N$ coefficients, as opposed to M^2+N^2 unknowns in the general case), the amount of independent channel realizations necessary to narrow down the solution subspace to dimension 1 is reduced, possibly down to $K=1$, which would enable the estimation of the reciprocity coefficients based on a single instantaneous channel measurement.

The simulated performance of this method for the case of diagonal $\mathbf{P}_A$ and $\mathbf{P}_B$ is presented in Figure 6.2. The complex diagonal coefficients of the reciprocity matrices are i.i.d, with phases uniformly distributed over $[0, 2\pi]$, and $\mathcal{N}(1, 1/25)$ radii. The average estimation error per component is plotted versus the noise variance of the channel estimates. For the simulations using multiple channel measurements, $K=5$ independent channel and noise realizations were used.

6.4.2 SISO frequency-selective case

In this particular case, the product in eq. (6.4) commutes because all factors are 1×1 matrices. Therefore, letting

$$P(\tau) \triangleq P_B(\tau) * P_A(\tau), \quad (6.13)$$

we can rewrite (6.4) as (in the time domain, using convolutions)

$$G(t,\tau) = H(t,\tau) * P(\tau). \quad (6.14)$$

Let us consider K pairs of measurements of the discretized complex channel impulse responses in both directions, $\underline{\mathbf{g}}_k \triangleq \left(g_1^{(k)}, \ldots, g_L^{(k)}\right)^T \in \mathbb{C}^L$,

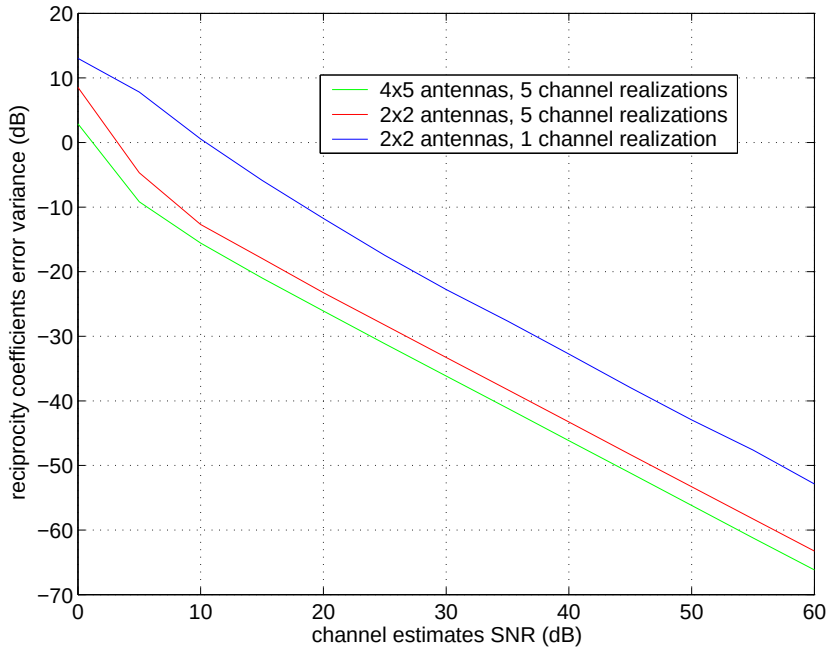


Figure 6.2: Estimation error variance for diagnoal $\mathbf{P}_A$ and $\mathbf{P}_B$ (narrowband, flat fading case).

$\underline{\mathbf{h}}_k \triangleq \left(h_1^{(k)}, \ldots, h_{L'}^{(k)}\right)^T \in \mathbb{C}^{L'}$. Assuming that both the channel impulse responses and the filters P_A and P_B have finite length, the convolution in the reciprocity condition (6.14) can be written as

$$\mathbf{H}_k \underline{\mathbf{p}} = \underline{\mathbf{g}}_k, \tag{6.15}$$

with $\underline{\mathbf{p}} \triangleq (p_1, \ldots, p_{L'-L+1})^T$ and

$$\mathbf{H}_k \triangleq \begin{bmatrix} h_{L'-L+1}^{(k)} & h_{L'-L}^{(k)} & \ldots & h_1^{(k)} \\ h_{L'-L+2}^{(k)} & h_{L'-L+1}^{(k)} & \ldots & h_2^{(k)} \\ \vdots & & \ddots & \vdots \\ h_{L'}^{(k)} & \ldots & & h_{L'-L+1}^{(k)} \end{bmatrix}. \tag{6.16}$$

Note that the length of the impulse response of the filter, $L' - L + 1$, must be a sensible value. It is possible to solve for $\underline{\mathbf{p}}$ in eq. (6.15), through *e.g.* least-squares if the system is overdetermined, however this method assumes

that only $\underline{\mathbf{g}}_k$ is noisy, and that $\underline{\mathbf{h}}_k$ is known perfectly. Since in practice, only the noisy versions $\widehat{\underline{\mathbf{g}}}_k$ and $\widehat{\underline{\mathbf{h}}}_k$ are known, we look for $\underline{\mathbf{p}}$ as the solution of the optimization problem

$$\min_{\mathbf{E}_k, \underline{\mathbf{x}}} \underline{\alpha}_k^H \underline{\alpha}_k + \underline{\beta}_k^H \underline{\beta}_k, \quad \text{s.t.} \left(\widehat{\mathbf{H}}_k + \mathbf{E}_k \right) \underline{\mathbf{x}} = \widehat{\underline{\mathbf{g}}}_k + \underline{\beta}_k \tag{6.17}$$

where $\mathbf{E}_k$ is Toeplitz and contains the coefficients of $\underline{\alpha}_k$:

$$\mathbf{E}_k \triangleq \begin{bmatrix} \alpha_{L'-L+1}^{(k)} & \alpha_{L'-L}^{(k)} & \cdots & \alpha_1^{(k)} \\ \alpha_{L'-L+2}^{(k)} & \alpha_{L'-L+1}^{(k)} & \cdots & \alpha_2^{(k)} \\ \vdots & & \ddots & \vdots \\ \alpha_{L'}^{(k)} & \cdots & & \alpha_{L'-L+1}^{(k)} \end{bmatrix}. \tag{6.18}$$

The vectors $\underline{\alpha}_k \in \mathbb{C}^{L'}$ and $\underline{\beta}_k \in \mathbb{C}^{L}$ represent the corrections of the noise present on $\underline{\mathbf{h}}_k$ and $\underline{\mathbf{g}}_k$ respectively. Under the assumption that the noise is Gaussian i.i.d., the ML solution of (6.17) can again be obtained numerically: although the simple SVD solution used in Section 6.4.1 does not apply since the coefficients in the (Toeplitz) noise matrix $\mathbf{E}_k$ are not i.i.d., it can be solved numerically, as recognized by Mastronardi in [76], since this formulation defines a Structured Total Least-Squares (STLS) problem.

In order to guarantee the identifiability of $\underline{\mathbf{p}}$, and since the measurements are noisy, it is preferable to over-determine the problem. To this aim, the linear system (6.15) can be extended by concatenating the successive channel measurements, since $\underline{\mathbf{p}}$ is assumed to remain constant over all measurements. Therefore,

$$\begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_K \end{bmatrix} \underline{\mathbf{p}} = \begin{bmatrix} \underline{\mathbf{g}}_1 \\ \vdots \\ \underline{\mathbf{g}}_K \end{bmatrix}. \tag{6.19}$$

The STLS algorithm can be straightforwardly extended to estimate $\underline{\mathbf{p}}$ using all the measurements, by solving the optimization problem (with $\underline{\alpha} \triangleq [\underline{\alpha}_1^T \ldots \underline{\alpha}_K^T]^T$)

$$\min_{\mathbf{E}_1, \ldots, \mathbf{E}_K, \underline{\mathbf{x}}} \underline{\alpha}^H \underline{\alpha} + \underline{\beta}^H \underline{\beta}, \tag{6.20}$$

$$\text{s.t.} \quad \begin{bmatrix} \widehat{\mathbf{H}}_1 + \mathbf{E}_1 \\ \vdots \\ \widehat{\mathbf{H}}_K + \mathbf{E}_K \end{bmatrix} \underline{\mathbf{x}} = \begin{bmatrix} \widehat{\underline{\mathbf{g}}}_1 \\ \vdots \\ \widehat{\underline{\mathbf{g}}}_K \end{bmatrix} + \underline{\beta}. \tag{6.21}$$

6.4.3 SIMO and MISO frequency-selective cases

Let us now consider the situation where only one side of the link is equipped with multiple antennas. This applies to both SIMO (Single-Input Multiple-Output, *i.e* $M = 1$, $N > 1$) and MISO (Multiple-Input Single-Output, *i.e.* $M > 1$, $N = 1$) channels.

In both cases, either $P_A(\tau)$ or $P_B(\tau)$ is a 1×1 filter, and therefore the commutation property can still be used to transform eq. (6.4) into

$$G(t,\tau) = P_{SIMO}(\tau) * H(t,\tau)^T \tag{6.22}$$

for the SIMO case, where $P_{SIMO}(\tau) \triangleq P_B(\tau) * (\mathbf{I}_N \otimes P_A(\tau))$, or

$$G(t,\tau) = H(t,\tau)^T * P_{MISO}(\tau) \tag{6.23}$$

with $P_{MISO}(\tau) \triangleq (\mathbf{I}_M \otimes P_B(\tau)) * P_A(\tau)$ in the MISO case. In each case, the reciprocity parameters are grouped into one single linear filter (P_{SIMO} or P_{MISO}), which again can be efficiently (ML-)estimated using the STLS method.

Furthermore, if no crosstalk is present on the side equipped with the multiple antennas, the problem merely degenerates into several parallel SISO channels whose reciprocity parameters can be estimated independently. This is evidenced by the fact that, for the respective SIMO and MISO cases, if P_B (resp. P_A) is diagonal, P_{SIMO} (resp. P_{MISO}) are also diagonal filters.

6.4.4 MIMO frequency-selective case

Estimation of the reciprocity parameters in the MIMO case ($M > 1$ and $N > 1$) is less straightforward, since in this case eq. (6.4) is not jointly linear in the unknowns (P_A,P_B), and can not be made linear by commutation of the filters.

First proposed method: overparameterization. This method is applicable only in the case where P_A and P_B are both diagonal. It represents the reciprocity parameters using more filters that necessary (*i.e.* one filter per antenna pair instead of one filter per antenna).

In order to do this, first note that eq. (6.4) is equivalent to the MN equations obtained from

$$[G(t,\tau)]_{i,j} = [P_B(\tau)]_{i,i} * [H(t,\tau)]_{j,i} * [P_A(\tau)]_{j,j} \tag{6.24}$$

for $i = 1 \dots N, \quad j = 1 \dots M$ (recall that $[\cdot]_{i,j}$ denotes the $(i,j)^{th}$ element of a matrix, in this case a linear filter). Since $[P_A(\tau)]_{j,j}$ and $[P_B(\tau)]_{i,i}$ represent SISO filtering operations, letting

$$P_{(i,j)} \triangleq [P_B(\tau)]_{i,i} * [P_A(\tau)]_{j,j} \tag{6.25}$$

yields

$$[G(t,\tau)]_{i,j} = [H(t,\tau)]_{j,i} * P_{(i,j)}, \tag{6.26}$$

where $P_{(i,j)}$ can again be estimated using the technique developed for the SISO case (compare (6.26) to (6.14)). However, this method fails to take into account the fact that the MN linear filters $P_{(i,j)}$ are generated from only $M+N$ impulse responses $[P_A(\tau)]_{j,j}$, $j = 1 \dots M$, and $[P_B(\tau)]_{i,i}$, $i = 1 \dots N$.

Second proposed approach: alternating estimation. This is an iterative method whereby $P_A(\tau)$ and $P_B(\tau)$ are alternatively assumed to be known perfectly, while the other is estimated under this assumption (the STLS algorithm can be used again, since assuming that $P_A(\tau)$ is perfectly known makes the problem linear in $P_B(\tau)$, and vice-versa). The algorithm is initialized by assuming *e.g.* that $P_A(\tau) = \mathbf{I}_M$ at the first iteration. The proposed algorithm is therefore the following:

1. initialization: $\widehat{P}_A(\tau) = \mathbf{I}_M$.
2. assuming $\widehat{P}_A(\tau)$ perfectly known, compute $\widehat{P}_B(\tau)$ as the STLS solution of a linear system identification problem expressed by equation (6.4).
3. assuming $\widehat{P}_B(\tau)$ perfectly known, compute $\widehat{P}_A(\tau)$, by applying STLS to eq. (6.4) again (note that the same equation describes different problems since the role of the unknowns and the constants have changed).
4. iterate (go to step 2) until convergence.

This approach is applicable even for non-diagonal reciprocity parameters, and does not over-parameterize the system. Unfortunately, although this method has been observed to work satisfactorily, its convergence has not been proved, and no optimality claim can be made about the results.

6.5 Experimental investigation of the reciprocity principle

In order to validate the linear system model of Section 6.2, and to assess the validity of the calibration process of Section 6.3, channel measurements were performed in the case of a SISO system. The measurements were performed in the framework of an actual UMTS connection, using Eurecom's in-house prototype UMTS platform [77]. The UMTS TDD link operates on a 3.84MHz wide channel in the 1900-1920MHz band. A rooftop antenna fitted with a power amplifier and a low-noise amplifier is connected via low-loss cables to the RF modules of the base station. Another RF module of similar design, directly connected to a portable antenna, was used inside the building as the mobile terminal (MT) side of the link.

Channel estimates were obtained on both sides using conventional channel estimation techniques, by exploiting the training sequences embedded in the UMTS traffic. During the calibration phase, the feedback link was assumed to be of infinite bandwidth, *i.e.* the estimates performed on both sides were made available to one single place for computation with no further degradation. In a practical calibration process, since this information has to be conveyed over a finite-capacity channel, some approximation errors would appear, and should be accounted for. Perfect feedback eases this requirement.

6.5.1 SISO reciprocity characteristics estimation

The measurements consist in a series of estimates $\widehat{G}(t,\tau) = G(t,\tau) + \widetilde{G}(t,\tau)$ and $\widehat{H}(t,\tau) = H(t,\tau) + \widetilde{H}(t,\tau)$ of the channel impulse responses. Assuming that $\widetilde{G}(t,\tau)$ and $\widetilde{H}(t,\tau)$ are both Gaussian i.i.d. random processes of equal variances, we used the STLS deconvolution algorithm presented in Section 6.4. One channel estimate for each direction was obtained every 10ms. The oscillators on both sides of the wireless link were synchronized through a wired link. Successive estimates of $\widehat{P}$ were performed, each estimate being based on 50 successive channel measurements (*i.e.* each estimate is made over a 500ms time span). Two series of measurements were performed:

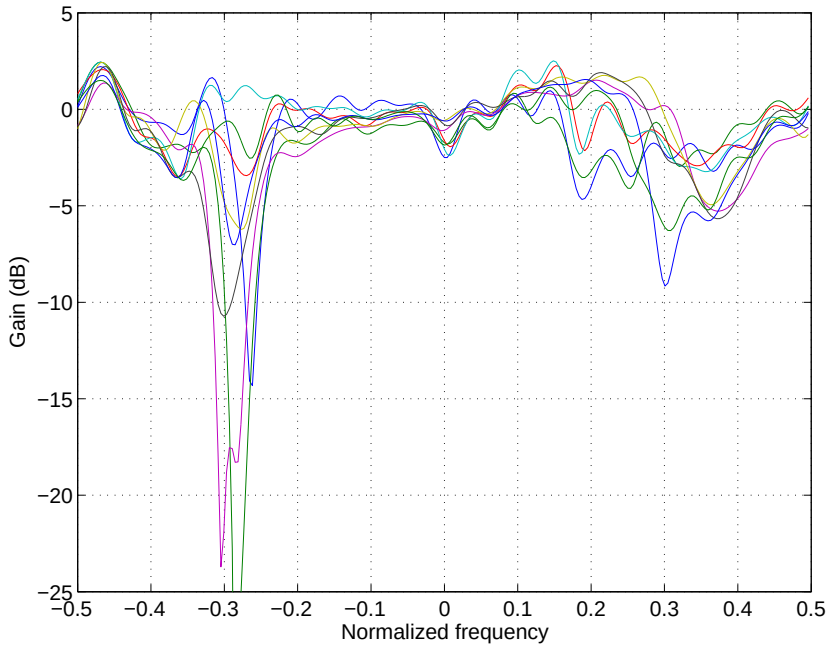
Fixed setting the MT antenna lies on a table, and human operators in the vicinity of the experiment make as little movement as possible.

Moving setting the MT antenna was hand-held and moved rapidly by a human operator.

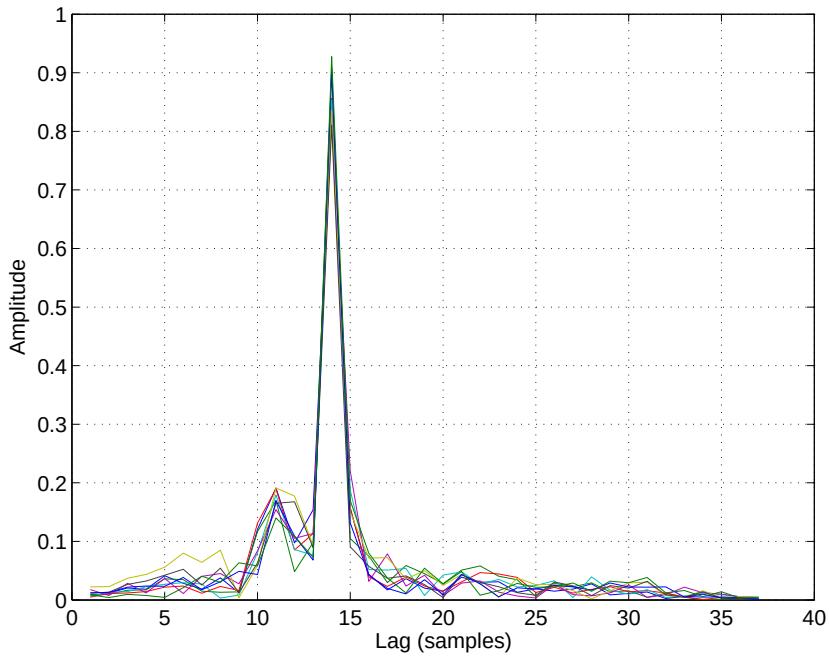
Fig. 6.3(a) and 6.3(b) show respectively the frequency-domain representation and the corresponding time-domain representation of the estimated reciprocity function $\widehat{P}$ over the considered 3.84MHz-wide frequency band. Every figure shows 9 curves, corresponding to successive estimates of $\widehat{P}$. The relative stability of the estimates of $P(f)$ in the series of 9 consecutive measurements indicates that the assumption that this is a slow-varying parameter is valid, and is therefore an encouraging sign that the proposed relative calibration process is possible. The first noticeable characteristic of these figures is their general non-flatness of the spectrum 6.3(a). A perfectly reciprocal setting would yield a Kronecker delta function, and the corresponding flat spectrum. The spectrum 6.3(a) indicates a high level of non-reciprocity in the absence of calibration, and thus the need to compensate for this effect. The impulse response 6.3(b) clearly shows two distinct peaks, instead of the ideal delta function. The deep fade in the spectrum around frequency -0.3 corresponds to a frequency band where both $G(t, f)$ and $H(t, f)$ have fades. Therefore, the determination of $P(f) \triangleq \frac{G(t,f)}{H(t,f)}$ is dominated by the noise.

Fig. 6.4(a) and 6.4(b) were obtained from the measurements in the mobile setting. The spectrum is definitely not flat, with a noticeable attenuation in the lower frequencies. The ripples, *e.g.* in the [.3 .4] frequency range, can not be attributed to noise or measurement artifacts, and clearly represent a deviation from ideal reciprocity.

Comparing Figs. 6.3(a) and 6.4(a), *i.e.* the estimates of $\widehat{P}$ for the fixed and mobile case, and bearing in mind that both experiments used the same hardware and were separated by only a few minutes, the discrepancy in the reciprocity functions is obvious. It can presumably be attributed to changes in the MT antenna coupling, due to the presence of the operator's hand near the antenna in the second case (moving setting). This is an important observation, since it means that more frequent calibration cycles might be required when human activity takes place very close to the antennas, for instance when the link involves a mobile telephone handset. Note that, contrarily to absolute calibration, both sides have to recalibrate if the characteristics are changing on either side of the link.

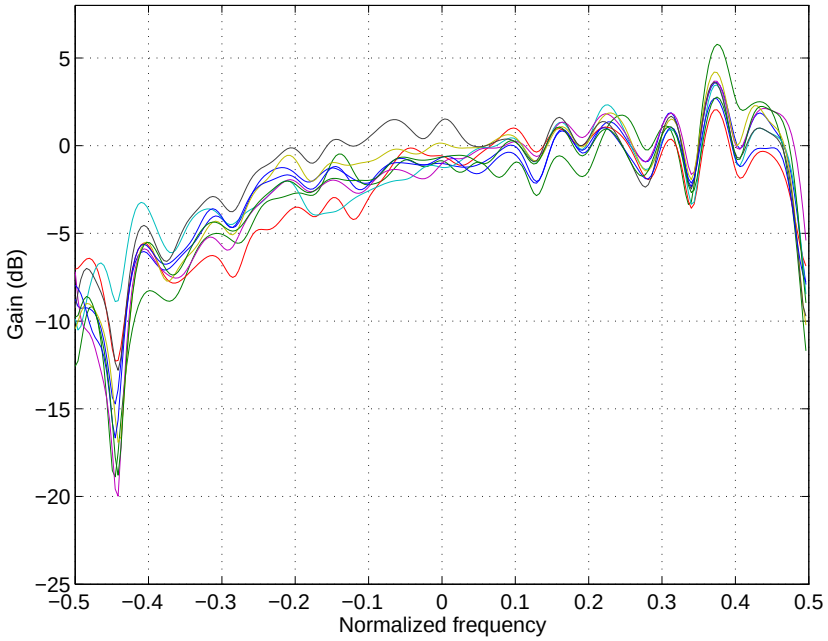


(a) Frequency-domain

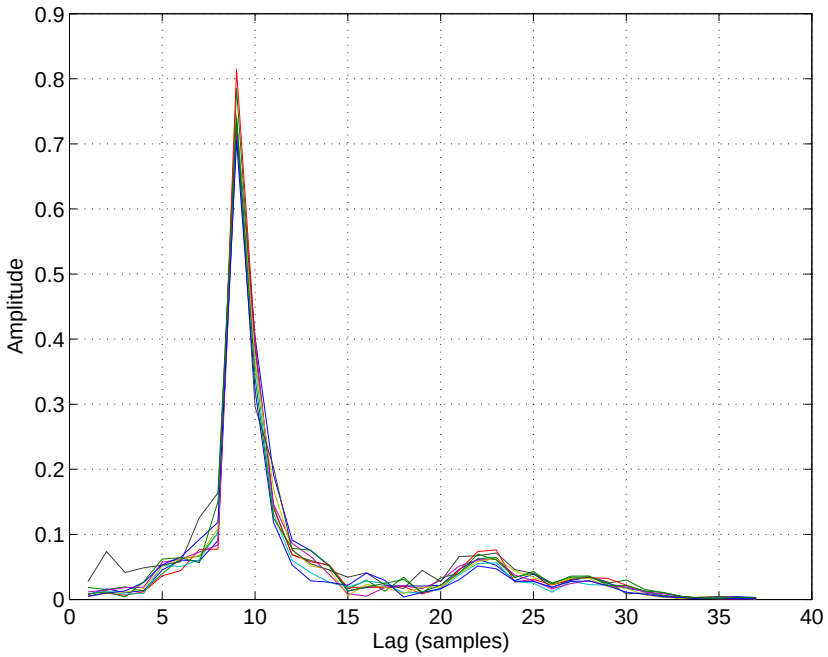


(b) Time-domain

Figure 6.3: Nine successive estimates of the reciprocity function, fixed setting



(a) Frequency-domain



(b) Time-domain

Figure 6.4: Nine successive estimates of the reciprocity function, moving setting

6.5.2 Performance of reciprocity-based channel estimation

Although it is interesting *per se* from an engineering point of view, the main interest of $\widehat{P}$ is its use for channel estimation, as described in Section 6.3. Therefore, in order to evaluate the level of performance attainable through relative calibration and reciprocal channel estimation, we reprocessed the same series of measurements was reprocessed in order to evaluate the accuracy of channel estimates.

In this scenario, after estimating $\widehat{P}$ only once over 50 2-tuples $(\widehat{G}(t), \widehat{H}(t))$, $\widehat{G}_{est}(t)$ was computed as

$$G_{est}(t,\tau) = \widehat{H}(t,\tau) * \widehat{P}(\tau) \tag{6.27}$$

for the subsequent 422 measurements of $\widehat{H}(t,\tau)$ (*i.e.* over a 4.22s time span). Since the noiseless value $G(t,\tau)$ was not available, $G_{est}(t)$ was compared to $\widehat{G}(t)$, using the noise amplification metric

$$\alpha \triangleq \frac{\mathrm{E}_t\left[\left|\left|G_{est}(t) - \widehat{G}(t)\right|\right|^2\right]}{\sigma^2_{\tilde{G}}}. \tag{6.28}$$

where the norm chosen here is the total energy in the impulse response, *i.e.*

$$||f|| \triangleq \int_{-\infty}^{+\infty} |f(\tau)|^2 \,\mathrm{d}\tau. \tag{6.29}$$

Obviously, even if the estimator is perfect, *i.e.* yields $G_{est} = G$, some noise is still present in $\widehat{G}$, and therefore α is lower-bounded by 1. For comparison, in the ideal reciprocity case, and with Gaussian i.i.d. noise, the metric (6.28)

becomes

$$\mathrm{E}_t\left[\left|\left|G_{est}(t)-\widehat{G}(t)\right|\right|^2\right] \tag{6.30}$$

$$= \mathrm{E}_t\left[\left|\left|H(t)*\widehat{P}+\widetilde{H}(t)*\widehat{P}-\left(H(t)*P+\widetilde{G}(t)\right)\right|\right|^2\right] \tag{6.31}$$

$$= \mathrm{E}_t\left[\left|\left|H(t)*\left(\widehat{P}-P\right)\right|\right|^2\right]+\mathrm{E}_t\left[\left|\left|\widetilde{H}(t)*\widehat{P}\right|\right|^2\right] +\mathrm{E}_t\left[\left|\left|\widetilde{G}(t)\right|\right|^2\right] \tag{6.32}$$

$$\geq \mathrm{E}_t\left[\left|\left|\widetilde{H}(t)*\widehat{P}\right|\right|^2\right]+\mathrm{E}_t\left[\left|\left|\widetilde{G}(t)\right|\right|^2\right] \tag{6.33}$$

where we used the independence of H, $\widetilde{G}$ and $\widetilde{H}$ in (6.32), and the bound in (6.33) is tight if the estimation of the reciprocity parameter is perfect ($\widehat{P}=P$). Furthermore, since both G and H are normalized to unit average energy ($\mathrm{E}_t\left[||G(t)||^2\right]=\mathrm{E}_t\left[||H(t)||^2\right]=1$), $||P||^2=1$ also, and therefore $||\widetilde{H}(t)*\widehat{P}||^2 \approx ||\widetilde{H}(t)||^2=\sigma^2_{\widetilde{H}}$. Plugging this expression back into (6.28) yields

$$\alpha \geq \frac{\sigma^2_{\widetilde{H}}+\sigma^2_{\widetilde{G}}}{\sigma^2_{\widetilde{G}}}=1+\frac{\sigma^2_{\widetilde{H}}}{\sigma^2_{\widetilde{G}}}=1+\frac{\mathsf{SNR}_G}{\mathsf{SNR}_H}. \tag{6.34}$$

This bound is tight if the estimators are perfect.

Table 6.1 shows the noise amplification metric α measured (the expectation is replaced by time-averaging over the measurement window) for both series of measurements (fixed and mobile setting), as well as the lower bound $1+\frac{\mathsf{SNR}_G}{\mathsf{SNR}_H}$, computed using the average measured SNRs.

For comparison, a third column presents simulation results obtained for a perfect reciprocity setting. The stochastic parameters in the simulation were set as close as possible to the experimental results: the downlink power density profile and the uplink and downlink SNRs were those from the measurements. The uplink channel was generated from the downlink channel by assuming perfect reciprocity. Since no statistical information is available on the reciprocity function P, it was set to be an FIR filter generated from its five random zeros, uniformly distributed in the zero-centered disc of radius .2. The rest of the simulation followed the method used for the measurements,

	α (measured)	$1 + \frac{\mathsf{SNR}_G}{\mathsf{SNR}_H}$	α (simulated)
Fixed setting	2.081	1.962	1.931
Mobile setting	3.692	1.728	1.727

Table 6.1: Noise figure α for the reciprocity channel estimation method

up to the computation of the noise metric α. The results presented in the third column of Table 6.1 is a Monte-Carlo average over the realizations of P.

It can be seen that the simulated results for α are slightly below the $1 + \frac{\mathsf{SNR}_G}{\mathsf{SNR}_H}$ bound. This can be explained by the fact that, conversely to what was assumed in the computation of the bound, the estimate of P is not perfect, and therefore $||\widehat{P}|| \leq 1$.

6.6 Conclusion

We presented a signal space (relative) calibration method that, after a calibration phase involving no hardware reference, enables the estimation of the downlink channel from the knowledge of the uplink channel estimate. This method has been shown effective by measurements on a UMTS link operating in a realistic setting. The linear reciprocity model has been experimentally shown to hold in the SISO case.

Further investigation of the reciprocity property in the MIMO case is needed (in particular, evaluation of the amount of crosstalk). On the theoretical side, a method to separate the effects of the transmitter and the receiver, in order to reduce the amount of redundant calibration information, as well as an extension to the multi-user case, are still being investigated.

Chapter 7

Pathwise channel tracking and prediction

In this chapter, we introduce a paths-based model for MIMO frequency-selective channels, based on the separation of its time-varying – e.g. Doppler – and time-invariant characteristics. This model yields a parsimonious channel representation, which is shown to simplify linear estimation and prediction of the channel. Identifiability of specular channels is studied, and an algorithm achieving identification is proposed. The performance of the proposed method is evaluated through computer simulations, and on experimental channel measurements.

7.1 Introduction

In multiple-antenna systems, the independence of the channel coefficients, often assumed *e.g.* by Telatar to compute the ergodic channel capacity, does not hold. Some correlation remains, as measured in [78], and can indeed be exploited to improve the channel tracking. Assuming that the channel realizations on different antennas correspond to different combinations (the antennas effectively sample this signal in several points in space) of the same set of waves, knowledge of this structure can be used to improve channel tracking. Methods relying on channel decomposition have been proposed, *e.g.* by Chizhik [79] where a spatial sectorization is used. However, as noticed by Svantesson *et al.* in [80], tracking is more efficient (in that the description of the channel is more parsimonious) if it can be decomposed according to a base where each vector corresponds to an impinging ray – we denote this a pathwise decomposition, although the term "beamspace representation" is sometimes used. Different decompositions and estimation methods were suggested, addressing channel estimation [81], efficient representation and transmission [70], or prediction [82, 83, 84, 85].

Pathwise methods constitute viable candidates for channel tracking and prediction, since the insight they provide into the actual channel structure – namely, separation of the channel variation into its space and time components – can improve the performance and decrease the complexity of channel tracking and prediction. Various methods have been proposed to estimate the underlying parameters, including MUSIC in [82], ESPRIT in [83] and SAGE in [81].

We use a blind identification method to analyze the time-variations of the channel: we try to separate the paths, under the assumption that the time-variation associated with the paths are independent, and fairly predictable random variables. Blind channel identification methods have been proposed before, including methods exploiting the specular structure of the channel, *e.g.* [86, 87]. However, our goal is different: we do not aim at channel equalization, but at channel tracking. Equalization consists in the identification (and, as much as possible, inversion) of a channel impulse response assumed constant (or having slow-enough fading). However, we propose to apply a similar method to channel tracking. In this context, the channel is not considered stationary but rather time-varying, and the identification method is

used to find a structure in the temporal variations of the channel. The channel realizations constitute the observed random variable, whereas the sources are the Doppler-generated complex exponentials.

In this chapter, we propose to use a specular (pathwise) approach in order to gain access to a reduced parameter set representing the channel state, in order to improve channel estimation (smoothing) and prediction. After recalling the specular channel model in Section 7.2, and outlining in Section 7.3 how it makes channel estimation and prediction easier, we provide sufficient conditions for identifiability of a specular channel in Section 7.4, and an algorithm, based on simultaneous diagonalization of the covariance matrices, that achieves identification is proposed in Section 7.5. Section 7.7 presents simulation results, and the algorithm is then checked using experimental channel measurements in Section 7.8.

Results presented in this chapter were published in [16, 17, 18].

7.2 Specular channel model

Let us consider a Multiple-Input Multiple-Output (MIMO) frequency-selective channel, with N_t transmit (Tx) and N_r receive (Rx) antennas. The impulse response of the channel between the i^{th} Tx antenna and the j^{th} Rx antenna is denoted by $h_{i,j}(t,\tau)$, where t is the time and τ is the lag.

We will henceforth work under the assumption that the channel state evolves according to a specular model. In such a model, each impulse response $h_{i,j}(t,\cdot)$ is the superposition of a finite number P of discrete paths at lag $\tau_p^{(i,j)} = l_p^{(i,j)} T_s$, $p = 1 \dots P$, resulting from either line-of-sight propagation, or one or several reflections. This model relies upon the fact that the paths between all the Tx-Rx antenna pairs have most of their characteristics in common, except for what happens near the antenna arrays. Hence, they share some properties, namely their speed w.r.t. the reflectors, and the reflection characteristics (hence their Doppler and gain are the same whatever antenna pair is considered). Each path coefficient can be decomposed into a product of two components:

- a space component $\alpha_p^{(i,j)}$, which depends on the physical properties of

path p between Tx antenna i and Rx antenna j, including antennas and reflectors position, path loss, etc.

- a time component $\beta_p(t)$ which includes the Doppler due to reflectors motion and the relative speed of the transmitter w.r.t. the receiver.

The time components $\beta_p(t)$ are assumed to be independent between paths, hence

$$\forall t' \in \mathbb{R}, p \neq p' \Rightarrow \mathrm{E}_t\left[\beta_p(t)\,\beta_{p'}(t+t')\right] = 0. \tag{7.1}$$

Note that we consider a time scale where they evolve significantly over time, *e.g.* due to the Doppler effect, and hence can be considered random processes, whereas the physical properties of the problem, comprised of the $\alpha_p^{(i,j)}$, do not vary.

In discrete time, the specular channel model yields

$$h_{n,l}^{(i,j)} \triangleq h_{i,j}(nT_s, lT_s) = \sum_{p=1}^{P^{(i,j)}} \alpha_p^{(i,j)} \beta_{n,p} \delta_{l_p^{(i,j)}}(l), \tag{7.2}$$

where we used the discretized version of the time component $\beta_{n,p} \triangleq \beta_p(nT_s)$ where T_s is the sampling interval at the receiver. Let us assume that the impulse response has finite support, and consider its discretized version

$$\underline{\mathbf{h}}_n^{(i,j)} \triangleq \left[h_{n,0}^{(i,j)}, \ldots h_{n,L-1}^{(i,j)}\right]^T, \tag{7.3}$$

with L chosen such that all the channel coefficients outside the lag interval $[0 \ldots (L-1)T_s]$ are zero. Let us further stack these into a row vector with $N_t N_r L$ coefficients

$$\underline{\mathbf{h}}_n \triangleq \left[\underline{\mathbf{h}}_n^{(1,1)^T} \ldots \underline{\mathbf{h}}_n^{(1,N_r)^T}, \underline{\mathbf{h}}_n^{(2,1)^T} \ldots \underline{\mathbf{h}}_n^{(N_t,N_r)^T}\right]^T. \tag{7.4}$$

We emphasize the fact that $\underline{\mathbf{h}}_n$ constitutes a snapshot of all the channel impulse response coefficients at time nT_s. With this notation, (7.2) can be rewritten in more compact form as

$$\underline{\mathbf{h}}_n = \mathbf{G}\underline{\mathbf{b}}_n \tag{7.5}$$

where

$$\underline{\mathbf{a}}_p^{(i,j)} \triangleq \alpha_p^{(i,j)} \left(\delta_{l_p^{(i,j)}}(0) \dots \delta_{l_p^{(i,j)}}(L-1) \right), \tag{7.6}$$

$$\underline{\mathbf{a}}_p \triangleq \left(\underline{\mathbf{a}}_p^{(1,1)} \dots \underline{\mathbf{a}}_p^{(1,N_r)}, \underline{\mathbf{a}}_p^{(2,1)} \dots \underline{\mathbf{a}}_p^{(N_t,N_r)} \right)^T, \tag{7.7}$$

$$\mathbf{G} \triangleq \left[\underline{\mathbf{a}}_1, \dots, \underline{\mathbf{a}}_P \right], \tag{7.8}$$

$$\underline{\mathbf{b}}_n \triangleq \left(\beta_{n,1}, \dots, \beta_{n,P} \right)^T. \tag{7.9}$$

7.3 Spectral factorization and linear estimation

In this section, we outline the possible improvements in channel tracking that can be achieved by deconstructing a specular channel, *i.e.* by separating the time and space properties as enounced in the previous section before doing any kind of smoothing or prediction. We seek to model the discrete-time random process $\{\underline{\mathbf{h}}_n\}$ from its noisy measurements $\tilde{\underline{\mathbf{h}}}_n = \underline{\mathbf{h}}_n + \underline{\mathbf{v}}_n$ where the noise $\{\underline{\mathbf{v}}_n\}$ is white Gaussian, i.i.d., independent from $\{\underline{\mathbf{h}}_n\}$. Assuming that both $\{\underline{\mathbf{h}}_n\}$ and $\{\underline{\mathbf{v}}_n\}$ are wide-sense stationary (WSS), let us define the (matrix) covariances

$$\mathrm{R}_{\tilde{\underline{\mathbf{h}}}\tilde{\underline{\mathbf{h}}}}(u) \triangleq \mathrm{E}_n \left[\tilde{\underline{\mathbf{h}}}_{n+u} \tilde{\underline{\mathbf{h}}}_n^H \right] \quad \text{and} \quad \mathrm{R}_{\underline{\mathbf{h}}\tilde{\underline{\mathbf{h}}}}(u) \triangleq \mathrm{E}_n \left[\underline{\mathbf{h}}_{n+u} \tilde{\underline{\mathbf{h}}}_n^H \right] \tag{7.10}$$

where $\mathrm{E}_n[\cdot]$ is the expectation operator taken over n, and the z-transforms

$$\mathrm{S}_{\tilde{\underline{\mathbf{h}}}\tilde{\underline{\mathbf{h}}}}(z) \triangleq \sum_{u=-\infty}^{+\infty} \mathrm{R}_{\tilde{\underline{\mathbf{h}}}\tilde{\underline{\mathbf{h}}}}(u)\, z^{-u} \quad \text{and} \tag{7.11}$$

$$\mathrm{S}_{\underline{\mathbf{h}}\tilde{\underline{\mathbf{h}}}}(z) \triangleq \sum_{u=-\infty}^{+\infty} \mathrm{R}_{\underline{\mathbf{h}}\tilde{\underline{\mathbf{h}}}}(u)\, z^{-u}. \tag{7.12}$$

The best linear estimator (in terms of mean square error) of $\underline{\mathbf{h}}_{n+\lambda}, \lambda \geq 0$ given $\{\tilde{\underline{\mathbf{h}}}_k\}_{k=-\infty}^{n}$ is

$$\hat{\underline{\mathbf{h}}}_{n+\lambda} \triangleq \sum_{i=-\infty}^{+\infty} \mathrm{K}_{n-i} \underline{\mathbf{h}}_i \tag{7.13}$$

where the matrix filter coefficients K_u are determined in the z-transform domain

$$\mathrm{K}(z) \triangleq \sum_{u=-\infty}^{+\infty} \mathrm{K}_u \; z^{-u} \tag{7.14}$$

by [88]

$$\mathrm{K}(z) = \left\{ z^{\lambda} \mathrm{S}_{\underline{\mathbf{h}}\tilde{\underline{\mathbf{h}}}}(z) \mathrm{L}^{-*}(z^{-*}) \right\}_{+} \mathrm{R_e}^{-1} \mathrm{L}^{-1}(z), \tag{7.15}$$

where $\mathrm{L}(z)$ and $\mathrm{R_e}$ come from the spectral factorization

$$\mathrm{S}_{\tilde{\underline{\mathbf{h}}}\tilde{\underline{\mathbf{h}}}}(z) = \mathrm{L}(z) \mathrm{R_e} \mathrm{L}^{*}(z^{-*}). \tag{7.16}$$

In general, spectral factorization is hard to compute in the case of vector-valued processes, and (7.15) is not feasible. Note that due to the independence between $\{\underline{\mathbf{v}}_n\}$ and $\{\underline{\mathbf{h}}_n\}$,

$$\mathrm{S}_{\underline{\mathbf{h}}\tilde{\underline{\mathbf{h}}}}(z) = \mathrm{S}_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(z) \quad \text{and} \quad \mathrm{S}_{\tilde{\underline{\mathbf{h}}}\tilde{\underline{\mathbf{h}}}}(z) = \mathrm{S}_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(z) + \mathrm{S}_{\underline{\mathbf{v}}\underline{\mathbf{v}}}(z). \tag{7.17}$$

7.3.1 Specular model and spectral factorization

Under the assumption that the channel variations follow the specular model (7.5), the covariances become

$$\begin{aligned} \mathrm{R}_{\tilde{\underline{\mathbf{h}}}\tilde{\underline{\mathbf{h}}}}(u) &= \mathbf{G} \mathrm{E}_n \left[\underline{\mathbf{b}}_{n+u} \underline{\mathbf{b}}_n^H \right] \mathbf{G}^H + \mathrm{E}_n \left[\underline{\mathbf{v}}_{n+u} \underline{\mathbf{v}}_n^H \right] && (7.18) \\ &= \mathbf{G} \mathrm{R}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(u) \mathbf{G}^H + \mathrm{R}_{\underline{\mathbf{v}}\underline{\mathbf{v}}}(u), && (7.19) \\ \mathrm{R}_{\underline{\mathbf{h}}\tilde{\underline{\mathbf{h}}}}(u) &= \mathbf{G} \mathrm{E}_n \left[\underline{\mathbf{b}}_{n+u} \underline{\mathbf{b}}_n^H \right] \mathbf{G}^H = \mathbf{G} \mathrm{R}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(u) \mathbf{G}^H. && (7.20) \end{aligned}$$

Note that in eqs. (7.18) and (7.20) the factor $\mathbf{G}$ is independent of the lag u. Therefore, the z-transforms can be factored as

$$\begin{aligned} \mathrm{S}_{\tilde{\underline{\mathbf{h}}}\tilde{\underline{\mathbf{h}}}}(z) &= \mathbf{G} \mathrm{S}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(z) \mathbf{G}^H + \mathrm{S}_{\underline{\mathbf{v}}\underline{\mathbf{v}}}(z) \quad \text{and} && (7.21) \\ \mathrm{S}_{\underline{\mathbf{h}}\tilde{\underline{\mathbf{h}}}}(z) &= \mathbf{G} \mathrm{S}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(z) \mathbf{G}^H. && (7.22) \end{aligned}$$

Let us define

$$s_{bb}^{(p)}(z) \triangleq \sum_{u=-\infty}^{+\infty} \mathrm{E}_n \left[\beta_{n+u,p} \beta_{n,p}^{*} \right] z^{-u} \quad \text{for} \quad p = 1 \ldots P. \tag{7.23}$$

Note that $\mathrm{S}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(z)$ is diagonal, since we assume that the $\beta_p(t)$ are independent:

$$\mathrm{S}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(z) = \mathrm{diag}\left(s_{bb}^{(1)}(z), \ldots, s_{bb}^{(P)}(z) \right). \tag{7.24}$$

Note that this structure allows to obtain the spectral factorization of $\mathrm{S}_{\underline{\tilde{\mathbf{h}}}\underline{\tilde{\mathbf{h}}}}(z)$ by performing P independent, scalar spectral factorizations of the $s_{bb}^{(p)}(z)$ (equivalently, this means that the random process $\{\underline{\mathbf{h}}_n\}$ can be accurately tracked by tracking P scalar processes).

In the following sections, we will address the following identifiability problem: assuming the knowledge of the spectrum $\mathrm{S}_{\underline{\tilde{\mathbf{h}}}\underline{\tilde{\mathbf{h}}}}(z)$, we show that if the $s_{bb}^{(p)}(z)$ are linearly independent polynomials, it is possible to identify them up to a permutation and a complex scalar coefficient. Fortunately, this is sufficient for our needs, since all possible solutions yield the same predictor $\mathrm{K}(z)$. Then, we discuss an algorithm that achieves this identification.

7.4 Identifiability

In this section, we show that if the spectrums $s_{bb}^{(p)}(z)$ are linearly independent polynomials, the spectral factorization (7.22) is unique up to a permutation and a scalar coefficient applied to the columns of $\mathbf{G}$.

Let us assume that $\mathbf{G}$ has full column rank, and let $(\underline{\mathbf{c}}_1, \ldots, \underline{\mathbf{c}}_P)$ be an orthonormal base of the column subspace of $\mathbf{G}$. Let $\mathbf{C} \triangleq [\underline{\mathbf{c}}_1, \ldots, \underline{\mathbf{c}}_P]$. Let $\mathbf{R}$ denote the representation of $\mathbf{G}$ in this base, such that $\mathbf{G} = \mathbf{C}\mathbf{R}$. Let us assume that $\mathrm{S}_{\underline{\mathbf{h}}\underline{\tilde{\mathbf{h}}}}(z)$ has an alternative factorization

$$\mathrm{S}_{\underline{\mathbf{h}}\underline{\tilde{\mathbf{h}}}}(z) = \mathbf{H}\mathrm{S}_{\underline{\mathbf{b}}'\underline{\mathbf{b}}'}(z)\mathbf{H}^H, \tag{7.25}$$

and show that $\mathbf{G}$ and $\mathbf{H}$ are identical up to a permutation and a linear scaling of their columns. Using the fact that $\mathbf{C}^H\mathbf{C} = \mathrm{I}_P$, the decomposition in (7.22) yields

$$\mathrm{S}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(z) = \mathbf{R}^{-1}\mathbf{C}^H\mathrm{S}_{\underline{\mathbf{h}}\underline{\tilde{\mathbf{h}}}}(z)\mathbf{C}\mathbf{R}^{-H}. \tag{7.26}$$

Hence, using (7.25) and defining the $P \times P$ matrix $\mathbf{S} \triangleq \mathbf{R}^{-1}\mathbf{C}^H\mathbf{H}$,

$$\mathrm{S}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(z) = \mathrm{S}\mathrm{S}_{\underline{\mathbf{b}}'\underline{\mathbf{b}}'}(z)\mathrm{S}^H. \tag{7.27}$$

Therefore we need to prove that S is the product of a permutation matrix and a diagonal matrix. Let $\underline{\mathbf{s}}_i, i = 1 \ldots P$ denote the columns of $\mathbf{S}$. The

diagonal structure of $\mathrm{S}_{\underline{\mathbf{b}}'\underline{\mathbf{b}}'}(z)$ lets us rewrite

$$\mathrm{S}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(z) = \sum_{i=1}^{P} \underline{\mathbf{s}}_i \underline{\mathbf{s}}_i^H s^{(i)}_{\underline{\mathbf{b}}'\underline{\mathbf{b}}'}(z). \tag{7.28}$$

This implies that each $\underline{\mathbf{s}}_i\underline{\mathbf{s}}_i^H$ is diagonal, otherwise the off-diagonal terms would yield an identically zero linear combination of $s^{(i)}_{\underline{\mathbf{b}}'\underline{\mathbf{b}}'}(z)$'s, which contradicts the linear independence assumption. This implies that each $\underline{\mathbf{s}}_i$ has at most one non-zero coefficient, which is equivalent to saying that S represents the product of a permutation matrix and a diagonal matrix.

7.5 Practical identification method

The previous discussion has shown that any factorization of the form of (7.25) is an equally good way of decomposing $\{\underline{\mathbf{h}}_n\}$ into scalar, independent processes. In this section, we present an algorithm to find one of these decompositions.

Let us assume that the noise level is known, or has been estimated, hence $\mathrm{S}_{\underline{\mathbf{h}}}(z)$ is known, or equivalently, $\mathrm{R}_{\underline{\mathbf{hh}}}(u)$ is known for $u \in \mathbb{Z}$. The algorithm that we present here provides a matrix $\mathbf{B}$ that decomposes $\mathrm{S}_{\underline{\mathbf{h}}}(z)$ into independent processes, *i.e.* $\mathbf{B}\mathrm{S}_{\underline{\mathbf{h}}}(z)\mathbf{B}^H$ is diagonal. We restrict the problem to the signal subspace, and consider $\mathbf{C}^H\mathrm{R}_{\underline{\mathbf{hh}}}(0)\mathbf{C}$. Since it is positive semi-definite, it can be decomposed according to its eigenstructure:

$$\mathbf{C}^H\mathrm{R}_{\underline{\mathbf{hh}}}(0)\mathbf{C} = \mathbf{W}\mathbf{D}\mathbf{W}^H, \tag{7.29}$$

where $\mathbf{W}$ is a $P \times P$ unitary matrix, and $\mathbf{D}$ is diagonal, and contains the (non-negative) eigenvalues. Notice that $(\mathbf{W}\sqrt{\mathbf{D}})(\mathbf{W}\sqrt{\mathbf{D}})^H$ constitutes a Cholesky decomposition of $\mathbf{C}^H\mathrm{R}_{\underline{\mathbf{hh}}}(0)\mathbf{C}$. Since $\left(\mathbf{R}\sqrt{\mathrm{R}_{\underline{\mathbf{bb}}}(0)}\right)\left(\mathbf{R}\sqrt{\mathrm{R}_{\underline{\mathbf{bb}}}(0)}\right)^H$ is also a Cholesky decomposition of the same matrix, they are unitarily similar [89], *i.e.* there exist an unitary matrix $\mathbf{Q}$ s.t.

$$\left(\mathbf{R}\sqrt{\mathrm{R}_{\underline{\mathbf{bb}}}(0)}\right)^H = \mathbf{Q}\left(\mathbf{W}\sqrt{\mathbf{D}}\right)^H. \tag{7.30}$$

Obviously, finding $\mathbf{Q}$ would let us identify $\mathbf{R}$ up to the scalar uncertainties contained in $\sqrt{\mathrm{R}_{\underline{\mathbf{bb}}}(0)}$. In order to find it, we use the fact that

$$\begin{aligned}\left(\mathbf{W}\sqrt{\mathbf{D}}\right)^{-1}\mathbf{C}^H\mathrm{R}_{\underline{\mathbf{hh}}}(u)\mathbf{C}\left(\mathbf{W}\sqrt{\mathbf{D}}\right)^{-H}\\ = \quad \mathbf{Q}^H\mathrm{R}_{\underline{\mathbf{bb}}}(u)\mathrm{R}_{\underline{\mathbf{bb}}}(0)^{-1}\mathbf{Q} \quad \forall u \in \mathbb{Z}.\end{aligned} \tag{7.31}$$

Under our assumptions, there is a unique way (up to a permutation $\mathbf{P}$) of diagonalizing the spectrum matrix, as demonstrated in Section 7.4, hence if $\mathbf{V}$ is a unitary matrix that diagonalizes $\mathbf{V}\sqrt{\mathbf{D}^{-1}}\mathbf{W}^H\mathbf{C}^H\mathrm{R}_{\underline{\mathbf{hh}}}(u)\mathbf{C}\mathbf{W}\sqrt{\mathbf{D}^{-1}}\mathbf{V}^H$ for all $u \in \mathbb{Z}$, then $\mathbf{Q} = \mathbf{P}^T\mathbf{V}$. Finding $\mathbf{V}$ is the well-known simultaneous diagonalization problem [90], and can be solved efficiently using an algorithm based on Jacobi angles [91]. It follows that $\{\underline{\mathbf{h}}_n\}$ is transformed into an arbitrary vector process $\{\underline{\mathbf{y}}_n\}$, with

$$\underline{\mathbf{y}}_n \triangleq \mathbf{V}\sqrt{\mathbf{D}^{-1}}\mathbf{W}^H\mathbf{C}^H\underline{\mathbf{h}}_n = \mathbf{P}\sqrt{\mathrm{R}_{\underline{\mathbf{bb}}}(0)^{-1}}\underline{\mathbf{b}}_n, \tag{7.32}$$

and $\mathrm{S}_{\underline{\mathbf{y}}}(z)$ is diagonal: $\mathbf{V}\sqrt{\mathbf{D}^{-1}}\mathbf{W}^H\mathbf{C}^H$ is a possible $\mathbf{B}$. The uncertainties outlined in Section 7.4 appear clearly in (7.32): each component of $\{\underline{\mathbf{y}}_n\}$ is normalized to unitary variance, and the permutation $\mathbf{P}$ is unknown.

Obviously, the theoretical identifiability outlined in Section 7.4, is not realistic in practice, since implementation constraints would restrict the knowledge of $\mathrm{R}_{\underline{\mathbf{hh}}}(u)$ to a limited range of u. Also, the requirement of linear independence of the columns of $\mathbf{G}$, as well as the linear independence of the z-spectrums of the time coefficients, are not guaranteed to be fulfilled in real life. However, our simulations show that these limitations do not seem to incur significant problems in practice.

7.6 Applications

One of the interest of a specular channel model is its long-term validity. Since it closely follows the physical channel structure, and hence separates the spatial and temporal properties of the channel, estimation or prediction of the time coefficients only (the $\beta_p(t)$) together with the knowledge of $\mathbf{G}$

provides knowledge of the channel state, following from

$$\begin{aligned}
\underline{\mathbf{h}}_n &= \mathbf{C}\mathbf{R}\underline{\mathbf{b}}_n && (7.33)\\
&= \mathbf{C}\mathbf{W}\sqrt{\mathbf{D}}\mathbf{Q}^H\sqrt{\mathrm{R}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(0)^{-1}}\underline{\mathbf{b}}_n && (7.34)\\
&= \mathbf{C}\mathbf{W}\sqrt{\mathbf{D}}\mathbf{Q}^H\mathbf{P}^{-1}\underline{\mathbf{y}}_n && (7.35)\\
&= \mathbf{C}\mathbf{W}\sqrt{\mathbf{D}}\mathbf{V}^H\underline{\mathbf{y}}_n && (7.36)
\end{aligned}$$

where we used successively (7.5), (7.30), (7.32), and the definition of $\mathbf{Q}$. Note that these matrices need to be estimated first, and in particular the choice of the number of independent components to consider. In practice, this parameter can be chosen as the number of non-noise eigenvalues of $\mathrm{R}_{\underline{\tilde{\mathbf{h}}}\underline{\tilde{\mathbf{h}}}}(0)$. Then, $\mathbf{C}$ can be obtained by orthogonalizing the corresponding set of eigenvectors. Subsequently, the channel estimate can be obtained as

$$\widehat{\underline{\mathbf{h}}}_n = \mathbf{C}\mathbf{W}\sqrt{\mathbf{D}}\mathbf{V}^H\widetilde{\underline{\mathbf{y}}}_n, \tag{7.37}$$

where $\widetilde{\underline{\mathbf{y}}}_n$ is obtained by any estimation method (for instance smoothing, linear prediction...) from the $\widehat{\underline{\mathbf{y}}}_n$. Since the coefficients in $\underline{\mathbf{y}}_n$ are independent processes, the burden of the estimation method is greatly decreased. Several kinds of processing can be applied at this point, for instance smoothing if the goal is to increase channel estimation accuracy. Prediction can be useful in systems relying on Channel State Information (CSI) at the Transmitter (CSIT) to enhance the link quality: since duplex systems mainly rely on a feedback scheme to transmit the channel state information from the receiver, and since the channel state information can not be fed back in a negligible amount of time, the ability for the transmitter to extrapolate CSI from past values can therefore be an important asset in the actual use of a CSIT-exploiting transmission scheme [16]. The nature of the underlying processes must also be considered. For instance, pure Doppler effect would yield an autoregressive process of order 1.

7.7 Simulation results

The algorithm proposed in Section 7.5 has been proven in Section 7.4 to identify perfectly the system in the noiseless case. We present here some

results obtained in a more practical simulation setting. We simulated a setting with 2×2 antennas, and a delay spread limited to $L = 5$ samples. The channel is an actual specular channel with $P = 4$ paths, where each path is determined by randomly generated integer lags, uniform Directions of Departure (DoD) and Arrival (DoA), a Gaussian gain and a random AR3 process. The algorithm works with approximate covariance matrices estimated from a finite-length measurement interval of N successive realizations of $\{\underline{\mathbf{h}}_n\}$. The number P' of estimated independent random processes to estimate is set artificially, and several cases ($P' > P$, $P' < P$, $P' = P$) are presented here.

The figure of merit used in these simulations is derived from the fact that in the absence of noise, with the notations of Section 7.5,

$$\mathbf{P}^{\mathbf{T}}\mathbf{V}\sqrt{\mathbf{D}^{-1}}\mathbf{W}^H\mathbf{C}^H\mathbf{G} = \sqrt{\mathrm{R}_{\underline{\mathbf{b}}\underline{\mathbf{b}}}(0)^{-1}}, \tag{7.38}$$

i.e. it is a diagonal matrix. When noise is present, this matrix is computed from the true $\mathbf{G}$ and the specular model as estimated by the proposed algorithm, hence it is not perfectly diagonal. Hence, denoting $[x_{i,j}]_{i,j} \triangleq \mathbf{P}^{\mathbf{T}}\mathbf{V}\sqrt{\mathbf{D}^{-1}}\mathbf{W}^H\mathbf{C}^H\mathbf{G}$, $x_{p,p}$ is the amount of energy from $\{\beta_{n,p}\}_n$ that is correctly attributed to the pth estimated process, and the $x_{i,p}, i \neq p$ represent the crosstalk with other processes. Thus, we define a global signal-to-interference ratio (SIR) as

$$\mathsf{SIR} \triangleq \frac{\sum_{p=1}^{P} x_{p,p}^2}{\sum_{p=1}^{P}\sum_{i\neq p} x_{p,i}^2}. \tag{7.39}$$

This value is plotted on Figure 7.1 for various configurations, with respect to the SNR of the raw channel estimates $\underline{\tilde{\mathbf{h}}}_n$.

These results clearly show the influence of the the quality of estimation of the covariances: increasing the number of realizations used for estimating the channel statistics from $N = 20$ to $N = 100$, then 1000, has a relatively bigger influence than the SNR variations over the range pictured here. The influence of overestimating the number of paths can be estimated by comparing the $P' = P = 4$ to the $P' = 6$ case. Overestimation of the number of paths yields an almost negligible decrease in the SIR of the P correctly identified paths. The case where $P' < P$ is interesting in that it represents the resilience of the identification algorithm to model mismatch, *i.e.* when the signal does not conform to the assumptions that support our method. In this case,

the SIR criterion is only computed on the P' separated processes, which are statistically the strongest. This explains the fact that we observe a slight SIR increase in this case, for low input SNR values, and goes to prove that the P' strongest paths are correctly identified. However, this metric is hiding the fact that the channel is not fully analyzed, and hence not as predictable. Evidencing this would require a more involved simulation setup, where the modeling error variance would be considered.

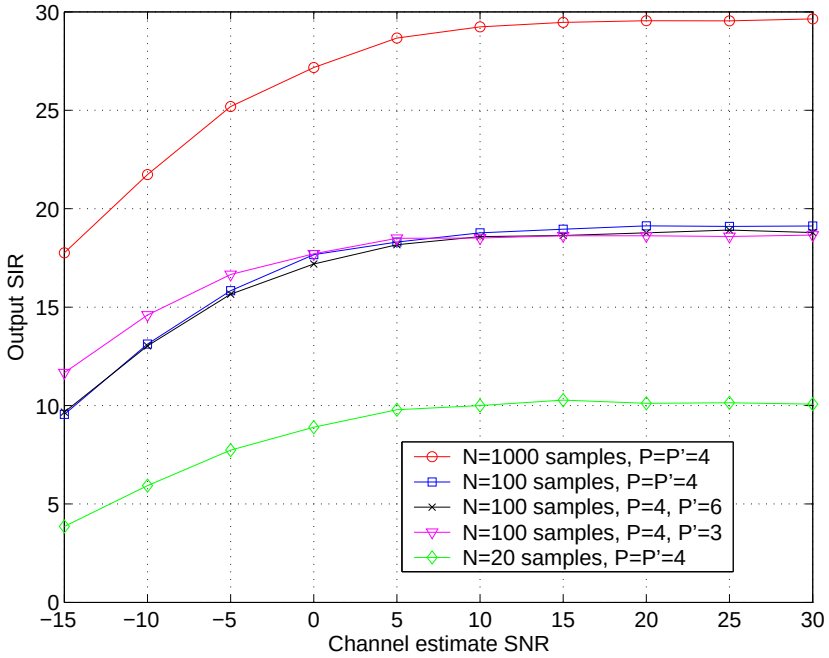


Figure 7.1: SIR of the process separation

7.8 Experimental results

The proposed identification method was applied to experimental data obtained from a prototype UMTS TDD link [77] operating on a 3.84MHz wide channel in the 1900-1920MHz band. The setting is comprised of two rooftop antennas connected to the base station, and a terminal connected to a portable antenna that was operated inside the building. The channel mea-

surements were performed on the uplink channel, in the framework of an actual UMTS connection. The considered channel has 1 Tx and 2 Rx antennas, and the length of the observed impulse responses is $L = 36$ samples. The complete SIMO channel impulse response is estimated every 10ms using conventional channel estimation techniques, by exploiting the training sequences embedded in the UMTS traffic. The oscillators of both Rx antennas are synchronized.

$\mathbf{B}$ is first computed by applying the proposed algorithm to a series of consecutive channel measurements, and the corresponding series $\underline{\mathbf{y}}_n$ is computed according to (7.32). The algorithm operates with approximate covariance matrices estimated from a finite-length measurement interval of 50 successive realizations (spanning 500ms) of $\{\underline{\mathbf{h}}_n\}$. The number P' of independent random processes to track is set arbitrarily. Two series of measurements were conducted. In the first setting (fixed setting), the MT antenna lies on a table and no noticeable movement is made around the antennas. In the second setting (moving setting), the MT antenna is hand-held and moved rapidly by a human operator.

Figure 7.2 presents an example of the output of the algorithm for the fixed setting. The top plot represents the angles of the coefficients of $\underline{\mathbf{y}}_n$ associated with the $P' = 3$ strongest modes. The bottom plot shows the profile of the modes as identified by the algorithm (each mode corresponds to a column of $\mathbf{C}\mathbf{W}\sqrt{\mathbf{D}}\mathbf{V}^H$, the generalized inverse of $\mathbf{B}$). Since $\underline{\mathbf{h}}_n$ contains the impulse responses from both Rx antennas, $\mathbf{B}$ has a similar structure, and therefore the left part of the plot represents the channel impulse response seen by the first Rx antenna, while the right part represents the channel impulse response seen by the second Rx antenna. Note that the arbitrary scaling has been partially resolved by normalizing each component of $\underline{\mathbf{y}}_n$ to unit average energy, so that the three modes can be plotted on the same scale.

Figure 7.3 was obtained for the second series of measurements (moving setting) through the same process. In both cases, the time-varying components exhibit a high predictability. In particular, the linear phase evolution exhibited by the fixed setting can be accurately predicted by an autoregressive (AR) model of order 1. As expected, the temporal evolution of the channel in the moving setting does not exhibit such a linear behaviour, although it could be predictable by an higher-order AR model.

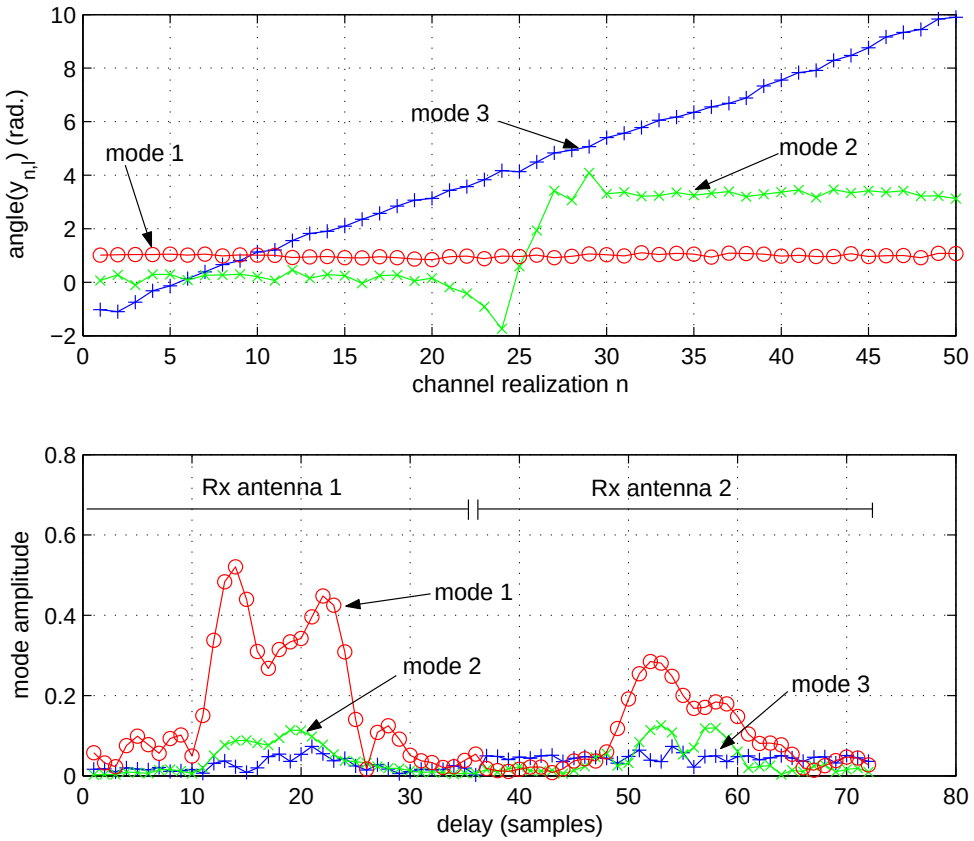


Figure 7.2: Example of modal decomposition, fixed setting.

It is noticeable from both Figures 7.2 and 7.3 that the modes spans both Rx antennas. This indicates that the channel structure is correlated between the antennas, and that the algorithm has correctly identified a common structure in their temporal evolution. This is not contradictory with the fact that this setting provides antenna diversity, since the actual impulse responses are linear combinations (with time-varying weights) of the modes plotted here. Although the result can vary rapidly, the knowledge of the underlying structure allows for a more accurate long-term tracking.

In order to evaluate the feasibility of channel prediction, one-step linear prediction was applied to the time-varying coefficients. For each time instant

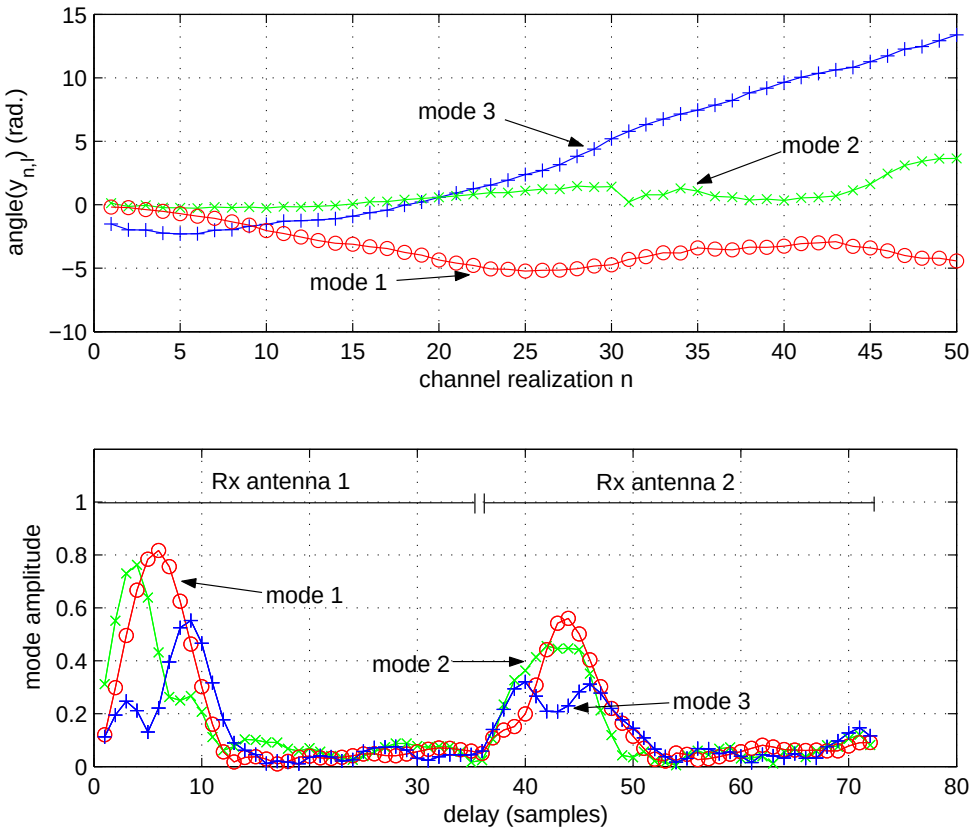


Figure 7.3: Example of modal decomposition, moving setting.

n, after computing $\mathbf{B}$, the time-series corresponding to the last 50 channel realizations $\underline{\mathbf{y}}_{n-49} \ldots \underline{\mathbf{y}}_n$ is extrapolated using an AR model. The resulting value $\widehat{\underline{\mathbf{y}}}_{n+1}$ is used to compute the predicted channel $\widehat{\underline{\mathbf{h}}}_{n+1} = \mathbf{C}\mathbf{W}\sqrt{\mathbf{D}}\mathbf{V}^H\widehat{\underline{\mathbf{y}}}_{n+1}$. Since the true channel value is not known, a noise metric defined as

$$\alpha \triangleq \frac{\mathrm{E}_n\left[|\widetilde{\underline{\mathbf{h}}}_{n+1}|^2\right]}{\mathrm{E}_n\left[|\widehat{\underline{\mathbf{h}}}_{n+1} - \widetilde{\underline{\mathbf{h}}}_{n+1}|^2\right]} \tag{7.40}$$

that can be computed from the noisy channel measurements, was used. There is a saturation effect associated with this metric, since even if the prediction is perfect ($\widehat{\underline{\mathbf{h}}}_{n+1} = \underline{\mathbf{h}}_{n+1}$), α goes to the input SNR. For comparison, the input SNR is 11.5dB for the fixed measurements, and 11.9dB for the moving

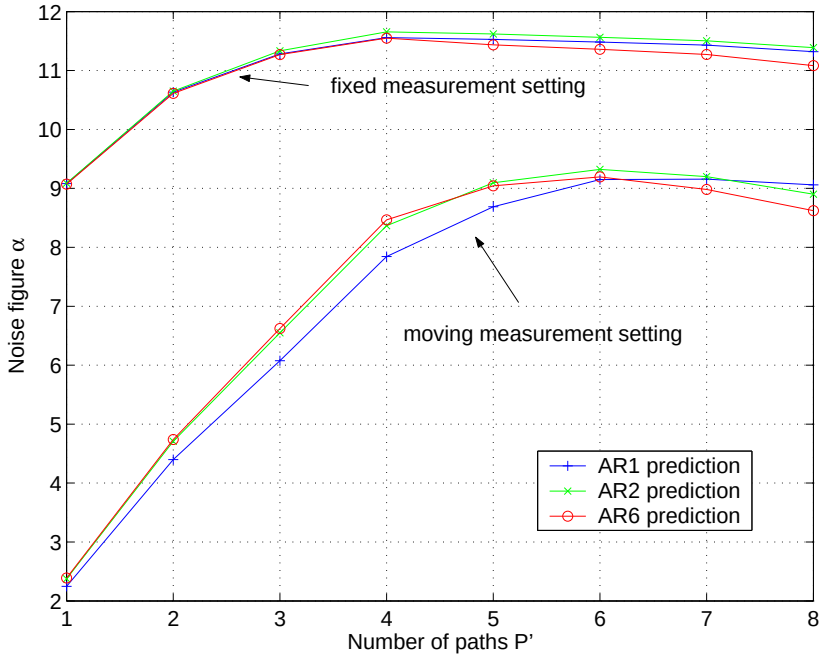


Figure 7.4: Predictor noise figure α.

measurements.

Figure 7.4 shows the evolution of α (in dB) for various orders of the AR predictor, and for various numbers of paths P', for both sets of measurements. As expected, the channel in the fixed setting exhibits a very high predictability, and the predictor reaches the optimality ($\alpha = \mathsf{SNR}$) for $P' = 4$ tracked paths. Increasing P' over this values slightly increases the amount of noise on the output, although for P' ranging from 3 to 8 paths, α is only a fraction of a dB from the optimality. Varying the order of the AR predictor has little influence on the performance in the fixed setting, as evidenced by the three almost superimposed curves.

This optimality is not reached for the moving setting, with α getting no closer than 2.5dB from the bound set by the SNR. The maximum accuracy is reached by tracking 6 subspaces. Note that this does not necessarily mean that 6 paths are identified, since this can be caused by the fact that the channel model can not be assumed stationary over the .5s analyzed here. The

order of the AR predictor seems more critical in this case, as exemplified by the fact that the AR2 performs significantly better than the AR1, especially when P' is underestimated. This could be explained by the fact that an AR2 model can better track a mixture of AR processes than an AR1, although no definitive conclusion can be drawn at this point.

7.9 Conclusion

We presented a channel modeling method based on the assumption that the channel follows a specular structure. We showed how this structured model, by separating space and time-components, lends itself to simplified tracking, including smoothing and prediction, once the underlying space and time characteristics are separated. We showed that under mild assumptions on the channel characteristics, these components are identifiable, and proposed a method based on simultaneous diagonalization of the covariance matrices that achieves the identification. We evaluated the performance of the proposed method through simulations.

General conclusion

In this thesis, we presented various problems and their solutions for coding and signal processing associated with MIMO wireless communications. They can be broadly classified into two categories, coding and equalization methods, and channel modeling methods.

In the first part, coding and equalization methods for OFDM systems were proposed:

- a new STFC named Space-Time-Frequency Spreading was proposed. STFS applies to OFDM systems using multiple Tx antennas, and maximally exploits the channel space and frequency diversity, as well as time diversity if some coding delay is acceptable.

- the streams-based structure, and the independence of the interfering streams of STFS enable the use of an iterative PIC decoder at the Rx side.

- STFS has been shown to perform comparatively better than Threaded STC codes, and, for small constellation sizes, as well as BICM, while using only low-complexity (Viterbi-based) decoding.

- a representation for the Doppler spectrum in CP-OFDM systems was proposed, which let the receiver trade ISI for ICI. Depending on which of the delay spread and the Doppler spread is lower, equalization can therefore take place in both domains interchangeably, allowing for lower-complexity equalization.

In the second part, the issues of CSIT and MIMO channel time-evolution modeling were addressed.

- we showed that CSIT is only valuable for relatively low SNR s. At asymptotically high SNR, CSIT only provides a constant mutual information increase w.r.t. the CSIR-only case, this constant representing a vanishing fraction of the achieved mutual information. For $N_t > N_r$, this constant has been shown to be zero.

- the resilience of eigenwaterfilling methods w.r.t. inaccurate CSIT was studied. We showed that a relatively low level (-5dB) of AWGN on the channel estimate is acceptable.

- the resilience of eigenwaterfilling methods w.r.t. improper Tx and Rx absolute calibration (phase and amplitude mismatch between antennas) was studied. It was shown that improper Tx calibration greatly decreases the mutual information, whereas Rx calibration has virtually no influence on the achievable performance.

- a channel reciprocity model was introduced, for the case of TDD systems. It applies to any number of Tx and Rx antennas, and frequency-flat as well as frequency selective channels. The frequency-selective model was verified to apply to a good precision using SISO channel measurements.

- a *relative calibration* method, suitable for the proposed reciprocity model, was introduced. It enables the exploitation of reciprocity while lifting the requirement for absolute (hardware-based) calibration at both the Tx and Rx. After a collaborative training phase (relative calibration), the downlink channel can be estimated directly from the uplink channel estimate, without the need for continuous channel feedback.

- a pathwise model was introduced to represent the time-variations of MIMO frequency-selective channels. It does not rely on DoD/DoA estimation, but rather uses blind methods to decompose the time-varying process into, for each path,

  - a constant matrix representing the combined physical characteristics of the environment (angles and delays, space correlation, antenna gains, etc.)

  - a time-varying, structured process (such as a Doppler series) that can be easily tracked or predicted

  The low number of time-varying parameters ensures that they are estimated with more accuracy than the channel coefficients.

- an identification algorithm for the proposed specular model was introduced, and shown to properly identify the channel characteristics, provided that the paths have linearly independent space-time signatures, and that the time-varying processes have linearly independent spectra.

- the performance of the identification algorithm was evaluated using synthetic data, and examples of space-time signatures and time series obtained for measured data were presented in Section 7.8.

This thesis has combined contributions from both signal processing and information theory, in order to optimize the use of the wireless MIMO propagation channel. In particular, the use of the mutual information criterion facilitated evaluation of various linear and additive impairments to the channel, and we feel that a number of similar problems can be transposed and better understood by using this method. In the scope of channel modeling, we felt the need to add our specular channel decomposition method to the already large corpus of wireless channel models, since every other method lacks either the realistic structure provided by a specular model, or the simplicity in the estimation of the slow-fading parameters. We feel that our method constitutes a competitive candidate for small time-range channel prediction. Finally, we can see a great number of application of the relative calibration method it general-purpose applications, mostly due to the fact that it achieves the same goals as absolute calibration, while requiring only minor protocol modifications and no extra hardware.

Bibliography

[1] J. H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE Journal on Selected Areas in Communications*, vol. 5, pp. 871–878, June 1987.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," technical memorandum, Bell Laboratories, Lucent Technologies, October 1995. available at `http://mars.bell-labs.com/papers/proof/`.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, March 1998.

[4] C. Shannon, "Communications in the presence of noise," in *Proc.of the IRE*, vol. 37, pp. 10–21, 1949.

[5] M. Guillaud and D. T. M. Slock, "Channel modeling and associated inter-carrier interference equalization for OFDM systems with high Doppler spread," in *Proc.* 28^{th} *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. IV, pp. 237–240, April 2003.

[6] M. Guillaud and D. T. M. Slock, "Full-rate full-diversity space-frequency coding for MIMO OFDM systems," in *Proc.* 3^{rd} *IEEE Benelux Signal Processing Symposium (SPS-2002)*, (Leuven, Belgium), March 2002.

[7] M. Guillaud and D. T. M. Slock, "Multi-stream coding for MIMO OFDM systems with space-time-frequency spreading," in *Proc.* 5^{th} *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, (Honolulu, HI, USA), October 2002.

[8] M. Guillaud and D. T. M. Slock, "Influence du compromis entre exploitation de la diversité et gain de codage dans les systèmes OFDM multi-entrées multi-sorties," in *Proc. 9th GRETSI Symposium on Signal and Image Processing (GRETSI)*, vol. 2, (Paris, France), pp. 359–362, Sept. 2003.

[9] G. Caire, K. Gosse, M. Guillaud, A. Guillén i Fàbregas, A. Ribeiro Dias, S. Rouquette-Léveil, and D. T. M. Slock, "Investigation of multiple transmit multiple receive antenna techniques without channel knowledge at the emitter, in the Hiperlan/2 context," project report, RNRT, October 2002. Project Antipode, Deliverable 2.1.

[10] G. Caire, K. Gosse, M. Guillaud, A. Guillén i Fàbregas, A. Ribeiro Dias, S. Rouquette-Léveil, and D. T. M. Slock, "Investigation of multiple transmit multiple receive antenna techniques with channel knowledge at the transmitter, in the Hiperlan/2 context," project report, RNRT, October 2002. Project Antipode, Deliverable 2.2.

[11] G. Caire, K. Gosse, M. Guillaud, A. Guillén i Fàbregas, A. Ribeiro Dias, S. Rouquette-Léveil, and D. T. M. Slock, "Evaluation of multiple transmit multiple receive techniques in a realistic propagation environment at 5GHz," project report, RNRT, October 2002. Project Antipode, Deliverable 2.3.

[12] G. Caire, K. Gosse, M. Guillaud, A. Guillén i Fàbregas, A. Ribeiro Dias, S. Rouquette-Léveil, and D. T. M. Slock, "Analysis of the impact of a realistic propagation environment at 5GHz on MTMR techniques : Simulation results," project report, RNRT, September 2003. Project Antipode, Deliverable 2.6.

[13] M. Guillaud and D. T. M. Slock, "Investigation of channel knowledge influence in the HIPERLAN/2 context," project report, RNRT, October 2003. Project Antipode, Deliverable 2.11.

[14] M. Guillaud, D. T. M. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration," in *Proc. Eighth International Symposium on Signal Processing and Its Applications (ISSPA '05)*, (Sydney, Australia), August 2005.

[15] D. T. M. Slock, M. Guillaud, and R. Knopp, "Process for estimating the channel state in a transmitter of a digital communication system and apparatus for doing the same." EU patent application #EP05368008, March 2005.

[16] M. Guillaud and D. T. M. Slock, "A specular approach to MIMO frequency-selective channel tracking and prediction," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, (Lisbon, Portugal), July 2004.

[17] M. Guillaud and D. T. M. Slock, "MIMO frequency-selective channel modeling based on pathwise dynamics," in *Proc.* 38^{th} *Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA, USA), November 2004.

[18] M. Guillaud and D. T. M. Slock, "Pathwise MIMO channel modeling and estimation," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, (New York, NY, USA), June 2005.

[19] C. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[20] T. S. Rappaport, *Wireless Communications, 2nd Edition*. Prentice Hall, 2002.

[21] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, Inc., 1968.

[22] Y. Li and L. J. Cimini Jr., "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Transactions on Communications*, vol. 49, pp. 401–404, March 2001.

[23] A. Gorokhov and J.-P. Linnartz, "Robust OFDM receivers for dispersive time varying channels: Equalization and channel acquisition," in *Proc. IEEE International Conference on Communications (ICC)*, vol. 1, pp. 470–474, 2002.

[24] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Estimation of fast fading channels in OFDM," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 1, pp. 465–470, 2002.

[25] E. N. Onggosanusi, A. M. Sayeed, and B. D. Van Veen, "Canonical space-time processing for wireless communications," *IEEE Transactions on Communications*, vol. 48, pp. 1669–1680, October 2000.

[26] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, pp. 29–48, May 2000.

[27] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Transactions on Information Theory*, vol. 50, pp. 1893–1909, September 2004.

[28] ETSI Normalization Committee, "Broadband radio access network (BRAN); HIPERLAN type 2; system overview," Tech. Rep. TR 101 683 v1.1.1, European Telecommunications Standards Institute, Sophia-Antipolis, France, Feb. 2000.

[29] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation – part I: Coherent detection," *IEEE Transactions on Communications*, vol. 45, pp. 563–572, May 1997.

[30] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation – part II: Differential block detection," *IEEE Transactions on Communications*, vol. 45, pp. 676–686, June 1997.

[31] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation – part III: Co-channel interference," *IEEE Transactions on Communications*, vol. 45, pp. 809–818, July 1997.

[32] J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Transactions on Information Theory*, vol. 44, pp. 1453–1467, July 1998.

[33] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. IEEE Globecom Conference*, (San Francisco, CA), November 2000.

[34] L. Zheng and D. N. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 1073–1096, May 2003.

[35] A. Medles and D. T. M. Slock, "Achieving the optimal diversity-vs-multiplexing tradeoff for MIMO flat channels with QAM space-time spreading and DFE equalization," *IEEE Transactions on Information Theory*. Submitted.

[36] A. Medles and D. T. M. Slock, "Optimal diversity vs. multiplexing tradeoff for frequency selective MIMO channels," in *International Symposium on Information Theory (ISIT '05)*, 2005. Submitted.

[37] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, October 1998.

[38] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, pp. 1456–1467, July 1999.

[39] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc.* 48^{th} *Annual Vehicular Technology Conference*, (Ottawa, Canada), May 1998.

[40] Z. Liu, G. B. Giannakis, S. Barbarossa, and A. Scaglione, "Transmit-antennae space-time block coding for generalized OFDM in the presence of unknown multipath," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1352–1364, July 2001.

[41] H. Bölcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. IEEE Wireless Communications and Networking Conference*, vol. 1, pp. 1–6, September 2000.

[42] P. Laspougeas, "Single and dual multi-sensor channel characterisation – analysis and models," project report, France Télécom R&D, October 2001. IST SATURN Project, D523, Part 2.

[43] H. E. Gamal and A. R. Hammons, Jr., "A new approach to layered space-time coding and signal processing," *IEEE Transactions on Information Theory*, vol. 47, pp. 2321–2334, September 2001.

[44] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM with sub-carrier grouping and constellation precoding," in *Proc. 27^{th} International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. III, pp. 2205–2208, May 2002.

[45] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.

[46] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 451–460, March 1999.

[47] R. R. Müller and W. H. Gerstacker, "On the capacity loss due to separation of detection and decoding," *IEEE Transactions on Information Theory*, vol. 50, pp. 1769–1778, Aug. 2004.

[48] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, pp. 498–519, Feb. 2001.

[49] C. Berrou, C. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo codes," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 1064–1070, 1993.

[50] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, pp. 429–445, Mar. 1996.

[51] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, pp. 284–287, March 1974.

[52] J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft-decision outputs and its applications," in *Proc. Globecom 89*, pp. 47.1.1–47.1.7, November 1989.

[53] ETSI Normalization Committee, "Broadband radio access network (BRAN); HIPERLAN type 2; physical (PHY) layer," Tech. Rep. TS 101 475 v1.3.1, European Telecommunications Standards Institute, Sophia-Antipolis, France, Dec. 2001.

[54] A. Guillén i Fàbregas and G. Caire, "Analysis and design of natural and threaded space-time codes with iterative decoding," in *Proc. 36th Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA, USA), November 2002.

[55] ETSI Normalization Committee, "Channel models for HIPERLAN/2 in different indoor scenarios," Tech. Rep. 3ERI085B, European Telecommunications Standards Institute, Sophia-Antipolis, France, 1998.

[56] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Globecom '98 Conference*, vol. 1, pp. 579–584, Nov. 1998.

[57] H. Vikalo and B. Hassibi, "Low-complexity iterative detection and decoding of multi-antenna systems employing channel and space-time codes," in *Proc. 36th Asilomar Conference on Signals, Systems and Computers*, vol. 1, (Pacific Grove, CA, USA), pp. 294–298, Nov. 2002.

[58] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 1639–1642, July 1999.

[59] G. Lebrun, J. Gao, and M. Faulkner, "MIMO transmission over a time-varying channel using SVD," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 757–764, Mar. 2005.

[60] P. Kyritsi, G. Papanicolaou, P. Eggers, and A. Oprea, "Time reversal techniques for wireless communications," in *Proc. IEEE Vehicular Technology Conference Fall*, vol. 1, pp. 47–51, Sept. 2004.

[61] A. Medles, S. Visuri, and D. T. M. Slock, "On MIMO capacity with partial channel knowledge at the transmitter," in *Proc. 36th Asilomar*

Conference on Signals, Systems and Computers, (Pacific Grove, CA, USA), November 2002.

[62] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Communications Letters*, vol. 6, pp. 322–324, August 2002.

[63] R. W. Heath Jr. and A. Paulraj, "A simple scheme for transmit diversity using partial channel feedback," in *Proc. 32nd Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA, USA), November 1998.

[64] A. Hottinen, O. Tirkkonen, and R. Wichman, *Multi-Antenna Transceiver Techniques for 3G and Beyond*. John Wiley & Sons, 2003.

[65] M. Médard, "The effect upon channel capacity in wireless communication of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, vol. 46, pp. 933–946, May 2000.

[66] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links ?," *IEEE Transactions on Information Theory*, vol. 49, pp. 951–963, Apr. 2003.

[67] B. Friedlander and A. J. Weiss, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. V, pp. 2681–2684, April 1988.

[68] R.-C. Li and G. W. Stewart, "A new relative perturbation theorem for singular subspaces," *Linear Algebra and its Application, Elsevier*, vol. 313, pp. 41–51, July 2000.

[69] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. Chapman & Hall/CRC, 1999.

[70] C. Jötten, P. Baier, M. Meurer, T. Weber, and M. Haardt, "Efficient representation and feedback signaling of channel state information in frequency division duplexing MIMO systems," in *Proc.* 5^{th} *International Symposium on Wireless Personal Multimedia Communications (WPMC), vol. 2*, (Honolulu, HI, USA), pp. 444 – 448, October 2002.

[71] S. Simon and A. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 406–417, Apr. 2003.

[72] G. S. Smith, "A direct derivation of a single-antenna reciprocity relation for the time domain," *IEEE Transactions on Antennas and Propagation*, vol. 52, pp. 1568–1577, June 2004.

[73] A. Bourdoux, B. Come, and N. Khaled, "Non-reciprocal transceivers in OFDM/SDMA systems: Impact and mitigation," in *Proc. IEEE Radio and Wireless Conference (RAWCON)*, (Boston, MA, USA), pp. 183 – 186, August 2003.

[74] V. Jungnickel, V. Krüger, G. Istoc, T. Haustein, and C. von Helmolt, "A MIMO system with reciprocal transceivers for the time-division duplex mode," in *Proc. IEEE Antennas and Propagation Society Symposium*, vol. 2, pp. 1267–1270, June 2004.

[75] S. Van Huffel, *Analysis of the Total Least Squares problem and its use in parameter estimation*. PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium, 1987.

[76] N. Mastronardi, P. Lemmerling, and S. Van Huffel, "Fast structured total least squares algorithm for solving the basic deconvolution problem," *SIAM Journal on Matrix Analysis and Applications*, vol. 22, pp. 533–553, 2000.

[77] Collective, "Eurécom Institute mobile communications dept. activity report 2003," tech. rep., Eurécom Institute, Sophia-Antipolis, France, 2003.

[78] P. Kyritsi, D. C. Cox, R. A. Valenzuela, and P. W. Wolniansky, "Correlation analysis based on MIMO channel measurements in an indoor environment," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 713–720, June 2003.

[79] D. Chizhik, "Slowing the time-fluctuating MIMO channel by beam forming," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1554–1565, Sept. 2004.

[80] T. Svantesson and J. W. Wallace, "Performance evaluation of MIMO channel prediction algorithms using measurements," in *Proc. IFAC SYSID 2003*, (Rotterdam, The Netherlands), August 2003.

[81] B. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *Proc. International Zürich Seminar on Broadband Communications 2002*, (Zürich, Switzerland), February 2002.

[82] J. Winters and J.-K. Hwang, "Sinusoidal modeling and prediction of fast fading processes," in *Proc. IEEE Globecom '98 Conference*, vol. 2, pp. 892–897, 1998.

[83] J. Andersen, J. Jensen, S. Jensen, and F. Frederiksen, "Prediction of future fading based on past measurements," in *Proc. IEEE Vehicular Technology Conference Fall*, vol. 1, pp. 151–155, 1999.

[84] T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *IEEE Communications Letters*, vol. 2, pp. 254–256, September 1998.

[85] M. Chen and M. Viberg, "LMMSE channel prediction based on sinusoidal modeling," in *Proc. Third IEEE Sensor Array and Multichannel Signal Processing (SAM) Workshop*, (Barcelona, Spain), July 2004.

[86] P. Comon and E. Moreau, "Blind MIMO equalization and joint-diagonalization criteria," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, (Salt Lake City, UT, USA), May 2001.

[87] L. Perros-Meilhac, E. Moulines, K. Abed-Meraim, P. Chevalier, and P. Duhamel, "Blind identification of multipath channels: a parametric subspace approach," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1468–1480, July 2001.

[88] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[89] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 1985.

[90] A. Bunse-Gerstner, R. Byers, and V. Mehrmann, "Numerical methods for simultaneous diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 14, no. 4, pp. 927–949, 1993.

[91] J.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 17, pp. 161–164, Jan. 1996.