Thomas Edlich

Recursive Spatial Multiplexing

kassel university press

Thomas Edlich

Recursive Spatial Multiplexing

kassel university press

This work has been accepted by the faculty of Electrical Engineering and Computer Sciences of the University of Kassel acquiring the academic degree of Doktor der Ingenieurwissenschaften (Dr.-Ing.).

Supervisor: Prof. Dr. sc.techn. Dirk Dahlhaus
Co-Supervisor: Dr. sc.techn. Thomas Hunziker

Examiner: Prof. Dr.-Ing. Olaf Stursberg
Prof. Dr.-Ing. Klaus David
Defense day: 27th March 2013

Bibliographic information published by Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.dnb.de.

Zugl.: Kassel, Univ., Diss. 2013
ISBN 978-3-86219-608-1 (print)
ISBN 978-3-86219-609-8 (e-book)
URN: http://nbn-resolving.de/urn:nbn:de:0002-36092

Printing Shop: Print Management Logistics Solutions GmbH & Co. KG, Kassel
Printed in Germany

Meinen Eltern in Dankbarkeit gewidmet

Danksagung

Ich danke meinem Referenten Herrn Prof. Dr. sc.techn. Dirk Dahlhaus für die Möglichkeit, diese Dissertation im Fachgebiet Nachrichtentechnik der Universität Kassel anzufertigen. Herr Prof. Dahlhaus schuf den notwendigen Freiraum zur Erstellung dieser Arbeit und lieferte hilfreiche Hinweise und Anregungen. Außerdem danke ich ihm für die akribische Durchsicht des Manuskripts.

Mein besonderer Dank gilt Herrn Dr. sc.techn. Thomas Hunziker. Die vielen Diskussionen und Ratschläge haben einen wesentlichen Beitrag zu dieser Arbeit geleistet. Für die zahlreichen wertvollen Kommentare zur Erstellung der Dissertation und die Durchsicht der Niederschrift bin ich Herrn Dr. Hunziker sehr dankbar.

Außerdem möchte ich meinen Kollegen im Fachgebiet Nachrichtentechnik danken, besonders Marc Selig, der die vielen Messkampagnen begleitet hat und in Diskussionen wertvolle Hinweise gab. Darüber hinaus danke ich besonders meinen Eltern und meiner Frau Jenny, die durch ihre Rücksichtnahme und ihre grenzenlose Unterstützung diese Arbeit erst ermöglicht haben.

Unterhaching, den 18. Juli 2012

Zusammenfassung

In drahtlosen Kommunikationssystemen ermöglicht der Einsatz von mehreren Sende- und Empfangsantennen (multiple-input multiple-output, MIMO) erhebliche Kapazitätssteigerungen. Letztere erfordern oft aufwändige räumliche Multiplexverfahren in den entsprechenden MIMO-Systemen, die aufgrund beschränkter Komplexität und Batteriekapazität, z.B. in Mobilgeräten, jedoch oft nicht praktikabel sind. Lineare Empfängerstrukturen, wie z. B. der Zero-Forcing-Entzerrer (ZF), bieten eine niedrige Komplexität, erzeugen allerdings gewöhnlich eine Verstärkung des Empfängerrauschens.

In dieser Arbeit wird eine einfache MIMO-Architektur *mit Rückkanal* vorgestellt, deren Empfänger auf einer linearen Signalverarbeitung basiert. Im Gegensatz zu MIMO-Präkodierungsverfahren, bei denen dem Sender Kanalzustandsinformation zur Verfügung steht, wird im vorgeschlagenen Verfahren der Rückkanal dazu benutzt, eine erneute Übertragung von Signalteilen anzufordern, welche in kritischen Unterräumen empfangen wurden. Das angeforderte Stützsignal ermöglicht anschließend dem Empfänger, die durch die lineare Signalaufbereitung generierte Rauschverstärkung zu limitieren. Da das Stützsignal im nachfolgenden Senderahmen eingebettet ist, welcher wiederum ein anderes Stützsignal erfordern könnte, resultiert diese Prozedur in einem rekursiven räumlichen Multiplex-Verfahren (recursive spatial multiplexing, RSM).

Die Leistungsfähigkeit von RSM wird anhand der systembedingten ergodischen Kapazität untersucht und mit derjenigen anderer MIMO-Architekturen mit Rückkanal verglichen. In diesem Zusammenhang werden verschiedene Sendewiederholungsstrategien vorgeschlagen und einander gegenübergestellt. Da in realen Systemen die Datenrate des Rückkanals begrenzt ist, wird die dadurch entstehende Veringerung des Kapazitätsgewinns untersucht. Des Weiteren zeigt sich, dass RSM leicht mit einem MIMO-Präkodierungsverfahren kombiniert werden kann, da sich

die in den beiden Verfahren zum Sender zurück gesendeten Informationen ähneln. Obwohl die Nutzung dieser über den Rückkanal gesendeten Information in beiden Verfahren völlig unterschiedlich geschieht, ergänzen sich diese Verfahren und ermöglichen z. B. einen Anstieg der Kapazität gegenüber reiner Präkodierung in Fällen, wenn die Kanalzustandsinformation veraltet ist.

Ein Schwachpunkt von RSM ist die theoretisch unbegrenzte Verarbeitungsverzögerung, da die lineare Signalaufbereitung solang verschoben wird, bis ein empfangenes Signal *kein* Stützsignal mehr benötigt. Die rekursive Eigenschaft von einer RSM-basierten Sende-/Empfangsstruktur verhindert so, dass eine maximale Verzögerungszeit garantiert werden kann. Durch Ausnutzung der Empfangsdiversität in MIMO-Systemen, in denen die Zahl der Empfangsantennen die Zahl der Sendeantennen übersteigt, kann eine Reduzierung der räumlichen Sendesignaldimensionen (bei gleichbleibender Anzahl der Empfangsantennen) die Häufigkeit von benötigten Sendewiederholungen herabsetzen. Aus diesem Grund werden optimale Strategien zur Sendeantennenwahl vorgestellt, welche auf Algorithmen zum maschinellen Lernen basieren.

Schließlich wird ein RSM-basiertes Übertragungssystem in Kanälen untersucht, welche mit Hilfe eines in dieser Arbeit vorgestellten MIMO-Kanalmessgeräts in verschiedenen Ausbreitungsszenarien gemessen wurden. Die so aufgenommenen Kanalbedingungen werden in numerischen Simulationen genutzt, was den Vorteil bietet, RSM einer einfachen Systemverifikation zu unterziehen, ohne eine echtzeitfähige Implementierung des Verfahrens vornehmen zu müssen.

Abstract

Great capacity gains can be realized in wireless transmissions by using multiple antennas at the transmitter as well as the receiver side. Those multiple-input multiple-output (MIMO) systems require sophisticated spatial multiplexing techniques, while solutions of high complexity are often not feasible, e.g. due to limited battery resources in mobile devices. Linear receiver structures, such as zero-forcing (ZF) schemes, offer low complexity, but usually suffer from excessive noise amplification.

In this thesis, we introduce a simple closed-loop MIMO architecture with linear array signal processing. In contrast to MIMO precoding systems, where channel state information (CSI) is fed back to the transmitter end, the novel architecture uses the feedback channel to request a retransmission of signal parts in critical subspaces. The requested backup subsequently helps the receiver to evade an excessive noise amplification in the linear signal reconstruction. Since the backup signal is embedded in the successive transmission frame, which may require yet another backup, the procedure results in a recursive spatial multiplexing (RSM) scheme.

We investigate the performance of RSM and compare it with other closed-loop MIMO systems by means of the system constrained ergodic capacities. In this context, several retransmission policies are suggested and compared to each other. Since in real transmission systems the feedback channel is rate-limited, the resulting performance degradation is examined. Further, it turns out that RSM can be easily combined with a unitary precoding MIMO system, since the information fed back to the transmitter is similar. Although the exploitation of the feedback information is different, the two systems complement one another and help to increase the system performance when CSI is outdated.

A weakness of RSM is its theoretically unlimited transmission or processing delay, since the linear signal reconstruction at the receiver is postponed until a received transmission frame is *not* in need for a backup

signal. This recursive nature of the RSM transmitter/receiver architecture makes it impossible to guarantee a maximum delay. Exploiting the receive diversity in MIMO systems equipped with more receive antennas than transmit antennas, a reduction of the transmit signal dimensions for a constant number of receive antennas leads to less frequent retransmission requests. For that reason, we introduce optimal transmit antenna selection strategies, which are found by means of machine learning methods.

Finally, an RSM system is applied to real-world transmission channels which have been measured by a MIMO channel sounder introduced in the thesis. The measured MIMO channel data are used in numerical system simulations. This approach enables a simple verification of RSM without the need of a real-time implementation.

Contents

List of Figures

Chapter 1.

Introduction

1.1. Motivation

In the recent two decades the demand for digital communications has increased tremendously. The world wide web has introduced new communication opportunities, such as email or chat, to a broad public. Most information exchange services, e.g. voice over internet protocol (VoIP) or video on demand (VoD), are based on the internet protocol. Social networks enable their users to stay permanently connected, sending messages, location information or status updates. The growing number of pervasive services leads to a correspondingly growing number of mobile communication systems. Simultaneously, the demand for high-rate data services on these mobile platforms increases.

Wireless transmission plays an important role in mobile communication, but often suffers from poor performance due to long and short term channel fading effects. As a result the capacity of the wireless system decreases. Great capacity gains have been realized by using multiple antennas at the transmitter as well as the receiver. Exploiting the capabilities of these so-called multiple-input multiple-output (MIMO) systems requires encoding/decoding methods which are often quite complex. Due to the limited energy in battery driven mobile devices, complex spatial multiplexing techniques are mostly very unattractive solutions.

Channel state information (CSI) at the transmitter end can be used for precoding the multidimensional signal, in order to decompose the MIMO channel into independent single-input single-output (SISO) channels [1], [2]. Each of the resulting signal layers is then affected by an individual SISO channel, which allows the use of off-the-shelf SISO encoders/decoders and enables data rates close to the MIMO channel capac-

ity with a limited processing complexity. This kind of vertical encoding (VE) architecture has the advantage of achieving diversity, which in turn lowers the probability of outages. However, the improvement of this approach is sensitive to inaccuracies of the CSI. Usually, data symbols are transmitted in frames, while the CSI estimated at the receiver end is fed back to the transmitter between two consecutive frame transmissions in a time-division duplex fashion. The system performance suffers significantly if the CSI is outdated due to time-variance of the channel [3], [4] or the CSI is inaccurate due to limited capacity of the feedback channel [5], [6].

Imperfect CSI at the transmitter necessitates additional methods to recover the parallel signal layers at the receiver end. Without any channel knowledge at the transmitter, optimal or near-optimal MIMO decoding becomes a complex task. Solutions of low complexity would be receivers with simple linear processing based on zero-forcing (ZF) or linear minimum mean-squared error (LMMSE) criteria, for instance, to restore the parallel signal layers. However, the linear signal reconstruction at the receiver results in noise amplification over all signal layers in case of a badly conditioned channel matrix [7]. Successive interference cancellation (SIC), e.g. as part of the Vertical Bell Lab Layered Space-Time (V-BLAST) proposal [8], is able to limit this noise amplification, but the necessary successive layer-wise decoding has the drawback of increasing the receiver complexity and introducing a possible error propagation.

In summary, the increasing demands for high data rate services in mobile devices in conjunction with a long battery life call for simple yet efficient transmission architectures. While MIMO channels offer great capacity gains, their exploitation often necessitates encoding/decoding algorithms of high complexity. Closed-loop MIMO systems, which feed back channel state estimates to the transmitter, often suffer from outdated CSI. This thesis introduces a novel approach for a closed-loop MIMO architecture, the performance of which does not depend on accurate CSI. Furthermore, the scheme is based on simple linear signal reconstruction methods and is therefore suitable for the aforementioned battery-constrained receiver architectures.

1.2. State of the Art

Since the first proposals of using multiple antennas at the transmitter as well as the receiver, e.g. [9] or [10], for increasing the capacity of wireless communication systems, many improvements and extensions have been suggested. While fundamental research on the capacity potential has been provided in [11] and [12], several first concrete system architectures have been suggested, as for instance Diagonal-BLAST (D-BLAST) in [13] and V-BLAST in [8]. Many modifications of the BLAST scheme have been proposed thereafter, such as a turbo-like iterative decoding at the receiver [14]. This T-BLAST scheme uses an inter-layer interleaving following the encoder.

Without CSI at the transmitter end the maximum-likelihood (ML) receiver is the optimal decoding method in terms of minimizing the average error probability [15]. Since the ML receiver performs a search through all vector constellations for the most likely sent signal vector, the decoding procedure is of high complexity. Different approaches have been suggested to reduce complexity significantly [16], e.g. sphere decoding [17, 18], in which only a certain part of all possible lattice points are considered in the ML decoder.

Linear signal reconstruction methods, such as ZF and LMMSE, require substantially less complexity, but they suffer from noise amplification over all signal layers [7]. A common solution is the ordered SIC used in V-BLAST [8], but this again leads to an increasing decoder complexity by the successively applied decoding and regenerative interference cancellation for each layer. In [19] the authors suggest a poly-diagonalization for limiting inter-layer interference in order to decrease the noise amplification.

For a perfect diagonalization of the MIMO channel at the receiver, CSI knowledge at the transmitter end can be used. This preprocessing of the transmit signal presumes accurate knowledge of the present channel conditions. Therefore, a precoding of the transmit signal vector is applied [1]. In practice, this assumption is often violated due to time-variance of the channel [3, 4]. Even in a time-invariant environment practical considerations, as for instance a limited capacity of the feedback channel [5, 20], can cause inaccuracies in the precoding matrix and, thus, limit the system performance. Finding appropriate codebooks of precoding matrices

can be represented as the problem of packings in Grassmannian spaces [21], as applied in [22], and bounds of these packings have been analyzed in [23]. Other approaches to reduce feedback information have been examined by interpolating the CSI [24] or by forming a precoding matrix based on means of channel averages and transmit antenna correlation [25].

A very simple closed-loop transmission protocol, the application of which is not restricted to MIMO systems, is given by the family of automatic repeat request (ARQ) architectures, as presented in [26]. The transmit signal is equipped with an error detection code, which enables the receiver to validate the correctness of the received message. In case of an erroneously received signal a retransmission of the original message or redundant information is requested. The latter principle is called a type-II Hybrid-ARQ (HARQ) [27]. The redundancy can be gradually increased [28] or a soft combining of the successively received messages can be applied [29], where the latter is also known as Chase combining. A Chase combining scheme applied to MIMO systems has been examined in [30]. Layer-wise HARQ and a single HARQ in the context of a MIMO transmission have been compared in [31], where the first exhibit a much larger complexity.

During the last decade, the field of *machine learning* has become of great interest for improving communication systems. As machine learning is used in Sect. 4.3 for improving the novel MIMO system approach introduced in this thesis, a brief outline of machine learning applications in MIMO systems is provided in the following. Reinforcement learning is a type of machine learning with the objective of maximizing a cumulative reward. A broad overview on reinforcement learning problems and algorithms has been provided in [32]. Finding optimal strategies for a reinforcement learning problem can be done by using the theory of Markov Decision Processes (MDP) [33]. In [34], the authors suggest an MDP based V-BLAST transmit power and rate control in order to minimize the transmit power constrained on limiting an average transmit delay time, which can be interpreted as a quality-of-service constraint. Bit loading and power allocation per antenna by means of solving an MDP is examined in [35]. While MDPs are suitable for queue management tasks, two-dimensional optimization problems are solved in [36, 37], where a cross-layer approach considers the transmit buffer and the over-

all throughput of a MIMO system. A similar MDP-modeled problem is presented in [38] to find an optimal precoding. Reinforcement learning can also be used in the context of MIMO link adaptation [39].

The evaluation of new algorithms and system architectures is often performed by means of numerical simulations. Since simulations usually employ a relatively simple model of the considered transmission channels, the achieved system performance results depend critically on the accuracy of the channel models. The weakness of such an approach is the assumption of a correctly modelled environment. In order to improve the accuracy of the channel models used in simulations, parts of the new transmission schemes can be implemented in experimental testbeds or hardware modules, which enable a verification in "real-world" scenarios. Clearly, depending on the constraints of the testbed the achievable accuracy in turn may be limited. For instance, in [40] the authors introduce 4×4 MIMO hardware that operates basically in baseband. It is realized by a combination of a digital signal processor (DSP) and a field programmable gate array (FPGA), which enables real-time signal processing, such as an artificial noise generation. MATLAB is a very common solution employed as user interface of verification systems for test algorithm implementations, as applied in the 4×4 MIMO systems suggested in [41, 42], both operating at a center frequency of $2.4 - 2.5$ GHz, but with different bandwidth restrictions. Specialized multi-user MIMO architectures with orthogonal frequency-division multiplexing (OFDM), mostly used for Long Term Evolution (LTE) verification tests, are shown in [43, 44, 45]. While the aforementioned testbeds basically focus on real-time operation, channel sounders provide a more general approach in terms of testing an architecture under realistic transmission conditions. In [46] a 4×4 MIMO channel sounder based on two off-the-shelf network analyzers is used. This set-up usually provides a large sounding bandwidth, but lacks sufficient resolution in time domain to capture time-variant channels. A widely-used channel sounder is based on the MEDAV RUSK system. Early *virtual* MIMO measurements have been performed by means of mechanical array positioning systems [47], with successively sounding the channel between only one transmit and receive antenna pair, where the location of the antennas is changed after each measurement. Nowadays, MEDAV RUSK is equipped with a very fast switching antenna array for sounding and modeling MIMO channels [48, 49]. However, it

has been shown in [50] that this time-division multiplexed switching of a single transmit/receive chain within the elements of a transmit/receive antenna array can cause measurement errors potentially leading to an overestimation of the channel capacity.

1.3. Contribution of this Thesis

The content of this thesis is focused on the introduction and analysis of a novel spatial multiplexing system approach, its application and several practical considerations regarding the architecture. As a performance criterion the system constrained MIMO capacity is used and compared with related transmission architectures. The main contribution of this thesis can be summarized as follows.

- In Sect. 2.4.3 a novel recursive spatial multiplexing (RSM) scheme of low complexity is proposed.
- The performance of RSM in terms of the constrained ergodic capacity is examined. The performance degradation when using a feedback channel with a limited capacity is shown in Sect. 3.3.
- The ease of simultaneously exploiting the advantages of unitary precoding and RSM in one system without further effort is demonstrated in Sect. 3.5.
- Since the RSM approach without further measures can theoretically result in an unlimited recursion process, a system extension to limit the maximum transmit delay is proposed in Sect. 4.3. Antenna selection strategies are derived by means of machine learning, which enables the MIMO system to adapt to changing channel conditions, e.g. transmit or receive antenna correlation.
- In Sect. 5.1 a flexible MIMO channel sounder is presented which is used to characterize real-world transmission channels. In contrast to other testbeds, this measurement system uses the same local oscillator frequency at each transmit and each receive antenna, respectively, but a time-division demultiplexing at only the transmitter side. The MIMO receiver is able to simultaneously measure the signals at the separate receive antennas. This reduces measurement

errors caused by phase noise considerations, as they may occur in virtual or switching MIMO systems.

- The performance indices of RSM derived in previous sections and based on simplified channel models are verified by measurements of system implementations in real-world scenarios using the MIMO channel sounder.

1.4. Outline of this Thesis

Chapter 2 provides an insight into the considered MIMO channel, vertical encoding MIMO and some linear signal reconstruction methods. Several existing closed-loop MIMO systems are reviewed and a detailed consideration of the RSM approach is presented, where the overall transceiver architecture is shown and the optimal feedback information is derived.

Chapter 3 deals with the ergodic performance in terms of the system throughput per channel use. Different RSM policies are introduced and their performance is compared with the constrained capacity of other closed-loop MIMO systems. The practical case of a feedback channel with limited data rate is considered and the inherent throughput degradation is examined. It is further shown that RSM can be even applied as an extension in a precoding MIMO system, where the advantages of both systems can be combined with each other. A theoretical treatment of RSM for large numbers of transmit/receive antennas is presented at the end of this chapter.

Chapter 4 addresses the throughput performance of RSM on condition that the maximum transmit delay is limited, i.e., the recursion depth is bounded. Transmission strategies are derived by means of MDP to minimize the number of outages. A simple machine learning algorithm is discussed which enables the RSM architecture to find adaptive antenna selection strategies in slowly time-varying MIMO channel conditions.

Chapter 5 considers some practical aspects of the RSM scheme. A new MIMO channel sounder is presented. Several transmission scenarios are described and the measurement results are briefly characterized. Finally, the different aforementioned RSM schemes, e.g. with/without precoding or limited maximum delay, are tested using the MIMO channels being measured by the channel sounder in the corresponding scenarios.

Chapter 6 contains the main conclusions and provides an outlook to possible future research topics in the area of RSM.

Throughout the thesis matrices are represented by bold capital letters, while vectors are written as small letters in boldface. $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real and imaginary part, respectively, of a complex value and $(\cdot)^*$ the complex conjugation. $E[\cdot]$ is the expectation of a random variable. The Hermitian transpose of $\mathbf{A}$ is written as $\mathbf{A}^\dagger$ and the Moore-Penrose pseudo-inverse of $\mathbf{A}$ as $\mathbf{A}^+$. $\mathrm{tr}\,(\mathbf{A})$ denotes the trace of the square matrix $\mathbf{A}$ and $\mathrm{diag}\,\{\mathbf{a}\}$ is a diagonal matrix with the elements of the vector $\mathbf{a}$ on its main diagonal. $[\mathbf{A}]_{i,j}$ defines the jth element of the ith row of $\mathbf{A}$ and $\mathbf{I}_a$ is the $a \times a$-identity matrix. The Frobenius norm of $\mathbf{A}$ is denoted as $\|\mathbf{A}\|_{\mathrm{F}}$.

Chapter 2.

MIMO Baseband Architectures

This chapter describes the proposed MIMO transceiver architecture. In general, MIMO transmission systems offer great capacity gains for reliable communication over fading channels. In the latter, the different spatial propagation paths introduce diversity due to independent signal components in the dimensions of the signal space. Encoding/Decoding methods of low complexity enable an orthogonalization of the signal components, i.e., the MIMO signaling is equivalent to a transmission over multiple spatial data streams in parallel.

Clearly, fading affects certain parts of the orthogonal signal dimensions, such that these are in bad condition. In closed-loop architectures, the receiver feeds back data to the transmitter, in particular channel state information (CSI). At the transmitter, exploiting CSI and appropriately precoding (cf. Sect. 2.4.1) the transmit data is a common method to cope with fading. However, this requires the CSI being up-to-date. In the case of imperfect or outdated CSI at the transmitter, further processing is necessary to recover the parallel signal layers at the receiver.

We propose a novel closed-loop MIMO transceiver architecture in which a recursive multiplexing of spatial signal components is applied at the receiver. In contrast to a (CSI based) precoding at the transmitter, the fed back information in the recursive spatial multiplexing (RSM) approach cannot be outdated, as described in the following. A block diagram shown in Fig. 2.1 illustrates the proposed architecture.

The novel system comprises a demultiplexer at the transmitter side and a multiplexer at the receiver side. At the transmitter, the serial input data stream is demultiplexed into array signal frames, while these are multiplexed back to a serial output data stream at the receiver. This corresponds to a vertical encoded MIMO (cf. Sect. 2.2) architecture. Thus,

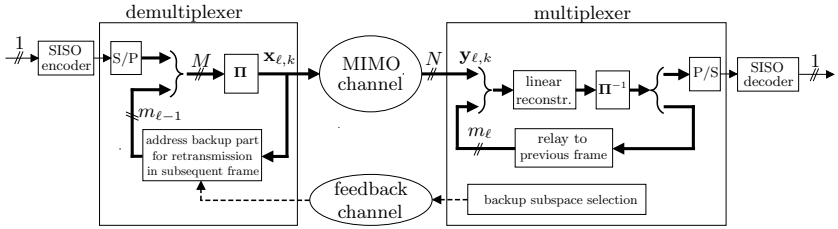

Figure 2.1 Basic principle of a MIMO architecture with recursive spatial multiplexing sketched as a block diagram.

the outer data interfaces support the transmission of a serial data stream, but the inner transmission architecture employs the great benefit of a MIMO communication system. For limiting the complexity of the receiver, a linear signal reconstruction is employed.

We consider an $N \times M$ MIMO channel, with $N \geq M$, for transmitting information data, where M denotes the number of transmit antennas and N the number of receive antennas. Throughout the thesis, we restrict the number of transmit antennas not to exceed the number of receive antennas, such that $M \leq N$, in order to be able to fully restore all transmitted parallel data streams. A total of L frames transmitted sequentially in time are indexed by ℓ with $\ell \in \{1, \ldots, L\}$. Each frame incorporates K signal vectors, such that the $M \times 1$ vectors $\mathbf{x}_{\ell,1}, \ldots, \mathbf{x}_{\ell,K}$ represent the array signal transmitted in the ℓth frame. The received signal is denoted as $\mathbf{y}_{\ell,k}$. A feedback channel is used to request m_ℓ critical signal parts for retransmission in the subsequent frame.

In the following, the individual parts of the considered MIMO architecture, as sketched in Fig. 2.1, are described in detail.

2.1. MIMO Channel

Before describing the inner system architecture, a brief look at the transmission channel will facilitate the understanding of the new approach. A point-to-point transmission over a narrowband $N \times M$ MIMO propagation channel is considered. We employ a block-fading propagation channel model, i.e., the propagation channel is stationary over the period

of a frame transmission, but it may vary arbitrarily from frame to frame. The $N \times M$ matrix $\mathbf{H}_\ell$ shown in Fig. 2.2 contains the narrowband MIMO propagation channel coefficients in the complex baseband describing the ℓth frame.

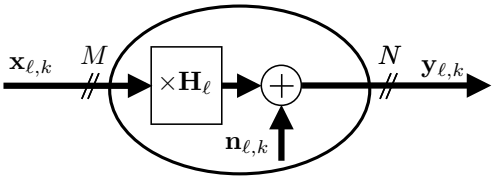

Figure 2.2 Considered MIMO channel.

The output of the propagation channel is superimposed by NK independent front end noise terms, that is, $\mathbf{n}_{\ell,1}, \ldots, \mathbf{n}_{\ell,K}$ are zero-mean complex Gaussian random vectors with covariance matrix $E\left[\mathbf{n}_{\ell,k}\mathbf{n}_{\ell,k}^\dagger\right] = \vartheta_{\text{noise}}\mathbf{I}_N$, with $\mathbf{I}_N$ and ϑ_{noise} denoting the $N \times N$ identity matrix and the noise power spectral density, resp., such that the receiver observes

$$\mathbf{y}_{\ell,k} = \mathbf{H}_\ell \mathbf{x}_{\ell,k} + \mathbf{n}_{\ell,k}, \quad k = 1, \ldots, K. \tag{2.1}$$

The transmit signal vector $\mathbf{x}_{\ell,k}$ is assumed to comprise uncorrelated zero-mean complex random variables with variance ε_{sig}, thus

$$E\left[\mathbf{x}_{\ell,k}\mathbf{x}_{\ell,k}^\dagger\right] = \varepsilon_{\text{sig}}\mathbf{I}_M. \tag{2.2}$$

The signal-to-noise ratio (SNR) at the input of the receiver is defined as $\gamma_{\text{inp}} = \varepsilon_{\text{sig}}/\vartheta_{\text{noise}}$. In addition to the MIMO channel, the existence of a feedback channel is assumed to enable the receiver signaling requests at the transmitter. First, the feedback channel is considered to be perfect with unlimited capacity. In Sect. 3.3, the effect of a limited feedback channel is considered.

Using a singular value decomposition, $\mathbf{H}_\ell$ can be represented as

$$\mathbf{H}_\ell = \mathbf{U}_\ell \mathbf{S}_\ell \mathbf{V}_\ell^\dagger, \tag{2.3}$$

where $\mathbf{S}_\ell$ is an $N \times M$ matrix containing the non-negative real singular

values $\lambda_i(\mathbf{H}_\ell)$ of $\mathbf{H}_\ell$ on its main diagonal and all other elements are zero. Clearly, in the case $N = M$, $\mathbf{S}_\ell$ is a diagonal matrix. For arbitrary values M and N, $\mathbf{U}_\ell$ and $\mathbf{V}_\ell$ are unitary matrices of corresponding dimensions and thus $\mathbf{U}_\ell^\dagger \mathbf{U}_\ell = \mathbf{U}_\ell \mathbf{U}_\ell^\dagger = \mathbf{I}_N$ and $\mathbf{V}_\ell^\dagger \mathbf{V}_\ell = \mathbf{V}_\ell \mathbf{V}_\ell^\dagger = \mathbf{I}_M$. The singular value decomposition in (2.3) can be used to express the eigendecomposition of $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$, such that

$$\mathbf{H}_\ell^\dagger \mathbf{H}_\ell = \left(\mathbf{U}_\ell \mathbf{S}_\ell \mathbf{V}_\ell^\dagger\right)^\dagger \mathbf{U}_\ell \mathbf{S}_\ell \mathbf{V}_\ell^\dagger = \mathbf{V}_\ell \mathbf{S}_\ell \mathbf{U}_\ell^\dagger \mathbf{U}_\ell \mathbf{S}_\ell \mathbf{V}_\ell^\dagger = \mathbf{V}_\ell \mathbf{S}_\ell^2 \mathbf{V}_\ell^\dagger. \quad (2.4)$$

For $\mathbf{H}_\ell$ having full rank M, the matrix $\mathbf{S}_\ell^2 = \operatorname{diag}\left\{\lambda_1^2(\mathbf{H}_\ell), \lambda_2^2(\mathbf{H}_\ell), \ldots, \lambda_M^2(\mathbf{H}_\ell)\right\}$ contains the strictly positive eigenvalues of $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ in ascending order, i.e., $0 < \lambda_1^2(\mathbf{H}_\ell) \leq \ldots \leq \lambda_M^2(\mathbf{H}_\ell)$. To each eigenvalue $\lambda_i^2(\mathbf{H}_\ell)$, with $i \in \{1, \ldots, M\}$, the according eigenvectors constitute the columns of the unitary matrix $\mathbf{V}_\ell$.

Considering the fluctuation of the received signal power caused by the MIMO channel, the channel is said to be in a fade if the power drops significantly [7]. The diversity in MIMO channels can help to cope with fading by means of proper space-time encoding/decoding. The maximum achievable diversity order equals the product MN of transmit and receive antennas, if the attenuations of each transmit-receive antenna pair occur independently [7].

2.2. Vertical Encoding

The information data transmitted by the communication system depicted in Fig. 2.1 is represented as a serial bit sequence. Usually, SISO transmission systems employ forward error correction by convolutional encoding to improve the robustness of the transmission. Temporal interleaving of the encoded data stream usually copes with the effect of burst errors and improves the error performance of the convolutional decoder. The interleaved encoded bit sequence is then mapped to complex symbols according to the applied modulation scheme, e.g. binary phase-shift keying (BPSK) or quadrature amplitude modulation (QAM).

This signal processing of a serial bit stream is depicted as the SISO encoder in Fig. 2.3. The output of the symbol mapper is demultiplexed into M signal streams, which are then transmitted over M antennas.

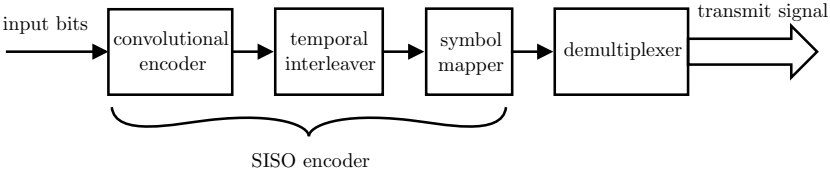

Figure 2.3 Vertical encoding MIMO architecture.

This structure is called vertical encoding (VE) and also employed in V-BLAST [8]. For simplification purposes, the convolutional encoder, the temporal interleaver and the symbol mapper in Fig. 2.3 are represented as the SISO encoder in Fig. 2.1 and in subsequent block diagrams.

The simplest variant of the demultiplexer is the application of a serial-to-parallel converter. Each information bit is spread over different complex symbols due to the convolutional encoding, thus, it is also spread over different transmit antennas by using a simple demultiplexing scheme as the serial-to-parallel converter[1]. In view of the latter, a diversity order greater than N can be achieved. It is also possible to transmit a linear combination of M different complex symbols at each antenna. Hence, in both cases, one information bit is spread over *all* transmit antennas, which results in VE potentially achieving the maximum diversity order of MN [7].

While VE benefits from the comparatively simple transmitter structure, the corresponding receiver architecture can be very complex, since it is required to decode all substreams jointly. Linear signal reconstruction methods are introduced in Sect. 2.3 offering low complexity, but suffering from noise amplification. Successive interference cancellation, as, for instance, part of the V-BLAST proposal [13], can limit noise enhancement at the expense of increasing system complexity due to the necessary successive layer-wise decoding and a potential error propagation. Therefore, the V-BLAST receiver structure will not be considered in this thesis.

[1] It is assumed that the minimum free distance of the convolutional code [51] is sufficiently large in comparison to the number of transmit antennas. Thus, each information bit is spread across all transmit antennas.

2.3. Linear Signal Reconstruction

An observed signal frame at the receiver input consists of K vectors of dimension $N \times 1$ each. The N elements of one of these vectors are superpositions of the different signals sent by each transmit antenna. For recovering the transmitted signal, *linear* estimation methods are applied to the received signal, in order to keep the receiver complexity low. To this end, a matrix $\mathbf{G}_\ell$ is used to estimate the transmit signal, such that

$$\widehat{\mathbf{x}}_{\ell,k} = \mathbf{G}_\ell \mathbf{y}_{\ell,k}. \tag{2.5}$$

VE requires the joint decoding of *all* substreams at the receiver. If some signal layers are of insufficient quality, the linear signal reconstruction may fail, as shown below. However, the reviewed algorithms are of low complexity. In the following, perfect knowledge of the CSI at the receiver only is assumed. Furthermore, at first the transmitter is considered to have no CSI.

2.3.1. Zero Forcing

In SISO systems, a ZF filter is usually applied to the received signal for the purpose of reducing intersymbol interference (ISI), e.g. in [51]. Here, the effect of ISI is neglected, since narrowband channels are considered only, i.e., the channel is nondistorting due to constant amplitude response and envelope delay characteristics of the propagation channel. Although no interference of consecutively transmitted symbols is present, ZF is used in a MIMO receiver to overcome the multistream interference (MSI) resulting from interfering substreams transmitted at the same time.

The simple idea of the ZF algorithm for decomposing the sublayers in a MIMO receiver is based on equalizing the linear mapping caused by the channel. This is done by applying a filter to the received signal, with the property to invert the channel, as described in [7]. For $N = M$ where $\mathbf{H}_\ell$ is a square matrix, the ZF algorithm results in multiplying the received signal by ${\mathbf{H}_\ell}^{-1}$. For $N > M$, the Moore-Penrose pseudoinverse [52] $\mathbf{H}_\ell^+ = \left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1} \mathbf{H}_\ell^\dagger$ of $\mathbf{H}_\ell$ is used. Hence, the equalized signal is given by

$$\widehat{\mathbf{x}}_{\ell,k}^{(\mathrm{ZF})} = \mathbf{H}_\ell^+ \mathbf{y}_{\ell,k} \tag{2.6}$$

with $\mathbf{G}_\ell = \mathbf{H}_\ell^+$. Using (2.1) and assuming $\mathbf{H}_\ell$ has full rank, the output of the ZF equalizer is given by

$$\widehat{\mathbf{x}}_{\ell,k}^{(\mathrm{ZF})} = \mathbf{x}_{\ell,k} + \mathbf{H}_\ell^+ \mathbf{n}_{\ell,k}. \tag{2.7}$$

As expected, the ZF decouples the transmitted substreams to parallel channels with additive noise. However, it can be seen from (2.7) that the noise is enhanced by the ZF operator $\mathbf{H}_\ell^+$. Quantifying this noise enhancement of the ZF can be done by calculating the postprocessing noise power matrix

$$\begin{aligned}
\mathbf{\Theta}_{\mathrm{noise}}^{(\mathrm{ZF})} &= E\left[\mathbf{H}_\ell^+ \mathbf{n}_{\ell,k} \left(\mathbf{H}_\ell^+ \mathbf{n}_{\ell,k}\right)^\dagger\right] \\
&= E\left[\left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1} \mathbf{H}_\ell^\dagger \mathbf{n}_{\ell,k} \mathbf{n}_{\ell,k}^\dagger \mathbf{H}_\ell \left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1}\right] \\
&= \left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1} \mathbf{H}_\ell^\dagger E\left[\mathbf{n}_{\ell,k} \mathbf{n}_{\ell,k}^\dagger\right] \mathbf{H}_\ell \left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1} \\
&= \vartheta_{\mathrm{noise}} \left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1}.
\end{aligned} \tag{2.8}$$

By comparison of the preprocessing noise covariance matrix $\vartheta_{\mathrm{noise}}\mathbf{I}_N$ and $\mathbf{\Theta}_{\mathrm{noise}}^{(\mathrm{ZF})}$, it is obvious that the postprocessing noise is correlated across the different substreams due to the ZF equalization. The SNR of each substream is calculated by considering the elements on the main diagonal of $\mathbf{\Theta}_{\mathrm{noise}}^{(\mathrm{ZF})}$, such that the ZF constrained SNR of the mth substream in the ℓth frame results to

$$\gamma_{\ell,m}^{(\mathrm{ZF})} = \frac{\gamma_{\mathrm{inp}}}{\left[\left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1}\right]_{m,m}}, \tag{2.9}$$

where $[\cdot]_{a,b}$ denotes the bth element of the ath row of a matrix. In summary, the ZF linear signal estimator is able to decompose the inherent M substreams of a received MIMO signal, but at the expense of noise enhancement.

2.3.2. Linear Minimum Mean-Squared Error

The output of a ZF equalizer constitutes perfectly decoupled substreams, but suffers from noise enhancement. An alternative receiver concept for improving the performance of the ZF equalizer, especially in the low SNR range, can be derived by minimizing the mean-squared error between the linearly equalized and transmitted signals. As shown in [7], the LMMSE receiver matrix $\mathbf{G}_\ell^{(\mathrm{LMMSE})}$ has to satisfy the equation

$$\mathbf{G}_\ell^{(\mathrm{LMMSE})} = \arg\min_{\mathbf{G}_\ell} E\left[\left\|\mathbf{G}_\ell \mathbf{y}_{\ell,k} - \mathbf{x}_{\ell,k}\right\|_{\mathrm{F}}^2\right], \tag{2.10}$$

where $\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\mathrm{tr}\left(\mathbf{A}\mathbf{A}^\dagger\right)}$ denotes the Frobenius norm of $\mathbf{A}$. The solution in (2.10) results from the *orthogonality principle* [53]

$$E\left[\left(\mathbf{G}_\ell^{(\mathrm{LMMSE})}\mathbf{y}_{\ell,k} - \mathbf{x}_{\ell,k}\right)\mathbf{y}_{\ell,k}^\dagger\right] = \mathbf{0}_{M,N}, \tag{2.11}$$

stating the estimation error $\left(\mathbf{G}_\ell^{(\mathrm{LMMSE})}\mathbf{y}_{\ell,k} - \mathbf{x}_{\ell,k}\right)$ has to be uncorrelated to all observations $\mathbf{y}_{\ell,k}$ in order to minimize the mean-squared error (MSE) [54]. Referring to [7], $\mathbf{G}_\ell^{(\mathrm{LMMSE})}$ can be derived from (2.11) as

$$\mathbf{G}_\ell^{(\mathrm{LMMSE})} = \left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell + \frac{\mathbf{I}_M}{\gamma_{\mathrm{inp}}}\right)^{-1} \mathbf{H}_\ell^\dagger. \tag{2.12}$$

A simple comparison of the LMMSE and the ZF estimators has been performed in [7] by considering (2.12) for high and low SNR regions. For large SNR values the LMMSE operator converges to the matrix of the ZF, i.e., $\mathbf{G}_\ell^{(\mathrm{LMMSE})} \approx \mathbf{H}_\ell^+$. In contrast, the LMMSE operator approximates a matched filter matrix when considering low SNR values, such that $\mathbf{G}_\ell^{(\mathrm{LMMSE})} \approx \gamma_{\mathrm{inp}}\mathbf{H}_\ell^\dagger$.

The performance of the LMMSE estimator will be characterized in chapter 3 in terms of the signal to interference and noise ratio (SINR) on the mth substream [7]

$$\kappa_{\ell,m} = \frac{1}{\left[\left(\mathbf{I}_M + \gamma_{\mathrm{inp}}\mathbf{H}_\ell^\dagger \mathbf{H}_\ell\right)^{-1}\right]_{m,m}} - 1. \tag{2.13}$$

2.4. Closed Loop MIMO Systems

So far, we have introduced an open-loop MIMO system, i.e., a unidirectional communication without any feedback information sent from the receiver back to the transmitter. A very common method is to feed back CSI to the transmitter. This CSI is then used at the transmitter to calculate an appropriate precoding of the transmit signal according to the current state of the channel. The major problem of this approach concerns the processing delay, i.e., the precoding is useless, if it is out-of-date and does not match the current CSI. Outdated CSI problems usually arise in fast fading channels (e.g. mobile communication), or if the time delay between sensing the channel state at the receiver and the precoding at the transmitter is too long. A second problem in precoding is the limited capacity of the feedback channel, which leads to a limited accuracy of the returned CSI. Furthermore, if the feedback channel is non-perfect, the transmission also suffers from erroneous and thus mismatched precoding.

In a second closed-loop system architecture, only the reception *quality* is returned to the transmitter. The simplest way of indicating the quality of the received message is a binary signaling, as provided by automatic repeat request (ARQ) schemes. Here, the receiver checks if the message has been received without any error. In case the receiver acknowledges the error-free transmission, the transmitter continues to send the subsequent messages. In all other cases, the erroneous message is retransmitted until the correct reception is acknowledged.

Below, unitary precoding and ARQ schemes are discussed in the context of MIMO communication systems. Furthermore, a novel closed-loop MIMO system approach is introduced, where the differences to the existing systems are highlighted.

2.4.1. Unitary Precoding

Unitary precoding (UP) is a preprocessing of the transmit signal without changing the norm of the transmit MIMO signal. The discrete-time baseband signal of the ℓth frame at the MIMO channel input has the form $\mathbf{P}_\ell \mathbf{x}_{\ell,1}, \ldots, \mathbf{P}_\ell \mathbf{x}_{\ell,K}$, as a result of precoding the signals $\mathbf{x}_{\ell,1}, \ldots, \mathbf{x}_{\ell,K}$ by the unitary $M \times M$ matrix $\mathbf{P}_\ell$. Fig. 2.4 shows a block diagram of a MIMO system using UP. The following description assumes a sequential transmission of frames alternating with a feedback of CSI from the receiver to

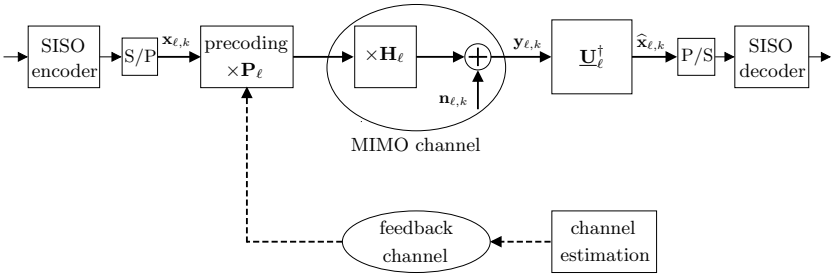

Figure 2.4 Unitary precoding in a MIMO system.

the transmitter in between two consecutively transmitted signal frames.

The aim of the precoding is to facilitate the decomposition of the MIMO channel into independent SISO channels by choosing $\mathbf{P}_\ell$ such that it is in line with the eigensystem of $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$. As it has already been shown in (2.3), the MIMO channel can be represented by $\mathbf{H}_\ell = \mathbf{U}_\ell\mathbf{S}_\ell\mathbf{V}_\ell^\dagger$. The diagonal matrix $\mathbf{S}_\ell$ contains the singular values of $\mathbf{H}_\ell$. Referring to Fig. 2.4 and supposing $\mathbf{P}_\ell$ to be the (yet unknown) precoding matrix, the input signal at the receiver is given as

$$\begin{aligned}\mathbf{y}_{\ell,k} &= \mathbf{H}_\ell\mathbf{P}_\ell\mathbf{x}_{\ell,k} + \mathbf{n}_{\ell,k} \\ &= \mathbf{U}_\ell\mathbf{S}_\ell\mathbf{V}_\ell^\dagger\mathbf{P}_\ell\mathbf{x}_{\ell,k} + \mathbf{n}_{\ell,k}.\end{aligned} \tag{2.14}$$

Since UP is a preprocessing of $\mathbf{x}_{\ell,k}$, a proper diagonalization and thus decomposing of the MIMO channel into independent SISO channels could be done by choosing $\mathbf{P}_\ell = \mathbf{V}_\ell$. A subsequent post-multiplication of $\mathbf{y}_{\ell,k}$ by $\mathbf{U}_\ell^\dagger$ results in a diagonal channel matrix as desired. As indicated in (2.3), the receiver has to determine the SVD of the MIMO channel, which in turn is estimated by means of a training sequence. Since the transmit data, and thus also the training sequence, is pre-processed by the precoding, the receiver estimates $\underline{\mathbf{H}}_\ell = \mathbf{H}_\ell\mathbf{P}_\ell$, i.e., the composite channel comprising both the MIMO channel and the precoding. This limits the receiver complexity, since no knowledge of the actual precoding matrix $\mathbf{P}_\ell$ is required. Similarly to (2.3), the composite channel can be represented

by the singular value decomposition

$$\underline{\mathbf{H}}_\ell = \underline{\mathbf{U}}_\ell \underline{\mathbf{S}}_\ell \underline{\mathbf{V}}_\ell^\dagger. \tag{2.15}$$

The matrix $\underline{\mathbf{V}}_\ell$ is then sent back to the transmitter, where it is used for the preparation of the subsequent frame. Since $\underline{\mathbf{V}}_\ell$ is part of the decomposition of the composite channel $\mathbf{H}_\ell \mathbf{P}_\ell$ and $\underline{\mathbf{V}}_\ell$ is unitary, the unitary matrix for precoding the signals transmitted in the subsequent frame reads

$$\mathbf{P}_{\ell+1} = \mathbf{P}_\ell \underline{\mathbf{V}}_\ell. \tag{2.16}$$

The following example shall illustrate the procedure. If no precoding is considered in the first frame, that is $\mathbf{P}_1 = \mathbf{I}_M$, the composite channel simplifies to $\underline{\mathbf{H}}_1 = \mathbf{H}_1$ and the SVD equals to the expression in (2.15) for $\ell = 1$. Hence, the precoding for the second frame is given by $\mathbf{P}_2 = \mathbf{V}_1 = \underline{\mathbf{V}}_1$. For $\ell = 2$, the SVD calculated by the receiver is based on the estimate of the composite channel $\underline{\mathbf{H}}_2 = \mathbf{H}_2 \mathbf{P}_2 = \mathbf{H}_2 \mathbf{V}_1 = \underline{\mathbf{U}}_2 \underline{\mathbf{S}}_2 \underline{\mathbf{V}}_2^\dagger$. The precoding matrix for the third frame results to $\mathbf{P}_3 = \mathbf{V}_1 \underline{\mathbf{V}}_2$.

To show the diagonalization property of UP, the case of a time-invariant channel is considered, i.e., $\mathbf{H}_{\ell+1} = \mathbf{H}_\ell$. In this case, the composite channel $\underline{\mathbf{H}}_{\ell+1}$ can be expressed as

$$\underline{\mathbf{H}}_{\ell+1} = \mathbf{H}_\ell \mathbf{P}_{\ell+1} = \mathbf{H}_\ell \mathbf{P}_\ell \underline{\mathbf{V}}_\ell = \underline{\mathbf{H}}_\ell \underline{\mathbf{V}}_\ell = \underline{\mathbf{U}}_\ell \underline{\mathbf{S}}_\ell, \tag{2.17}$$

such that $\underline{\mathbf{V}}_{\ell+1} = \mathbf{I}_M$ and a post-multiplication by $\underline{\mathbf{U}}_\ell^\dagger$ results in a diagonal channel matrix. Usually, such a signal reconstruction is impossible when $\mathbf{H}_{\ell+1} \neq \mathbf{H}_\ell$. UP is sometimes called SVD (based) precoding or optimal beamforming [11].

2.4.2. Automatic Repeat Request

When using an ARQ scheme, the receiver either acknowledges the error-free reception or requests a retransmission of the corresponding information symbols. An error detection code can be used to check the correctness of the received signal. Under bad channel conditions, a message would require several retransmissions. To avoid resulting inefficient retransmissions, in HARQ a forward error correction (FEC) code, e.g. Reed-Solomon [55], is applied to the transmit signal. It enables to find

and correct errors, but it also decreases the data rate by adding redundancy. In [26] two types of HARQ are introduced.

Type-I HARQ defines the aforementioned extension of the ARQ protocol by FEC. In cases of sufficient transmission quality, no retransmission is needed. Few errors of the received message can be corrected by the inherent FEC. This prevents the occupation of the channel by a retransmission at the expense of the aforementioned lower data rate, since FEC usually increases the message length by a factor of two or three. In type-II HARQ, the transmitter alternates between sending an uncoded message and providing parity bits (redundancy), if necessary. A very common type-II HARQ has been proposed in [27], where a message is equipped with an error-detection code only. If the message is received without any error, the channel capacity has been exploited efficiently. When the receiver does not acknowledge the error-free reception, the transmitter uses FEC on the original message to generate redundancy. Instead of retransmitting the original message, the redundancy only is transmitted to the receiver. This increases the robustness of the already received first message, which in turn increases the probability of an error-free decoding.

Especially in 3GPP Long Term Evolution, *HARQ with soft combining* is used. The erroneously received message is not discarded when requesting the transmitter for a retransmission. The original message and all retransmissions are then combined at the receiver, e.g. by a maximum-ratio combining scheme, as introduced by Chase in [29]. Different variants of this so-called Chase combining exist. While usually combining after the symbol demodulation, this can can be also performed in symbol domain. Chase combining can be considered as a special case of the principle of *incremental redundancy*, initially suggested by Mandelbaum in [28]. The FEC encoded information is separated into different blocks by puncturing. If the first transmitted block is not able to be decoded error-free, other blocks will support the decoding in succeeding transmissions. When all different groups have been transmitted, but still the correct decoding is impossible, several messages can be repeated. Chase combining is then applied on those groups, as shown in [56].

In MIMO communication links HARQ can be applied to the total transmit frame, i.e., at the SISO decoder in case of a VE, as examined in [30]. Here, two approaches of utilizing Chase combining have been studied, namely carrying out the cumulative combining *before* or *after* the

linear signal reconstruction. It has been shown that the pre-combining is superior to the post-combining scheme. Several MIMO cross-layer designs have been proposed, e.g. extending HARQ with adaptive modulation schemes [57]. In [31], a comparison between MIMO with a single HARQ scheme and a layer-wise HARQ design is drawn. While the throughput of the single HARQ approach is inferior, an HARQ on each layer suffers from its high complexity.

2.4.3. Recursive Spatial Multiplexing

Having introduced the commonly used closed-loop MIMO architectures in the previous section, namely UP and ARQ, RSM as a novel closed-loop system approach is proposed in the following. The description of RSM is divided into two parts. First, the demultiplexing structure at the transmitter is explained, while the receiver part is discussed in the following.

Transmitter

At the transmitter, the SISO encoded and interleaved signal stream is demultiplexed into M parallel signals to match the MIMO signal dimension. The block diagram of the transmitter is shown in Fig. 2.5.

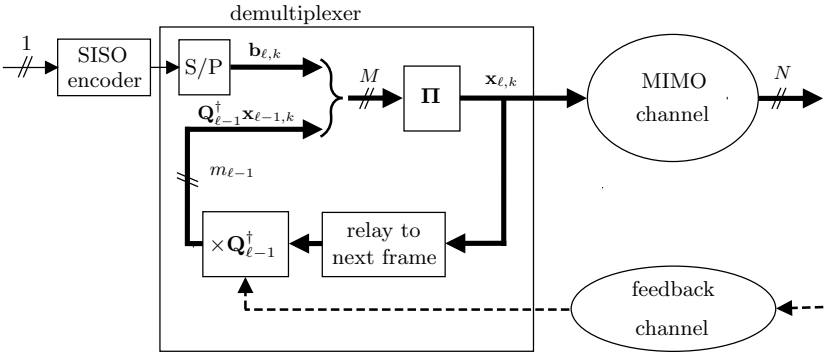

Figure 2.5 Transmitter architecture in RSM.

The serial/parallel converted signal $\mathbf{b}_{\ell,k}$, with $k = 1, \ldots, K$, is the SISO encoded data, which occupy $M - m_{\ell-1}$ of the available M MIMO dimen-

sions (layers). The remaining $m_{\ell-1}$ layers are dedicated to signal parts requested for retransmission by the previous $(\ell-1)$th frame. For this purpose, the transmit message $\mathbf{x}_{\ell,1}, \ldots, \mathbf{x}_{\ell,K}$ is relayed to the time instance (or time slot) of the next frame. The subspace selector $\mathbf{Q}_{\ell-1}$, which is an $M \times m_{\ell-1}$ matrix obtained from the receiver via the feedback channel, identifies a backup signal of the form $\mathbf{Q}_{\ell-1}^{\dagger}\mathbf{x}_{\ell-1,1}, \ldots, \mathbf{Q}_{\ell-1}^{\dagger}\mathbf{x}_{\ell-1,K}$ to be embedded in the next frame. In summary, this subsequent frame then contains

$$(\mathbf{x}_{\ell,1}, \ldots, \mathbf{x}_{\ell,K}) = \mathbf{\Pi}\left(\begin{bmatrix} \mathbf{b}_{\ell,1} \\ \mathbf{Q}_{\ell-1}^{\dagger}\mathbf{x}_{\ell-1,1} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{b}_{\ell,K} \\ \mathbf{Q}_{\ell-1}^{\dagger}\mathbf{x}_{\ell-1,K} \end{bmatrix}\right), \quad (2.18)$$

where $\mathbf{\Pi}$ denotes an inner interleaving of the signals over the M layers, which has the purpose of decorrelating the noise terms after the signal reconstruction at the receiver. The transmit signal $\mathbf{x}_{\ell,k}$ is again assumed to represent uncorrelated zero-mean complex random variables with variance ε_{sig}, as shown in (2.2). This holds true as long as the subspace identifier $\mathbf{Q}_{\ell}$ satisfies $\mathbf{Q}_{\ell}^{\dagger}\mathbf{Q}_{\ell} = \mathbf{I}_{m_\ell}$, i.e., $\mathbf{Q}_{\ell}$ is to be an $M \times m_{\ell}$ matrix with orthonormal columns, and as long as

$$E\left[\mathbf{b}_{\ell,k}\mathbf{b}_{\ell,k}^{\dagger}\right] = \varepsilon_{\text{sig}}\mathbf{I}_{M-m_\ell}, \quad (2.19)$$

which is a common assumption for the output signal of a SISO encoder.

Receiver

The transmit signal passes the narrow band MIMO channel and undergoes MN attenuations and phase shifts. Here, no particular channel model is assumed, since the basic receiver architecture is independent of the underlying channel statistics. MIMO channel models are specified in Chapt. 3 and Sect. 4.3 when examing the performance of the RSM approach.

Throughout this work, a linear signal reconstruction is assumed to be applied at the receiver in order to limit the receiver complexity (cf. Sect. 2.3). Hence, a linear estimator $\mathbf{G}_{\ell}$ can be formulated for the transmitted array signal, as it can be seen in Fig. 2.6. However, the linear processing results in noise enhancement over all parallel M layers in cases of badly conditioned channel matrices.

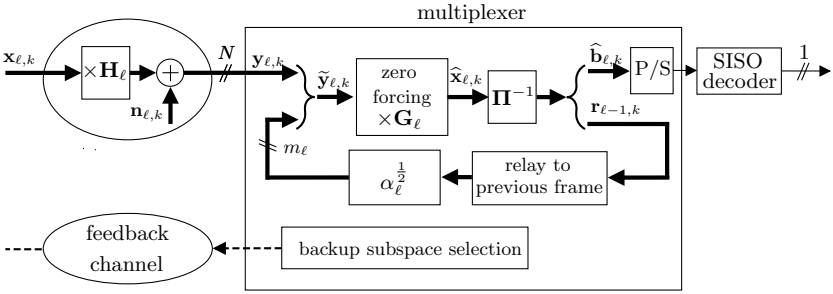

Figure 2.6 Receiver structure of the RSM approach.

We assume the availability of m_ℓ-dimensional backup vector signals $\mathbf{r}_{\ell,1},\ldots,\mathbf{r}_{\ell,K}$, where $\mathbf{r}_{\ell,k}$ has the form

$$\mathbf{r}_{\ell,k} = \mathbf{Q}_\ell^\dagger \mathbf{x}_{\ell,k} + \mathbf{n'}_{\ell,k} \tag{2.20}$$

where $\mathbf{n'}_{\ell,k}$ is a zero-mean random vector, which denotes the inherent noise of the backup signal whith $E\left[\mathbf{n'}_{\ell,k}\mathbf{n'}^\dagger_{\ell,k}\right] = \vartheta^{(\ell)}_{\text{backup}}\mathbf{I}_{m_\ell}$. The components in $\mathbf{n'}_{\ell,k}$ are considered uncorrelated, due to the previously processed deinterleaving of the signals over the M layers. The input signal to the zero forcer is composed by the observed signal $\mathbf{y}_{\ell,k}$ and the scaled version of the undecodable signal layers of the previous frame $\mathbf{r}_{\ell,k}$ according to

$$\widetilde{\mathbf{y}}_{\ell,k} = \widetilde{\mathbf{H}}_\ell \mathbf{x}_{\ell,k} + \widetilde{\mathbf{n}}_{\ell,k} \tag{2.21}$$

$$\text{with}\quad \widetilde{\mathbf{y}}_{\ell,k} = \begin{bmatrix} \mathbf{y}_{\ell,k} \\ \alpha_\ell^{\frac{1}{2}}\mathbf{r}_{\ell,k} \end{bmatrix},\quad \widetilde{\mathbf{H}}_\ell = \begin{bmatrix} \mathbf{H}_\ell \\ \alpha_\ell^{\frac{1}{2}}\mathbf{Q}_\ell^\dagger \end{bmatrix},\quad \widetilde{\mathbf{n}}_{\ell,k} = \begin{bmatrix} \mathbf{n}_{\ell,k} \\ \alpha_\ell^{\frac{1}{2}}\mathbf{n'}_{\ell,k} \end{bmatrix},$$

and α_ℓ a scaling factor as described in the following. The backup signal $\mathbf{r}_{\ell,k}$ is the result of a retransmitted signal in a *subsequent* frame and this is collocated with the received signal $\mathbf{y}_{\ell,k}$ of the *active* frame. Since we employ a block-fading propagation channel model (cf. Sect. 2.1), the propagation channel may vary arbitrarily from frame to frame and, thus, the variances of the noise vector elements in $\mathbf{n}_{\ell,k}$ and $\mathbf{n'}_{\ell,k}$ may differ.

Therefore, the vector $\mathbf{r}_{\ell,k}$ is scaled by a factor

$$\alpha_{\ell}^{\frac{1}{2}} = \sqrt{\frac{\vartheta_{\text{noise}}}{\vartheta_{\text{backup}}^{(\ell)}}}\,, \tag{2.22}$$

where the scaling by $\alpha_{\ell}^{\frac{1}{2}}$ leads to $E\left[\widetilde{\mathbf{n}}_{\ell,k}\widetilde{\mathbf{n}}_{\ell,k}^{\dagger}\right] = \vartheta_{\text{noise}}\mathbf{I}_{N+m_\ell}$.

The zero forcing matrix $\mathbf{G}_\ell$ can be derived on the basis of (2.21), assuming knowledge of $\widetilde{\mathbf{H}}_\ell$ at the receiver end. The output of the linear signal reconstruction is a signal vector of the form

$$\widehat{\mathbf{x}}_{\ell,k} = \mathbf{G}_\ell\widetilde{\mathbf{y}}_{\ell,k} = \mathbf{\Pi}\left(\begin{bmatrix} \widehat{\mathbf{b}}_{\ell,k} \\ \mathbf{r}_{\ell-1,k} \end{bmatrix}\right) = \mathbf{\Pi}\left(\begin{bmatrix} \mathbf{b}_{\ell,k} \\ \mathbf{Q}_{\ell-1}^{\dagger}\mathbf{x}_{\ell-1,k} \end{bmatrix}\right) + \mathbf{w}_{\ell,k}, \tag{2.23}$$

where the vector $\mathbf{w}_{\ell,k} = \widehat{\mathbf{x}}_{\ell,k} - \mathbf{x}_{\ell,k}$ represents the error signal vector. The reconstructed signal $\widehat{\mathbf{x}}_{\ell,k}$ is an $M \times 1$ vector, in which the upper $M - m_{\ell-1}$ elements, namely $\widehat{\mathbf{b}}_{\ell,k}$, are parallel/serial converted and sent to the SISO decoder. The remaining lower $m_{\ell-1}$ elements compose $\mathbf{r}_{\ell-1,k}$, i.e., the backup for the preceding frame.

The recursive structure of the receiver is summarized as follows. When a frame of signals is received, it is stored in a last-in first-out (LIFO) memory. At the same time, the condition of the current MIMO channel is estimated, e.g. by means of a training sequence which could be a part of $\mathbf{y}_{\ell,1}$. In the case of bad or insufficient channel conditions, the linear signal reconstruction is postponed unless a backup for the critical signal parts is available. Then, the receiver requests a retransmission of the signals at these critical layers by feeding back an identifier $\mathbf{Q}_\ell$ to address the requested signal part $\mathbf{Q}_\ell^{\dagger}\mathbf{x}_{\ell,k}$, for $k = 1, \ldots, K$. Fig. 2.7 demonstrates these details, which are only briefly shown in Fig. 2.6. If the received frame does not require a backup signal, i.e., $m_\ell = 0$, all frames stored in the LIFO are processed by the RSM decoder in reverse order. That is, the zero forcing is applied to the last frame stored in the LIFO memory first. As shown in Fig. 2.7, we introduce a delay of T_{LIFO}, defined as the RSM processing time for each frame contained in the LIFO. We assume that the time for processing the full LIFO content is less than the period of a transmitted frame.

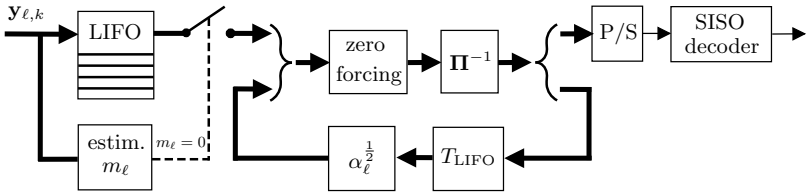

Figure 2.7 Simplified RSM receiver structure.

At the receiver, the multiplexing and decoding need to be delayed as long as the channel conditions do not meet the requirements. That is, the frame containing a backup signal may also necessitate a retransmission of certain signal parts from the following frame and so on. The RSM completes once a frame is received, which can be processed without the need for a backup. Fig. 2.8 points out the RSM operation principle.

The left hand side depicts the transmission and retransmission structure, where each transmission comprises the framing at the leftmost box, the interleaving at the middle and sending the frame over the channel on

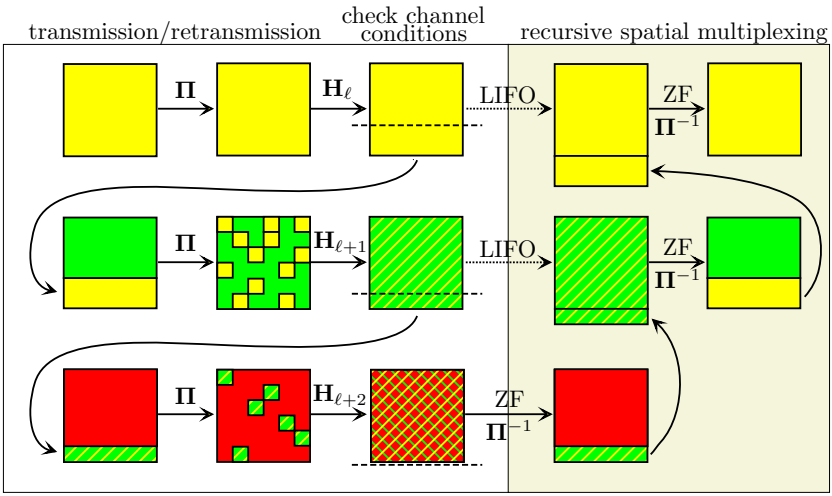

Figure 2.8 Operation principle of RSM.

the right. Frame 2 and 3 each contains a part of the previously transmitted signal as a backup, since certain parts of the preceding frame have been received with insufficient quality. The quality of the signals can be checked, for instance, by comparing the channel state with a threshold, marked as a dashed line here. The signal reconstruction and decoding start as soon as a complete signal vector is above that threshold, i.e., the channel conditions meet the requirements, as shown in the third frame. Signal reconstruction (shown on the gray shaded right hand side at Fig. 2.8) is then performed in reverse frame order, using the stored signal frame from the LIFO and the backup signal provided by the subsequent frame.

Backup Addressing

It is yet to be discussed how to address the backup signal to be transmitted in the subsequent frame. By providing a *partial* retransmission of the transmitted signal the receiver is able to decode the message of the previous frame while receiving new data in the current frame. For the $(\ell+1)$th frame at the transmitter, a linear combination of the signal vectors of the ℓth transmitted frame is created and thereafter combined with the new message $\mathbf{b}_{\ell+1,k}$ of the $(\ell+1)$th frame (cf. (2.23)).

The objective of the backup procedure is to provide signals with a sufficient SNR for subsequent decoding. In the following, we will focus on a ZF based signal reconstruction. The noise vector $\mathbf{w}_{\ell,k}$ observable at the output of the signal reconstruction is defined as

$$\mathbf{w}_{\ell,k} = \mathbf{G}_\ell \widetilde{\mathbf{y}}_{\ell,k} - \mathbf{x}_{\ell,k}. \tag{2.24}$$

In the following, it is assumed that the channel estimation at the receiver is perfect such that $E[\mathbf{w}_{\ell,k}] = 0$. According to (2.8), the postprocessing noise power matrix at the output of the RSM system results to

$$E\left[\mathbf{w}_{\ell,k}\mathbf{w}_{\ell,k}^\dagger\right] = E\left[\mathbf{G}_\ell \widetilde{\mathbf{n}}_{\ell,k}\widetilde{\mathbf{n}}_{\ell,k}^\dagger \mathbf{G}_\ell^\dagger\right] = \vartheta_{\text{noise}}\mathbf{G}_\ell\mathbf{G}_\ell^\dagger = \vartheta_{\text{noise}}\left(\widetilde{\mathbf{H}}_\ell^\dagger\widetilde{\mathbf{H}}_\ell\right)^{-1}, \tag{2.25}$$

with the zero forcing matrix $\mathbf{G}_\ell = \left(\widetilde{\mathbf{H}}_\ell^\dagger\widetilde{\mathbf{H}}_\ell\right)^{-1}\widetilde{\mathbf{H}}_\ell^\dagger$ according to (2.6). The

MSE of the estimated signal $\mathbf{G}_\ell \widetilde{\mathbf{y}}_{\ell,k}$ is defined by

$$\begin{aligned}\mu(\mathbf{Q}_\ell) &= \frac{1}{M} E\left[\|\mathbf{w}_{\ell,k}\|_{\mathrm{F}}^2\right] \\ &= \frac{\vartheta_{\mathrm{noise}}}{M} \mathrm{tr}\left(\left(\widetilde{\mathbf{H}}_\ell^\dagger \widetilde{\mathbf{H}}_\ell\right)^{-1}\right) \\ &= \frac{\vartheta_{\mathrm{noise}}}{M} \mathrm{tr}\left(\left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell + \alpha_\ell \mathbf{Q}_\ell \mathbf{Q}_\ell^\dagger\right)^{-1}\right). \end{aligned} \tag{2.26}$$

According to (2.4), the eigenvalue decomposition of the Hermitian matrix $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ can be represented by $\mathbf{V}_\ell \mathbf{S}_\ell^2 \mathbf{V}_\ell^\dagger$.

Proposition:
When $m_\ell = \{1, \ldots, M\}$ denotes the number of retransmitted layers in the backup, then the matrix

$$\bar{\mathbf{Q}}_\ell = \arg \min_{\left\{\mathbf{Q} \in \mathbb{C}^{M m_\ell} : \mathbf{Q}_\ell^\dagger \mathbf{Q}_\ell = \mathbf{I}_{m_\ell}\right\}} \mu(\mathbf{Q}_\ell) \tag{2.27}$$

minimizes the MSE of the estimated signal $\mathbf{G}_\ell \widetilde{\mathbf{y}}_{\ell,k}$, where $\bar{\mathbf{Q}}_\ell$ is a matrix constituted by the columns of $\mathbf{V}_\ell$ associated with the m_ℓ smallest eigenvalues in $\mathbf{S}_\ell^2$.

Proof:
First, the case of $m_\ell = 1$ is considered. Here, the matrix $\mathbf{Q}_\ell$ is a column vector $\mathbf{Q}_\ell = \mathbf{q}_\ell$ of size $N \times 1$. By using (2.4), the MSE in (2.26) can be expressed as

$$\begin{aligned}\mathrm{tr}\left(\left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell + \alpha_\ell \mathbf{q}_\ell \mathbf{q}_\ell^\dagger\right)^{-1}\right) &= \mathrm{tr}\left(\left(\mathbf{V}_\ell \mathbf{S}_\ell^2 \mathbf{V}_\ell^\dagger + \alpha_\ell \mathbf{q}_\ell \mathbf{q}_\ell^\dagger\right)^{-1}\right) \\ &= \mathrm{tr}\left(\left(\mathbf{V}_\ell \left(\mathbf{S}_\ell^2 + \alpha_\ell \mathbf{V}_\ell^\dagger \mathbf{q}_\ell \mathbf{q}_\ell^\dagger \mathbf{V}_\ell\right) \mathbf{V}_\ell^\dagger\right)^{-1}\right) \\ &= \mathrm{tr}\left(\left(\mathbf{S}_\ell^2 + \alpha_\ell \bar{\mathbf{q}}_\ell \bar{\mathbf{q}}_\ell^\dagger\right)^{-1}\right), \end{aligned} \tag{2.28}$$

with $\bar{\mathbf{q}}_\ell = \mathbf{V}_\ell^\dagger \mathbf{q}_\ell$. Applying the Sherman–Morrison matrix inversion iden-

tity [58] leads to

$$\begin{aligned}\operatorname{tr}\left(\left(\mathbf{S}_\ell^2+\alpha_\ell\bar{\mathbf{q}}_\ell\bar{\mathbf{q}}_\ell^\dagger\right)^{-1}\right)&=\operatorname{tr}\left(\mathbf{S}_\ell^{-2}\right)-\operatorname{tr}\left(\frac{\mathbf{S}_\ell^{-2}\bar{\mathbf{q}}_\ell\bar{\mathbf{q}}_\ell^\dagger\mathbf{S}_\ell^{-2}}{\alpha_\ell^{-1}+\bar{\mathbf{q}}_\ell^\dagger\mathbf{S}_\ell^{-2}\bar{\mathbf{q}}_\ell}\right)\\&=\operatorname{tr}\left(\mathbf{S}_\ell^{-2}\right)-\operatorname{tr}\left(\frac{\bar{\mathbf{q}}_\ell^\dagger\mathbf{S}_\ell^{-4}\bar{\mathbf{q}}_\ell}{\alpha_\ell^{-1}+\bar{\mathbf{q}}_\ell^\dagger\mathbf{S}_\ell^{-2}\bar{\mathbf{q}}_\ell}\right),\end{aligned}\tag{2.29}$$

where the latter expression is due to the fact that the trace is invariant under cyclic permutations. From (2.26) and (2.29), it can be easily seen that (2.27) for $m_\ell = 1$ is equivalent to the problem of finding a $\bar{\mathbf{q}}_{\text{opt}}$ which maximizes

$$f\left(\bar{\mathbf{q}}_\ell\right)=\frac{\bar{\mathbf{q}}_\ell^\dagger\mathbf{S}_\ell^{-4}\bar{\mathbf{q}}_\ell}{\alpha_\ell^{-1}+\bar{\mathbf{q}}_\ell^\dagger\mathbf{S}_\ell^{-2}\bar{\mathbf{q}}_\ell},\tag{2.30}$$

such that

$$\bar{\mathbf{q}}_{\text{opt}}=\arg\max_{\|\bar{\mathbf{q}}_\ell\|_{\text{F}}^2=1}f\left(\bar{\mathbf{q}}_\ell\right),\tag{2.31}$$

with $\|\bar{\mathbf{q}}_\ell\|_{\text{F}}^2=\bar{q}_{\ell,1}^2+\ldots+\bar{q}_{\ell,M}^2$. Rewriting $f\left(\bar{\mathbf{q}}_\ell\right)$ in terms of the elements of $\bar{\mathbf{q}}_\ell=\left[\bar{q}_{\ell,1},\bar{q}_{\ell,2},\ldots,\bar{q}_{\ell,M}\right]^{\text{T}}$ and exploiting that $\mathbf{S}_\ell^2$ is a diagonal matrix containing the positive eigenvalues $\lambda_1^2\left(\mathbf{H}_\ell\right),\ldots,\lambda_M^2\left(\mathbf{H}_\ell\right)$ of $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$ results to

$$f\left(w\right)=\frac{\sum\limits_{m=1}^{M}w_m\lambda_m^{-4}\left(\mathbf{H}_\ell\right)}{\alpha_\ell^{-1}+\sum\limits_{m=1}^{M}w_m\lambda_m^{-2}\left(\mathbf{H}_\ell\right)},\quad\text{with}\quad\sum_{m=1}^{M}w_m=1,\tag{2.32}$$

where $w_m=\left[\Re\left(\bar{q}_{\ell,m}\right)\right]^2+\left[\Im\left(\bar{q}_{\ell,m}\right)\right]^2$ denotes the mth element of the vector $w=\left[w_1,\ldots,w_M\right]^{\text{T}}$. Since $\lambda_1^2\left(\mathbf{H}_\ell\right)\leq\lambda_2^2\left(\mathbf{H}_\ell\right)\leq\ldots\leq\lambda_M^2\left(\mathbf{H}_\ell\right)$, an intuitive approach is to choose $w=\left[1,0,\ldots,0\right]^{\text{T}}$, i.e., $w_1=1$ while all other elements of w are equal to zero. According to (2.28), for this choice, the optimal backup addressing vector $\mathbf{q}_{\text{opt}}$ is given by the column of $\mathbf{V}_\ell$ associated with the smallest eigenvalue of $\mathbf{S}_\ell^2$, which is proven in the following.

Assuming two arbitrary vectors $w=\left[w_1,\ldots,w_M\right]^{\text{T}}$ and $\widetilde{w}=\left[w_1+w_k,w_2,\ldots,w_{k-1},0,w_{k+1},\ldots,w_M\right]^{\text{T}}$, which both comply with

$\sum_{m=1}^{M} w_m = 1$ and $w_m \geq 0$, but have different entries at the first and kth position, the inequality

$$f(\widetilde{w}) - f(w) \geq 0 \tag{2.33}$$

has to be fulfilled. If (2.33) is true for any $k \in \{2, \ldots, M\}$, each element w_m with $m > 1$ can, in particular, be chosen zero and yields $w = [1, 0, \ldots, 0]^{\mathrm{T}}$. For simplicity matters, the mth eigenvalue $\lambda_m^2(\mathbf{H}_\ell)$ is denoted as λ_m^2 and we obtain

$$\begin{aligned} f(\widetilde{w}) - f(w) &= \frac{\sum_{m=1}^{M} w_m \lambda_m^{-4} + w_k \lambda_1^{-4} - w_k \lambda_k^{-4}}{\alpha_\ell^{-1} + \sum_{m=1}^{M} w_m \lambda_m^{-2} + w_k \lambda_1^{-2} - w_k \lambda_k^{-2}} - \frac{\sum_{m=1}^{M} w_m \lambda_m^{-4}}{\alpha_\ell^{-1} + \sum_{m=1}^{M} w_m \lambda_m^{-2}} \\ &= \underbrace{\frac{1}{\left(\alpha_\ell^{-1} + \sum_{m=1}^{M} w_m \lambda_m^{-2}\right)\left(\alpha_\ell^{-1} + \sum_{m=1}^{M} w_m \lambda_m^{-2} + w_k \lambda_1^{-2} - w_k \lambda_k^{-2}\right)}}_{>0} \\ &\quad \times \left[\left(\sum_{m=1}^{M} w_m \lambda_m^{-4} + w_k \lambda_1^{-4} - w_k \lambda_k^{-4}\right)\left(\alpha_\ell^{-1} + \sum_{m=1}^{M} w_m \lambda_m^{-2}\right)\right. \\ &\quad \left. - \left(\sum_{m=1}^{M} w_m \lambda_m^{-4}\right)\left(\alpha_\ell^{-1} + \sum_{m=1}^{M} w_m \lambda_m^{-2} + w_k \lambda_1^{-2} - w_k \lambda_k^{-2}\right)\right]. \end{aligned} \tag{2.34}$$

Since the denominator of (2.34) is greater than zero, (2.33) is fulfilled if the numerator is non-negative. For $w_k = 0$, the numerator is zero. For $w_k > 0$, we have to prove that

$$\left(\lambda_1^{-2} - \lambda_k^{-2}\right)\left(\left(\alpha_\ell^{-1} + \sum_{m=1}^{M} w_m \lambda_m^{-2}\right)\left(\lambda_1^{-2} + \lambda_k^{-2}\right) - \sum_{m=1}^{M} w_m \lambda_m^{-4}\right) \geq 0. \tag{2.35}$$

Since $\lambda_1^{-2} \geq \lambda_k^{-2}$, (2.35) is equivalent to

$$\alpha_\ell^{-1}\left(\lambda_1^{-2} + \lambda_k^{-2}\right) + \sum_{m=1}^{M} w_m \lambda_m^{-2}\left(\lambda_1^{-2} + \lambda_k^{-2}\right) - \sum_{m=1}^{M} w_m \lambda_m^{-4} \geq 0. \tag{2.36}$$

Since $\alpha_\ell^{-1}\left(\lambda_1^{-2}+\lambda_k^{-2}\right)\geq 0$, a sufficient condition for (2.36) is

$$\sum_{m=1}^{M} w_m \lambda_m^{-2}\left(\lambda_1^{-2}+\lambda_k^{-2}\right)-\sum_{m=1}^{M} w_m \lambda_m^{-4}\geq 0, \tag{2.37}$$

which is true, since $\lambda_1^{-2}\geq\lambda_2^{-2}\geq\ldots\geq\lambda_M^{-2}$.

Depending on the applied policy (see Chapt. 3), multiple layers for transmitting the backup in the subsequent frame may be necessary, i.e., $m_\ell>1$. In this case, the different columns of $\mathbf{Q}_\ell$ have to be orthogonal to each other, thus any two columns $\mathbf{q}_i$ and $\mathbf{q}_j$ of $\mathbf{Q}_\ell$ have to satisfy

$$\langle \mathbf{q}_i,\mathbf{q}_j\rangle=\begin{cases}\|\mathbf{q}_i\|_{\mathrm{F}}\|\mathbf{q}_j\|_{\mathrm{F}} & \text{if } i=j\\ 0 & \text{else}\end{cases},\quad \forall i,j\in\{1,\ldots,m_\ell\}, \tag{2.38}$$

where $\langle \mathbf{q}_i,\mathbf{q}_j\rangle=\mathbf{q}_i^\dagger\mathbf{q}_j$ denotes the inner product of vector $\mathbf{q}_i$ and $\mathbf{q}_j$. This can be explained by referring to the proof of (2.31). If any optimal $\mathbf{q}_s$, which is an eigenvector of $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$, has been found, a matrix $\alpha_\ell\mathbf{q}_s\mathbf{q}_s^\dagger$ can be added to $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$ without changing the underlying set of eigenvectors. In general, if any new $\mathbf{q}_t$ is an eigenvector of

$$\left(\mathbf{H}_\ell^\dagger\mathbf{H}_\ell+\alpha_\ell\sum_{\substack{i=1\\ i\neq t}}^{m}\mathbf{q}_i\mathbf{q}_i^\dagger\right),\quad \forall t\in\{1,\ldots,m\} \tag{2.39}$$

then $\mathbf{q}_t$ is also an eigenvector of

$$\begin{aligned}\left(\mathbf{H}_\ell^\dagger\mathbf{H}_\ell+\alpha_\ell\left(\sum_{\substack{i=1\\ i\neq t}}^{m}\mathbf{q}_i\mathbf{q}_i^\dagger+\mathbf{q}_t\mathbf{q}_t^\dagger\right)\right)&=\left(\mathbf{H}_\ell^\dagger\mathbf{H}_\ell+\alpha_\ell\sum_{i=1}^{m}\mathbf{q}_i\mathbf{q}_i^\dagger\right)\\ &=\left(\mathbf{H}_\ell^\dagger\mathbf{H}_\ell+\alpha_\ell\mathbf{Q}_\ell\mathbf{Q}_\ell^\dagger\right)\end{aligned} \tag{2.40}$$

which is a proof for (2.38). Finally, the approach that any found $\mathbf{q}_t$ is an eigenvector of $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$ is true, since it has been shown above that $\mathbf{q}_t$ is an eigenvector of $\left(\mathbf{H}_\ell^\dagger\mathbf{H}_\ell+\alpha_\ell\mathbf{Q}_\ell\mathbf{Q}_\ell^\dagger\right)$ and of $\alpha_\ell\mathbf{Q}_\ell\mathbf{Q}_\ell^\dagger$.

Chapter 3.

Ergodic RSM Performance

In this chapter, different policies and their resulting performances are considered. A policy defines under which circumstances a retransmission is requested by the receiver and what the properties of this retransmission are, e.g. the dimension of the backup signal. In order to assess the performance of the proposed RSM scheme, it is compared with a number of standard MIMO transmission architectures relying on a conventional ARQ mechanism. Further, the impact of a limited feedback channel is investigated and the performance gain of combining UP and RSM is demonstrated. The last section extends the analysis to the case of systems with a large number of transmit and receive antennas.

3.1. Policies

It has been shown in Sect. 2.4.3 that the backup signal is addressed by the $M \times m_\ell$ matrix $\mathbf{Q}_\ell$ to minimize the postprocessing MSE $\mu(\mathbf{Q}_\ell)$. Hence, for a given value of m_ℓ, the ideal m_ℓ-dimensional subspace is known, but the question is *how to choose the dimension* m_ℓ of the backup signal subspace. In terms of the overall capacity, a trade-off has to be found. Larger $\ldots, m_\ell, m_{\ell+1}, \ldots$ reduce the noise at the decoder input, but at the cost of a lower payload. As the ZF receiver enables a diagonalization of the channel, the RSM constrained capacity is the sum of the capacities of each layer [11]. We assume Gaussian noise terms [51] as considered in

Sect. 2.1 and, using (2.26), define the noise power amplification[1]

$$\varrho_\ell = \frac{\mu\left(\mathbf{Q}_\ell\right)}{\vartheta_{\text{noise}}} = \frac{1}{M}\text{tr}\left(\left(\mathbf{H}_\ell^\dagger\mathbf{H}_\ell + \alpha_\ell\mathbf{Q}_\ell\mathbf{Q}_\ell^\dagger\right)^{-1}\right), \tag{3.1}$$

which is the second moment of the noise terms in the reconstructed signal $\mathbf{G}_\ell\widetilde{\mathbf{y}}_{\ell,k}$ normalized by ϑ_{noise}. Further we define the RSM constrained postprocessing SNR $\gamma_{\text{dec}} = \gamma_{\text{inp}}/\varrho_\ell$ at the SISO decoder input. For the transmission of SISO encoded data $\mathbf{b}_{\ell,k}$ in the ℓth frame $g_\ell = M - m_{\ell-1}$ signal layers can be used, thus, g_ℓ defines the multiplexing gain of the transmission. Thus, the instantaneous capacity of the RSM transmission system for frame ℓ is given by

$$C_\ell = g_\ell \log_2\left(1 + \frac{\gamma_{\text{inp}}}{\varrho_\ell}\right). \tag{3.2}$$

In an $N \times M$ MIMO system with $N \geq M$, we have $0 \leq g_\ell \leq M$. Considering RSM, the multiplexing gain of the active frame can be reduced by the necessity of transmitting a backup signal for the preceding frame, which occupies a certain number of dimensions $m_{\ell-1}$. The demultiplexer accepts g_ℓ dimensions from the encoder for the ℓth frame, but this number varies from frame to frame, due to the time-variant channel.

In order to characterize the choice of the signals to be retransmitted, we define a *policy* φ, which is a mapping from the eigenvalue vector $\Lambda_\ell = \left[\lambda_1^2(\mathbf{H}_\ell), \ldots, \lambda_M^2(\mathbf{H}_\ell)\right]^{\text{T}} \in \{0, \mathbb{R}_+\}^M$ to $m_\ell \in \{0, \ldots, M\}$ so that $\varphi : \{0, \mathbb{R}_+\}^M \to \{0, \ldots, M\}$ and

$$m_\ell = \varphi\left(\left[\lambda_1^2(\mathbf{H}_\ell), \ldots, \lambda_M^2(\mathbf{H}_\ell)\right]^{\text{T}}\right) = \varphi\left(\Lambda_\ell\right). \tag{3.3}$$

There is no backup signal in the subsequent frame whenever $\varphi(\cdot) = 0$, whereas $\varphi(\cdot) = M$ indicates that a full retransmission of the active frame is requested.

[1] We distinguish ϱ_ℓ from the scaling factor α_ℓ^{-1}: while $\varrho_\ell = \mu\left(\mathbf{Q}_\ell\right)/\vartheta_{\text{noise}}$ considers the pre- and postprocessing noise terms of the same frame, $\alpha_\ell^{-1} = \vartheta_{\text{backup}}^{(\ell)}/\vartheta_{\text{noise}}$ corresponds to the backup noise content of the subsequent frame.

Policy 1:

It can be easily seen from (2.26) that the smallest eigenvalue $\lambda_1^2(\mathbf{H}_\ell)$ has the greatest impact to the resulting MSE of the reproduced signal. Hence, regarding the limitation of the noise power amplification it is more valuable to support the corresponding dimension associated with the smallest eigenvalue by a backup signal then any other dimension. Applying policy φ_1 is to request a *one-dimensional* backup to be transmitted in the subsequent frame, if the smallest eigenvalue of the active frame is below a certain threshold ξ, thus

$$\varphi_1(\Lambda_\ell) = \begin{cases} 1, & \text{if } \lambda_1^2(\mathbf{H}_\ell) < \xi \\ 0, & \text{if } \lambda_1^2(\mathbf{H}_\ell) \geq \xi \end{cases} . \tag{3.4}$$

The decoding, and thus the multiplexing at the receiver, starts as soon as the smallest eigenvalue $\lambda_1^2(\mathbf{H}_\ell) \geq \xi$.

Policy 2:

Applying policy φ_2 is to request a *multi-dimensional* backup to be transmitted in the subsequent frame for all eigenvalues of $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ below a certain threshold ξ, thus

$$\varphi_2(\Lambda_\ell) = \left|\left\{i = 1, \ldots, M : \lambda_i^2(\mathbf{H}_\ell) < \xi\right\}\right|, \tag{3.5}$$

with $|\{\cdot\}|$ denoting the cardinality of the set. The decoding, and thus the multiplexing at the receiver, starts as soon as *all* eigenvalues are greater or equal to the threshold ξ. This guarantees an SNR γ_{dec} at the SISO decoder input, which is at least equal to $\gamma_{\text{inp}}\xi$, as can be easily seen from the following example.

Considering the eigenvalues $\lambda_1^2(\mathbf{H}_L), \ldots, \lambda_M^2(\mathbf{H}_L)$ of the last of a sequence of transmitted frames to be greater or equal to a threshold ξ, the noise power amplification at this very frame is given by (3.1) for $\mathbf{Q}_\ell = \mathbf{0}$ and reads

$$\varrho_L = \frac{1}{M}\text{tr}\left(\left(\mathbf{H}_L^\dagger \mathbf{H}_L\right)^{-1}\right) = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{\lambda_i^2(\mathbf{H}_L)} \tag{3.6}$$

with $\lambda_i^2(\mathbf{H}_L) \geq \xi$ for $i = 1, \ldots, M$. In this case, it can be easily seen

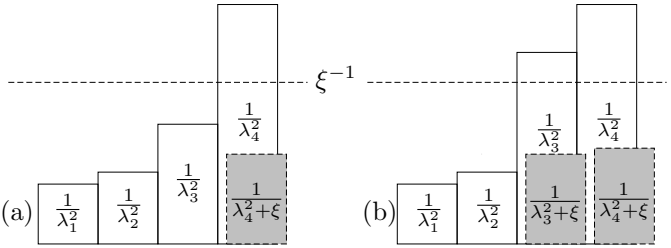

Figure 3.1 Comparing noise enhancement without RSM (white) and with RSM (gray) for policy (3.5), when (a) one eigenvalue or (b) two eigenvalues are below threshold ξ.

from (3.6) that $\varrho_L \leq \xi^{-1}$. Hence, the postprocessing SNR γ_{dec} of the estimated signal, which is also the SNR of the inherent backup signal for the previous frame, provides $\gamma_{\mathrm{dec}} = \gamma_{\mathrm{inp}}/\varrho_L \geq \gamma_{\mathrm{inp}}\xi$. Supporting the signal dimensions of the $(L-1)$th frame associated with eigenvalues below the threshold ξ, the backup also increases the postprocessing SNR of this very frame.

Fig. 3.1 shows the noise enhancement per layer caused by each eigenvalue of $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ when applying a ZF equalizer. Two different scenarios are considered. In (a) the large noise enhancement of layer 4 (white bar) is limited to the gray shaded value by providing a one-dimensional backup signal. The case (b) demonstrates the same issue, if two eigenvalues are below the threshold. It can be seen, that the noise power amplification ϱ_ℓ is limited to a maximum of ξ^{-1}, since the backup forces *each layer* to fulfill this requirement.

Policy 3:

Policy φ_3 is similar to policy φ_2, that is, the backup signal comprises as many dimensions of the MIMO transmit signal as necessary for a certain SNR of the reconstructed signal. But in contrast to policy φ_2, where the noise enhancement of each MIMO layer is below a certain limit, the aim of policy φ_3 is to limit the *average* noise amplification ϱ_ℓ of the received frame. Since the trace of a matrix is equal to the sum of its eigenvalues,

(3.1) can be modified to represent policy φ_3 as

$$\varphi_3(\Lambda_\ell) = \min\{m = 0, \ldots, M : \varrho_\ell(m) \leq \check{\varrho}\} \tag{3.7}$$

with

$$\varrho_\ell(m) = \frac{1}{M}\left(\sum_{i=1}^{m}\left(\lambda_i^2(\mathbf{H}_\ell) + \alpha_\ell\right)^{-1} + \sum_{i=m+1}^{M}\left(\lambda_i^2(\mathbf{H}_\ell)\right)^{-1}\right). \tag{3.8}$$

Multiplexing at the receiver starts as soon as a frame fulfills the condition $\varrho_\ell(0) \leq \check{\varrho}$, i.e., no retransmission of any signal part is needed to limit the MSE of the reconstructed signal. While policy φ_2 depends only on the number of eigenvalues that are below a certain threshold, policy φ_3 is a function of *all* eigenvalues of the received frame. This property is depicted in Fig. 3.2.

This policy ensures a mean postprocessing SNR per frame, which is greater or equal to $\gamma_{\text{inp}}/\check{\varrho}$. The most obvious difference to policy φ_2 can be seen in case (a) of Fig. 3.2: though two eigenvalues are below a level $\check{\varrho}^{-1}$, only the smallest eigenvalue requires a retransmission yielding the (layer-) average noise enhancement $\varrho_\ell \leq \check{\varrho}$. In (b) two eigenvalues are supported by a backup signal.

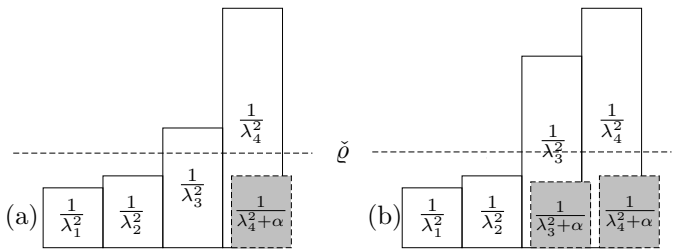

Figure 3.2 Comparing noise enhancement without RSM (white) and with RSM (gray) applying policy (3.7), when the backup signal is required to be (a) one-dimensional or (b) two-dimensional.

3.2. Performance Comparison

In this section, the previously proposed policies for RSM are compared with each other in terms of their performance. Further, a performance comparison with a number of standard transmission architectures is examined which are based on a conventional ARQ mechanism, that is, only full frame retransmissions are requested if necessary. We focus on the achievable long-term average capacity with SISO decoders attuned to a target SNR $\gamma_{\mathrm{inp}}/\breve{\varrho}$ or an outage capacity in the case of standard architectures.[2] The target SNR, threshold and outage capacities are optimally chosen with respect to the capacity.

3.2.1. Capacity

Ergodic MIMO Capacity: Based on Shannon's theory of communication [59] and referring to [12, 11], the instantaneous capacity of the MIMO channel for the ℓth frame without any channel knowledge at the transmitter is given by

$$C_{\mathrm{MIMO}}(\mathbf{H}_\ell) = \log_2\left(\det\left(\mathbf{I}_N + \gamma_{\mathrm{inp}}\mathbf{H}_\ell\mathbf{H}_\ell^\dagger\right)\right) \tag{3.9}$$

$$= \sum_{i=1}^{M} \log_2\left(1 + \gamma_{\mathrm{inp}}\lambda_i^2(\mathbf{H}_\ell)\right) \tag{3.10}$$

in bits per channel use. It can be easily seen from (3.10) that the instantaneous MIMO capacity results from summing the capacities of the SISO-equivalent layers of the channel. This implies optimal coding across the transmit antennas and optimal decoding at the receiver for perfect diagonalization of the MIMO channel. The MIMO channel is considered as a zero-mean circularly symmetric Gaussian distributed random matrix $\mathbf{H}$ with variance $1/M$ and independent elements. Referring to Sect. 2.1, a block fading MIMO channel model is assumed, where $\mathbf{H}_1, \mathbf{H}_2, \ldots$ are independent realizations of $\mathbf{H}$. The ergodic capacity of an unconstrained

[2] A non-ergodic treatment of the performance is given in Sect. 4.

MIMO system is thus defined as

$$\bar{C}_{\mathrm{MIMO}} = E\left[\sum_{i=1}^{M} \log_2\left(1+\gamma_{\mathrm{inp}}\lambda_i^2(\mathbf{H})\right)\right]. \tag{3.11}$$

Optimal MIMO Coding/ARQ: As discussed in Sect. 2.4.2, in a conventional ARQ system a (complete) retransmission of a transmit frame is requested, if the channel is in bad condition. In other words, if the capacity provided by the present channel is below a considered data rate R_{out} an outage of the transmit message occurs. The ϵ-outage capacity is given by $C_{\mathrm{ARQ}}(\epsilon) = \sup\{R_{\mathrm{out}} > 0 : \Pr\left[C_{\mathrm{MIMO}}(\mathbf{H}) < R_{\mathrm{out}}\right] \leq \epsilon\}$, where ϵ denotes the probability that the MIMO capacity $C_{\mathrm{MIMO}}(\mathbf{H})$ is below the predefined rate R_{out}, with $0 < \epsilon < 1$. The capacity per channel use of a conventional ARQ architecture is defined as

$$\bar{C}_{\mathrm{ARQ}} = \sup_{0<\epsilon<1} (1-\epsilon) C_{\mathrm{ARQ}}(\epsilon). \tag{3.12}$$

The factor $(1-\epsilon)$ represents the loss in capacity due to the requested retransmissions.

ZF Signal Reconstruction/ARQ: A ZF receiver has been introduced in Sect. 2.3.1. It has been shown that multiplying the received signal $\mathbf{y}_{\ell,k}$ with the (pseudo-) inverse of the channel matrix $\mathbf{H}_\ell^+$ leads to additive noise vectors after the ZF with covariance matrix $\vartheta_{\mathrm{noise}}(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell)^{-1}$. Proceeding as in (3.10) and using (2.9), the instantaneous constrained capacity of the ZF receiver is given as [60]

$$C_{\mathrm{ZF}}(\mathbf{H}_\ell) = \sum_{i=1}^{M} \log_2\left(1+\gamma_{\ell,i}^{(\mathrm{ZF})}\right). \tag{3.13}$$

When we embed a ZF receiver in a MIMO ARQ architecture, an outage occurs if $C_{\mathrm{ZF}}(\mathbf{H}_\ell) < R_{\mathrm{out}}$ and, thus, the ergodic capacity per channel use follows from (3.12) as

$$\bar{C}_{\mathrm{ZF}} = \sup_{0<\epsilon<1} (1-\epsilon) C_{\mathrm{ZF}}(\epsilon), \tag{3.14}$$

with

$$C_{\mathrm{ZF}}(\epsilon) = \sup\left\{R_{\mathrm{out}} > 0 : \Pr\left[C_{\mathrm{ZF}}(\mathbf{H}) < R_{\mathrm{out}}\right] \leq \epsilon\right\}. \tag{3.15}$$

LMMSE Signal Reconstruction/ARQ: Similarly to (3.13) the instantaneous constrained capacity in case of an LMMSE signal reconstruction is defined by the sum of the Shannon capacities of each layer of the channel. Due to the remaining inter-layer interference after the LMMSE signal reconstruction, the SINR on the mth layer, denoted by $\kappa_{\ell,m}$ from (2.13), is considered here, such that

$$C_{\mathrm{LMMSE}}(\mathbf{H}_\ell) = \sum_{i=1}^{M} \log_2\left(1 + \kappa_{\ell,i}\right). \tag{3.16}$$

According to the principal in (3.12) and (3.14) the ergodic capacity per channel use equals

$$\bar{C}_{\mathrm{LMMSE}} = \sup_{0<\epsilon<1} (1-\epsilon) C_{\mathrm{LMMSE}}(\epsilon), \tag{3.17}$$

with

$$C_{\mathrm{LMMSE}}(\epsilon) = \sup\left\{R_{\mathrm{out}} > 0 : \Pr\left[C_{\mathrm{LMMSE}}(\mathbf{H}) < R_{\mathrm{out}}\right] \leq \epsilon\right\}, \tag{3.18}$$

when we use an LMMSE receiver in a MIMO ARQ system.

RSM: The minimal MSE of the estimated signal $\widehat{\mathbf{x}}_{\ell,k}^{(\mathrm{ZF})}$ after the ZF is $\mu(\bar{\mathbf{Q}}_\ell)$ as given by (2.26) and (2.27). Since $\mathbf{H}$ is now considered a random variable, we define $\mu(\mathbf{H}_\ell)$ as the power of the postprocessing noise in the ℓth frame. Note that $\mu(\mathbf{H}_{\ell+1})$ of the succeeding frame is unknown, when the receiver requests an appropriate backup on the basis of the active channel conditions. For an ergodic performance analysis of an RSM system it is supposed that the receiver observes a large number of frames and starts the multiplexing procedure in reverse frame order, where the Lth frame fulfills $\varphi(\Lambda_L) = 0$. Without any delay limitation, the achievable performance depends on the ergodic properties of the sequence $\mu(\mathbf{H}_L), \mu(\mathbf{H}_{L-1}), \ldots$.

Since the spatial multiplexing at the receiver takes place in reverse frame order, the quantities $W_\ell = \mu(\mathbf{H}_{L-\ell})/\vartheta_{\text{noise}}$ are defined, where the sequence $W_0, W_1, \ldots$ represents the MSEs of the reconstructed signals normalized by ϑ_{noise}. With (2.26),

$$W_\ell = \frac{1}{M}\left(\sum_{i=1}^{\varphi(\mathbf{z}_\ell)} \frac{1}{z_{\ell,i} + W_{\ell-1}^{-1}} + \sum_{i=\varphi(\mathbf{z}_\ell)+1}^{M} \frac{1}{z_{\ell,i}}\right) \tag{3.19}$$

for $\ell \in \mathbb{N}$ with the initial value $W_0 = \mu(\mathbf{H}_L)/\vartheta_{\text{noise}}$, where $z_{\ell,i} = \lambda_i^2(\mathbf{H}_{L-\ell})$ and $\mathbf{z}_\ell = [z_{\ell,1}, \ldots, z_{\ell,M}]^{\mathrm{T}}$. The infinite sequence $W_0, W_1, \ldots$ represents a Markov chain with the range $\mathbb{R}_+$. The transition probabilities in this chain depend on the distribution of $\mathbf{z}_\ell$ and the retransmission policy φ. Here, a policy φ is said to be *permissible* if the two conditions

$$1. \quad \Pr\left[\varphi(\mathbf{z}_\ell) = M\right] < 1 \tag{3.20}$$

$$2. \quad E\left[\frac{1}{z_{\ell,\varphi(\mathbf{z}_\ell)+1}} \,|\{\varphi(\mathbf{z}_\ell) < M\}\right] < \infty, \tag{3.21}$$

where $E\left[\cdot \mid \cdot\right]$ denotes the conditional expected value, are fulfilled. (3.20) states that the policy has to be chosen, such that the probability of a full retransmission is less than one. The condition in (3.21) says that the expected noise power amplification of the layers not supported by a retransmission is less than infinity. Obviously, this holds true for all suggested policies given in Sect. 3.1 as long as the threshold/target SNR is below a certain limit. It has been shown in [61] that the Markov chain induced by a permissible policy φ is ergodic in the sense that the probability distributions of $W_1, W_2, \ldots$ converge (weakly) to a unique probability measure $\Psi^{(\varphi)}$, and

$$\lim_{L\to\infty} \frac{1}{L} \sum_{\ell=1}^{L} W_\ell \to S^{(\varphi)} \tag{3.22}$$

where

$$S^{(\varphi)} = \int_{\mathbb{R}_+} W \mathrm{d}\Psi^{(\varphi)}(W) < \infty. \tag{3.23}$$

Since φ_1,φ_2 and φ_3 are permissible policies, the noise power at the de-

coder input, averaged over the processed frames, equals $S^{(\varphi)}\vartheta_{\text{noise}}$ if $L \to \infty$. Averaging the noise power over the different signals for the SISO decoder (according to the transmission on different layers) results also to $S^{(\varphi)}\vartheta_{\text{noise}}$, due to the independence of W_ℓ and $\varphi(\mathbf{z}_{\ell+1})$. The mean multiplexing gain $g^{(\varphi)}$ achieved by the policy φ can be easily derived as

$$g^{(\varphi)} = E\left[g_\ell\right] = M - E\left[\varphi(\mathbf{z}_\ell)\right]. \tag{3.24}$$

The MIMO capacity given in (3.11) is modified to define the RSM constrained capacity in the following. Although the distribution of the noise terms at the output of the multiplexer may be non-Gaussian (due to the averaging), it does not impair the achievable capacity, at least for nearest neighbor decoding [62].

In the case of RSM, the multiplexing gain is reduced by retransmitting backup signals, whereas this very backup enables the reduction of the average noise power amplification after the signal reconstruction. Thus, using $\gamma_{\text{dec}} = \gamma_{\text{inp}}/S^{(\varphi)}$ and the multiplexing gain from (3.24), the RSM constrained ergodic capacity for a given policy can be expressed as

$$\bar{C}_{\text{RSM}}^{(\varphi)}(\gamma_{\text{inp}}) = \sup_{\varphi \in \mathcal{M}} \left(g^{(\varphi)} \log_2 \left(1 + \frac{\gamma_{\text{inp}}}{S^{(\varphi)}} \right) \right), \tag{3.25}$$

where $\mathcal{M}$ denotes a given family of permissible policies. The policies defined in Sect. 3.1 lead to three different families of permissible policies, namely

$$\mathcal{M}_1 = \{\varphi_1(\xi) : \xi > 0\}, \tag{3.26}$$

$$\mathcal{M}_2 = \{\varphi_2(\xi) : \xi > 0\}, \tag{3.27}$$

$$\mathcal{M}_3 = \{\varphi_3(\alpha) : \alpha > 0\}. \tag{3.28}$$

That is, (3.25) denotes the achievable capacity for a certain policy with an optimally chosen ξ or α, respectively. Note, that the notation in (3.26), (3.27), (3.28) refers to finding an optimal policy from a family for a given MIMO channel realization.

3.2.2. Simulation Results

The capacities reported in this section have been obtained by Monte-Carlo simulations[3]. In Chapt. 5 the following results are compared to the RSM performance provided by measurements on a 4×4 MIMO testbed. To this end, most results are presented for a 4×4 and an 8×8 MIMO channel. The latter shall demonstrate the system behavior for larger MIMO systems.

Fig. 3.3 and Fig. 3.4 compare the capacity resulting from policy φ_1 and policy φ_2 in the case of a 4×4 and an 8×8 MIMO system, respectively. The policies are comprised in the two families $\mathcal{M}_1 = \{\varphi_1(\xi) : \xi > 0\}$ and $\mathcal{M}_2 = \{\varphi_2(\xi) : \xi > 0\}$. Note that we focus on the *achievable* capacity of a family, i.e., the threshold ξ is optimized for maximizing the ergodic capacity. Along with the results the unconstrained ergodic MIMO capacity

[3] Consistent results have been obtained by numerical computations of the induced Markov chain in [61].

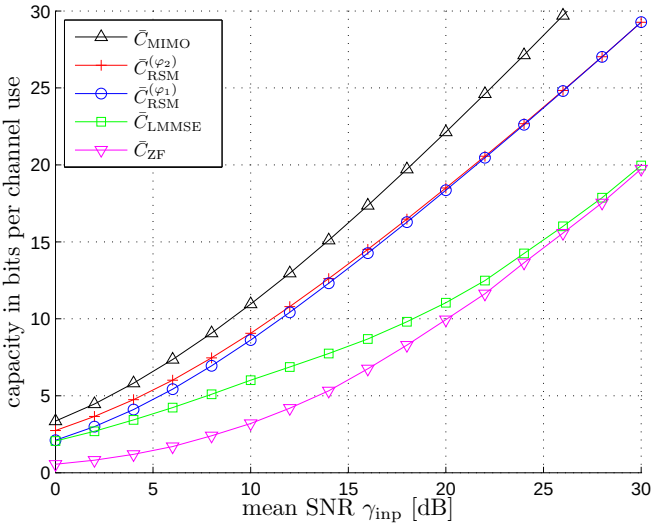

Figure 3.3 Capacity comparison of RSM using policy φ_1 and policy φ_2 over a 4×4 MIMO channel with Rayleigh fading.

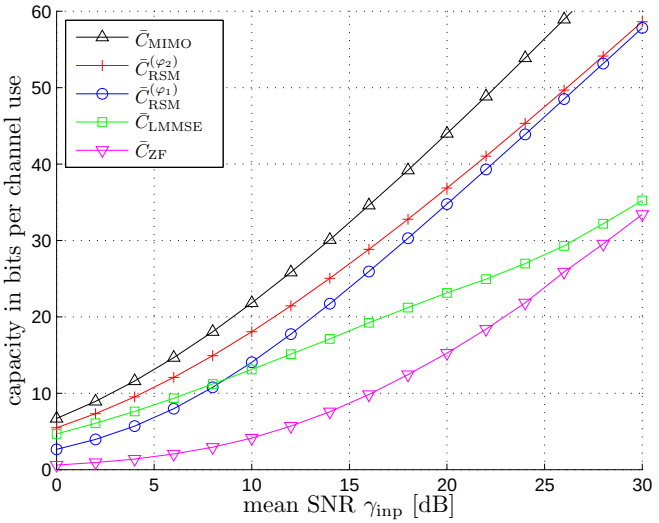

Figure 3.4 Capacity comparison of RSM using policy φ_1 and policy φ_2 over a 8×8 MIMO channel with Rayleigh fading.

$\bar{C}_{\mathrm{MIMO}}$, as denoted in (3.11), is considered as an upper bound. Further, the results are compared with the constrained capacity of two open-loop architectures with array signal reconstruction based on LMMSE and ZF, respectively.

As expected, policy φ_2 outperforms policy φ_1, because the application of φ_2 guarantees a minimal SNR in each frame. Interestingly, the performance difference becomes more apparent in the 8×8 MIMO system. The results in Fig. 3.3 indicate, that the performance is similar for policy φ_1 and φ_2 for $\gamma_{\mathrm{inp}} \geq 12$ dB. As already mentioned, optimizing the expression in (3.25) is equivalent to balance the impact of the multiplexing gain and the noise power amplification. For increasing values of γ_{inp} the average capacity per layer increases, while the noise power amplification still depends on the channel and the applied policy only. That is, the multiplexing gain becomes more pronounced for large values of γ_{inp}. Thus, the necessity for requesting backup signal decreases with increasing SNR, what brings policy φ_2 to approach φ_1.

For MIMO systems of higher dimensions, as for instance an 8×8 MIMO channel, the performance improvement of policy φ_2 becomes more

apparent, as shown in Fig. 3.4. Obviously, with increasing number of considered signal layers, the probability $\Pr[m_\ell > 1]$ increases, which in turn lets policy φ_2 outperform policy φ_1. In case of a low SNR $\gamma_{\text{inp}} \leq 11$ dB even a LMMSE receiver provides a higher capacity than policy φ_1 for RSM.

Fig. 3.5 and Fig. 3.6 show the capacity for policy φ_2 in comparison to policy φ_3. Obviously, these two policies achieve similar ergodic performance, but they are part of different approaches. Referring to the explanations in Sect. 3.1, φ_2 limits the maximum noise power per layer, while φ_3 limits the *average* noise power by an upper bound. This leads to different values of the noise enhancement, but also to different multiplexing gains. In a Rayleigh fading MIMO channel the maximum RSM constrained ergodic capacity seems to be similar for the two policies, as the results suggest.

A more practical comparison of the two policies φ_2 and φ_3 is provided in the following. Both policies provide a guaranteed postprocessing SNR

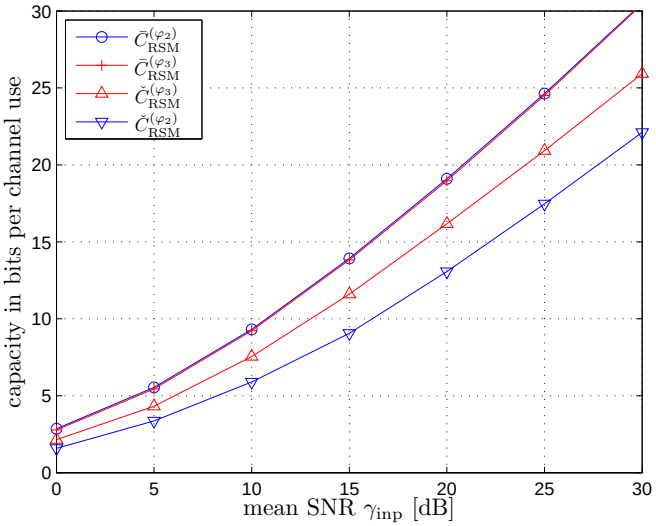

Figure 3.5 Capacity comparison of RSM using policy φ_2 and policy φ_3 over a 4×4 MIMO channel with Rayleigh fading.

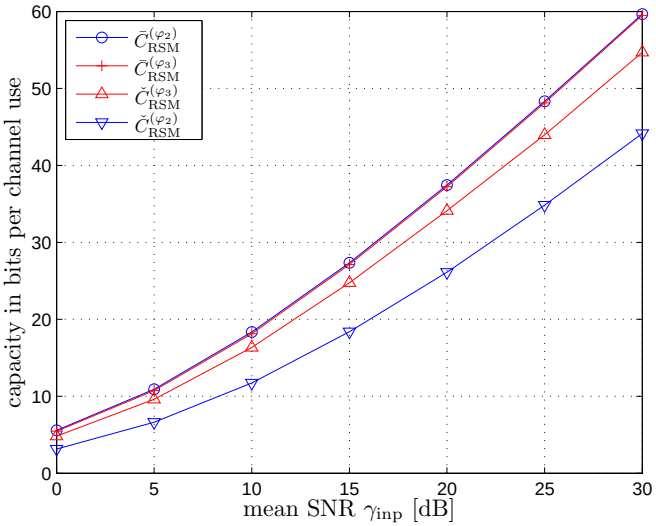

Figure 3.6 Capacity comparison of RSM using policy φ_2 and policy φ_3 over a 8×8 MIMO channel with Rayleigh fading.

$\gamma_{\mathrm{dec}} \geq \gamma_{\mathrm{target}}$ at the SISO decoder input, with $\gamma_{\mathrm{target}} = \gamma_{\mathrm{inp}}/\xi$ for policy φ_2 and $\gamma_{\mathrm{target}} = \gamma_{\mathrm{inp}}/\check{\varrho}$ when applying policy φ_3. As can be easily seen from (3.5) and (3.7), the result of a policy depends on the threshold ξ or the noise power amplification $\check{\varrho}$, respectively. For the moment and according to (3.24), we denote the multiplexing gain $g^{(\varphi)}(\gamma_{\mathrm{target}}) = M - E\left[\varphi(\mathbf{z}_\ell)\right]$ depending on the required postprocessing SNR γ_{target}. Thus, the ergodic capacity considering a guaranteed postprocessing SNR γ_{target} is defined as

$$\check{C}_{\mathrm{RSM}}^{(\varphi)} = \sup_{\gamma_{\mathrm{target}}>0} \left(g^{(\varphi)}(\gamma_{\mathrm{target}}) \log_2 \left(1 + \gamma_{\mathrm{target}}\right) \right). \tag{3.29}$$

This denotes the RSM constrained capacity for which an error-free decoding of the transmitted signal is assured, comparable to an outage capacity. These RSM constrained capacities are also shown in Fig. 3.5

and Fig. 3.6. From the policy explanation in Fig. 3.2 it is evident that policy φ_3 provides a higher guaranteed capacity than policy φ_2 because of the larger achievable multiplexing gain. This comparison cannot be performed for policy φ_1, since a minimum capacity cannot be guaranteed by a retransmission of only one dimension.

Finally, RSM is compared to other simple closed-loop MIMO systems, i.e., ARQ based architectures as described in Sect. 2.4.2, by means of the system constrained capacity. In this comparison, RSM is applied for policy φ_2. Fig. 3.7 and Fig. 3.8 show $\bar{C}_{\mathrm{RSM}}$, $\bar{C}_{\mathrm{ARQ}}$, $\bar{C}_{\mathrm{ZF}}$ and $\bar{C}_{\mathrm{MMSE}}$ versus γ_{inp} for the cases of a 4×4 MIMO channel and a 8×8 MIMO channel, respectively. Additionally the ergodic MIMO capacity $\bar{C}_{\mathrm{MIMO}}$ is shown for reference. At higher values of the SNR the RSM scheme seems to be much closer to the architectures with optimal MIMO coding in terms of capacity than to the ZF and MMSE-based schemes.

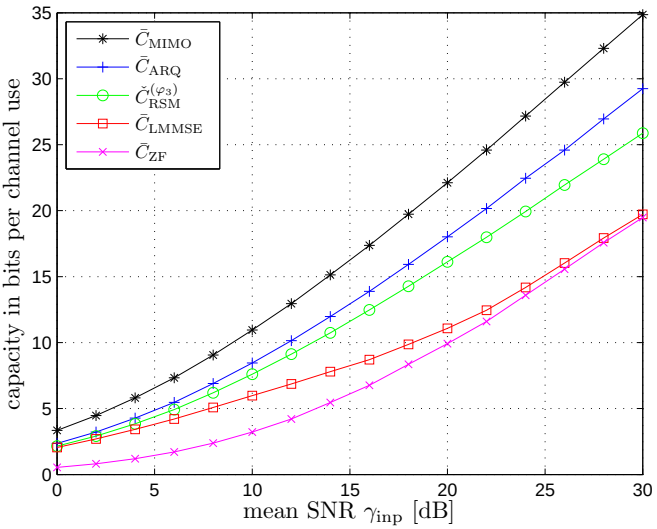

Figure 3.7 Capacity by the RSM-enhanced multiplexing and standard ARQ-based scheme over a 4×4 MIMO channel with Rayleigh fading.

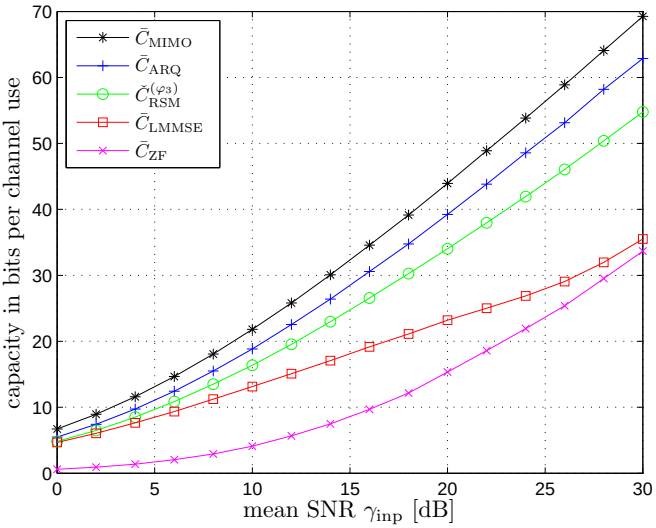

Figure 3.8 Capacity by the RSM-enhanced multiplexing and standard ARQ-based scheme over a 8×8 MIMO channel with Rayleigh fading.

3.3. Limited Feedback Channel

So far, the feedback channel has been considered to be a perfect channel with unlimited capacity. This assumption enables a perfect backup signal addressing, since the optimal retransmission identifier $\bar{\mathbf{Q}}_\ell$ is fed back to the transmitter without any loss. However, in practice, a feedback channel offers only a limited capacity. Hence, the subspace selector $\bar{\mathbf{Q}}_\ell$ has to be quantized, which in turn affects the system performance.

In other closed-loop MIMO systems, where CSI is fed back to the transmitter for precoding purposes, adequate codebooks of precoders are defined, which has been investigated e.g. in [5]. This procedure is equivalent to defining proper codebooks of subspace identifiers in the context of RSM. Finding optimal codebooks leads to the problem of *packings in Grassmannian manifolds*, which has been considered e.g. in [21] for packings of n-dimensional subspaces in $\mathbb{R}^m$, with $m \geq n$. Some examples of these real-valued packings can be found in [63]. In the case of defining codebooks for unitary precoders and subspace selectors in RSM,

the problem is to find *complex-valued* packings in complex Grassmannian manifolds.

Let $\mathcal{S}(M,m)$ represent the Stiefel manifold of all $M \times m$ matrices with orthonormal columns, where $m \in \{1,\ldots,M\}$. Considering two such matrices $\mathbf{U},\mathbf{Q} \in \mathcal{S}(M,m)$, the *chordal distance* between these m-dimensional subspaces (spanned by the columns of $\mathbf{U}$ and $\mathbf{Q}$) is denoted by

$$d_{\text{chord}}(\mathbf{U},\mathbf{Q}) = \frac{1}{\sqrt{2}} \left\| \mathbf{U}\mathbf{U}^{\dagger} - \mathbf{Q}\mathbf{Q}^{\dagger} \right\|_{\text{F}}. \tag{3.30}$$

To request a backup signal for the ℓth frame, rather than $\bar{\mathbf{Q}}_\ell$ the receiver can only choose an element $\widehat{\mathbf{Q}}_\ell$ from a finite codebook $\mathcal{C}_{m_\ell}$. Each set $\mathcal{C}_m \subset \mathcal{S}(M,m)$ comprises a finite number of m-dimensional subspaces of $\mathbb{C}^M$. Note that different codebooks for different values of m have to be predefined, the exact number depends on the range of the applied policy φ.

Of course, the retransmission identifier $\widehat{\mathbf{Q}}_\ell$ from the finite codebook addresses a suboptimal backup signal at the transmitter. The question is *what is the impact of choosing this suboptimal addressing*?

If the sub-optimal backup addressing $\widehat{\mathbf{Q}}_\ell$ is applied at the transmitter, the MSE from (2.26) changes to

$$\begin{aligned} \mu\left(\widehat{\mathbf{Q}}_\ell\right) &= \frac{\vartheta_{\text{noise}}}{M} \text{tr}\left(\left(\mathbf{H}_\ell^{\dagger}\mathbf{H}_\ell + \alpha_\ell \widehat{\mathbf{Q}}_\ell \widehat{\mathbf{Q}}_\ell^{\dagger} \right)^{-1} \right) \\ &= \frac{\vartheta_{\text{noise}}}{M} \text{tr}\left(\left(\mathbf{H}_\ell^{\dagger}\mathbf{H}_\ell + \alpha_\ell \bar{\mathbf{Q}}_\ell \bar{\mathbf{Q}}_\ell^{\dagger} + \alpha_\ell \mathbf{E}_\ell \right)^{-1} \right), \end{aligned} \tag{3.31}$$

where the $M \times M$ Hermitian matrix $\mathbf{E}_\ell = \widehat{\mathbf{Q}}_\ell \widehat{\mathbf{Q}}_\ell^{\dagger} - \bar{\mathbf{Q}}_\ell \bar{\mathbf{Q}}_\ell^{\dagger}$ is non-zero in the case $\widehat{\mathbf{Q}}_\ell \neq \bar{\mathbf{Q}}_\ell$. In the following a reasonable choice for $\widehat{\mathbf{Q}}_\ell$ from a finite codebook $\mathcal{C}_{m_\ell}$ is described. Considering (3.31), the vector containing the

positive eigenvalues of $\left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell + \alpha_\ell \bar{\mathbf{Q}}_\ell \bar{\mathbf{Q}}_\ell^\dagger + \alpha_\ell \mathbf{E}_\ell\right)$ is given by

$$\begin{bmatrix} \lambda_1^2(\mathbf{H}_\ell) + \alpha_\ell + \zeta_{\ell,1} \\ \vdots \\ \lambda_{m_\ell}^2(\mathbf{H}_\ell) + \alpha_\ell + \zeta_{\ell,m_\ell} \\ \lambda_{m_\ell+1}^2(\mathbf{H}_\ell) + \zeta_{\ell,m_\ell+1} \\ \vdots \\ \lambda_M^2(\mathbf{H}_\ell) + \zeta_{\ell,M} \end{bmatrix}$$

with $\zeta_{\ell,1}, \ldots, \zeta_{\ell,M}$ representing the deviations due to a non-zero $\mathbf{E}_\ell$. Because both $\widehat{\mathbf{Q}}_\ell$ and $\bar{\mathbf{Q}}_\ell$ are orthonormal matrices of the same size, it follows that $\mathrm{tr}(\mathbf{E}_\ell) = 0$ and therefore $\sum_{i=1}^M \zeta_{\ell,i} = 0$. It follows from the Wielandt-Hoffman theorem [64] that $\sum_{i=1}^M \zeta_{\ell,i}^2 \leq \|\alpha_\ell \mathbf{E}_\ell\|_{\mathrm{F}}^2$, while $\|\alpha_\ell \mathbf{E}_\ell\|_{\mathrm{F}}^2$ is equal to $2\alpha_\ell^2 d_{\mathrm{chord}}^2(\widehat{\mathbf{Q}}_\ell, \bar{\mathbf{Q}}_\ell)$ according to (3.30). Hence, a small chordal distance $d_{\mathrm{chord}}(\widehat{\mathbf{Q}}_\ell, \bar{\mathbf{Q}}_\ell)$ limits the eigenvalue deviations, such that a reasonable choice of the subspace selector $\widehat{\mathbf{Q}}_\ell$ from a given codebook $\mathcal{C}_{m_\ell}$ is the one with minimal distance to $\bar{\mathbf{Q}}_\ell$, i.e.,

$$\widehat{\mathbf{Q}}_\ell = \arg \min_{\mathbf{U} \in \mathcal{C}_{m_\ell}} d_{\mathrm{chord}}\left(\mathbf{U}, \bar{\mathbf{Q}}_\ell\right). \tag{3.32}$$

Following this choice of the subspace selector, the noise power propagation can be described by the extended Markov chain

$$\widetilde{W}_\ell = \frac{1}{M}\left(\sum_{i=1}^{\varphi(\mathbf{z}_\ell)} \frac{1}{z_{\ell,i} + \widetilde{W}_{\ell-1}^{-1} + D_{\ell,i}} + \sum_{i=\varphi(\mathbf{z}_\ell)+1}^{M} \frac{1}{z_{\ell,i} + D_{\ell,i}}\right) \tag{3.33}$$

for $\ell \in \mathbb{N}$ with some initial value $\widetilde{W}_0 = \widetilde{w}_0$ and $\mathbf{D}_\ell = [D_{\ell,1}, \ldots, D_{\ell,M}]^{\mathrm{T}}$ accounting for the eigenvalue deviations due to $\widehat{\mathbf{Q}}_{L-\ell} \neq \bar{\mathbf{Q}}_{L-\ell}$. According to the considerations of $\zeta_{\ell,1}, \ldots, \zeta_{\ell,M}$ above, the random vectors $\mathbf{D}_1, \mathbf{D}_2, \ldots$ are subject to

$$D_{\ell,1} + D_{\ell,2} + \ldots + D_{\ell,M} = 0 \tag{3.34}$$

$$D_{\ell,1}^2 + D_{\ell,2}^2 + \ldots + D_{\ell,M}^2 \leq 2\widetilde{W}_{\ell-1}^{-2} d_{\max,\varphi(\mathbf{z}_\ell)}^2, \tag{3.35}$$

where $d_{\max,m}$ represents the maximal distance between a certain matrix

$\bar{\mathbf{Q}}_\ell$ and its nearest neighbor $\widehat{\mathbf{Q}}_\ell \in \mathcal{C}_m$, such that

$$d_{\max,m} = \sup_{\mathbf{Q}\in\mathcal{S}(M,m)} \min_{\mathbf{U}\in\mathcal{C}_m} d_{\text{chord}}(\mathbf{U},\mathbf{Q}). \tag{3.36}$$

It has been shown in [61] that the Markov chain in (3.33) is ergodic if the applied policy φ is permissible and additionally there exists a constant $z_0 > 0$ such that $\Pr\left[\left\{z_{\ell,\varphi(\mathbf{z}_\ell)+1} \geq z_0\right\} | \{\varphi(\mathbf{z}_\ell) < M\}\right] = 1$.

Different ways for finding appropriate Grassmannian packings have been suggested by several authors. For instance, an optimization technique which uses a relaxation method in conjunction with gradient search algorithms is presented in [65], while in [66] the authors describe a numerical method based on alternating projection to find good packings in the Grassmannian manifold. A systematic Fourier-based design approach is proposed in [67], which has been used in [5] to create complex-valued codebooks. Some example codebooks are available for download at [68]. The latter method has the advantage of finding packings with a simplified design process. At least, the last two aforementioned techniques suffer from the rapidly increasing complexity, when the ambient dimension M or the number of subspaces in the codebook is large.

For examining the performance degradation in case of a limited feedback channel, we have devised codebooks according to the systematic approach described in [67]. The simplicity of this search algorithm is achieved by finding codebooks with large minimum chordal distances. But it is emphasized that these distances are usually far away from the optimum. Nevertheless, these codebooks are valuable means for demonstrating the effect of a feedback channel with limited capacity.

Fig. 3.9 and Fig. 3.10 show the the performance degradation resulting from a finite-rate feedback channel for the case of retransmission policy φ_1 applied in a 4×4 and a 8×8 RSM system, respectively. More precisely, the RSM constrained ergodic capacity is considered here. The previously designed codebook $\mathcal{C}_1$ consists of $M \times 1$ vectors, since $\varphi_1 \in \{0, 1\}$. The result for $|\mathcal{C}_1| = 256$ has been calculated with a randomly generated codebook, because the systematic codebook design method suffers from inefficiency when the number of codebook entries increases.

The limited capacity caused by a finite-rate feedback channel for the case of retransmission policy φ_2 is shown in Fig. 3.11 and Fig. 3.12, respectively. As mentioned above, a separate codebook is designed for

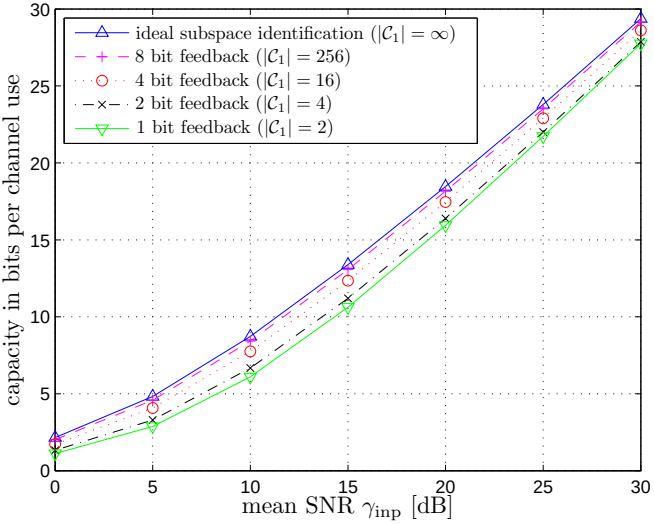

Figure 3.9 Performance degradation of $\bar{C}_{\text{RSM}}$ resulting from finite-rate feedback channel of 4×4 MIMO system for policy φ_1.

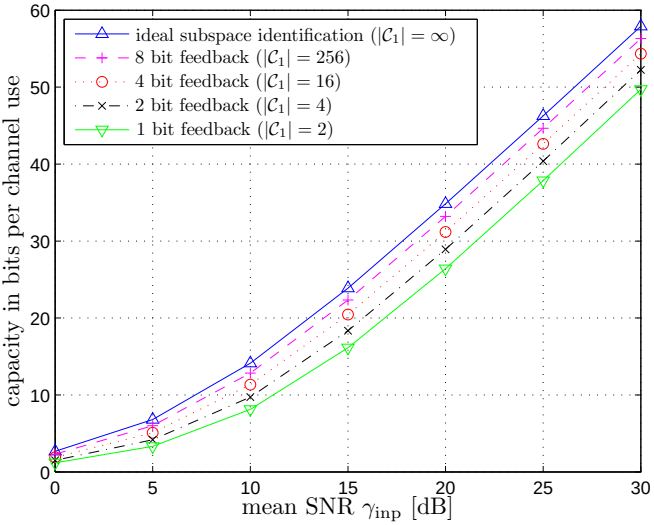

Figure 3.10 Performance degradation of $\bar{C}_{\text{RSM}}$ resulting from finite-rate feedback channel of 8×8 MIMO system for policy φ_1.

every $m \in \{1, \ldots, M-1\}$. Similarly to the results in Sect. 3.2.2, no significant difference between the RSM constrained ergodic capacities of φ_2 and φ_3 has been found, when the capacity of the feedback channel is limited. Hence, the performance plot for φ_3 has been omitted.

The results suggest that the performance degradation due to the limited feedback channel is small, if the codebook is sufficiently large. For the case of an 8×8 MIMO channel, the performance reduction is more significant for policy φ_2. As can be concluded from Fig. 3.12, this performance reduction is higher in the low SNR region, while it decreases slowly for increasing SNR. All results show that the impact of a limited feedback channel is small even if the capacity of the feedback is restricted to only a few bits per channel use.

Referring to the results depicted in Fig. 3.9, Fig. 3.10, Fig. 3.11 and Fig. 3.12, the required capacity of the feedback channel is equal to $\log_2(|\mathcal{C}|\, M + 1)$ bits per channel use, since the corresponding bits address M different codebooks each of $|\mathcal{C}|$ entries and an additional state for the cases "no retransmission" ($m_\ell = 0$).

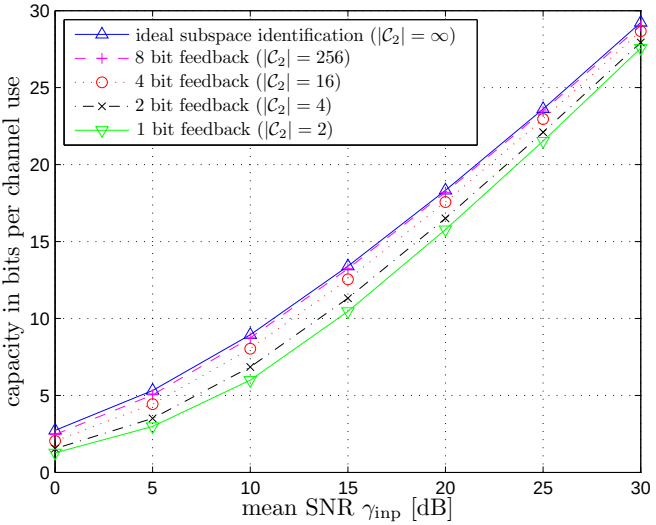

Figure 3.11 Performance degradation of $\bar{C}_{\text{RSM}}$ resulting from finite-rate feedback channel of 4×4 MIMO system for policy φ_2.

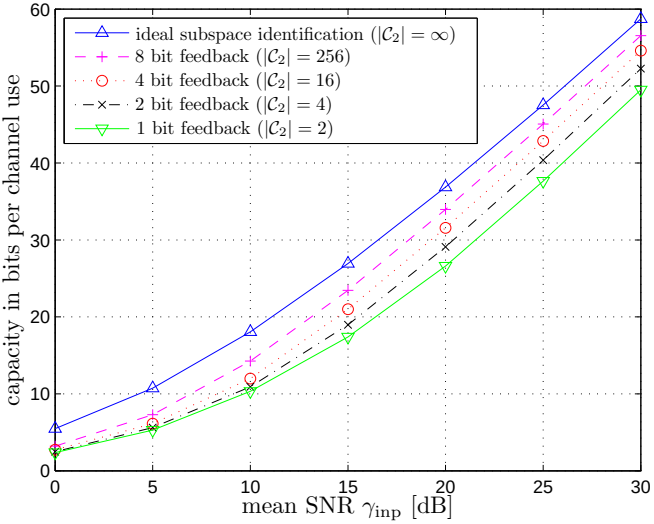

Figure 3.12 Performance degradation of $\bar{C}_{\mathrm{RSM}}$ resulting from finite-rate feedback channel of 8×8 MIMO system for policy φ_2.

3.4. Using Side Information at the SISO Decoder

The RSM system described in Sect. 2.4.3 involves a vertical interleaving/deinterleaving at both the transmitter and receiver. This interleaving across the signal layers has the purpose of averaging over different noise power levels on different layers. One can improve the attainable capacity of a MIMO transmission system by providing side information about these different noise levels to the SISO decoder.

As can be seen from (3.2), the instantaneous capacity depends on the multiplexing gain g_ℓ, the SNR γ_{inp} and the noise power amplification ϱ_ℓ. For the moment, we focus on the instantaneous capacity $C_{\ell,m}$ of each layer m, with $m = m_{\ell-1} + 1, \ldots, M$. Therefore, we define the noise power amplification $\varrho_{\ell,m}$ for the mth layer, such that

$$C_{\ell,m} = \log_2\left(1 + \frac{\gamma_{\mathrm{inp}}}{\varrho_{\ell,m}}\right), \qquad m = m_{\ell-1} + 1, \ldots, M. \tag{3.37}$$

For each frame transmission $M - m_{\ell-1}$ layers contribute to the RSM constrained capacity per MIMO channel use as shown in (3.2), while $m_{\ell-1}$ layers transmit a backup signal for the $(\ell-1)$th frame. For a given γ_{inp}, the capacity $C_{\ell,m}$ is a *convex* function depending on $\varrho_{\ell,m}$. Defining the average noise power amplification of the signal layers corresponding to the upper $M - m_{\ell-1}$ eigenvalues of $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ as

$$\breve{\varrho}_\ell = \frac{1}{M - m_{\ell-1}} \sum_{m=m_{\ell-1}+1}^{M} \varrho_{\ell,m}, \tag{3.38}$$

it can be seen by using the principle of Jensen's inequality [69] that the sum of the instantaneous capacities $C_{\ell,m}$, with $m_{\ell-1}+1, \ldots, M$, is

$$\sum_{m=m_{\ell-1}+1}^{M} C_{\ell,m} \geq (M - m_{\ell-1}) \log_2 \left(1 + \frac{\gamma_{\text{inp}}}{\breve{\varrho}_\ell}\right). \tag{3.39}$$

The aforementioned principle of providing side information is based on (3.39), while the side information is $\gamma_{\text{inp}}/\varrho_{\ell,m}$. Hence, the capacity of RSM depicted above can be even increased by exploiting the capacity of each signal layer separately. For this purpose, the SISO decoder needs to have side information about the noise levels or SNR, respectively, of each layer.

When applying the vertical interleaving/deinterleaving for decorrelating the noise terms of the backup signal, the decoder requires the noise level information for *each* element of $\mathbf{n}'_{\ell,k}$. It could be more advantageous to use a horizontal interleaving/deinterleaving instead, i.e., the signals *within each* of the M layers are reordered and, thus, the noise terms are decorrelated in time. Since the reordering pattern is different for each layer, this results in a vertical decorrelation of the noise terms after the signal reconstruction. But still, the noise terms remain in its corresponding signal layer. Clearly, information about the noise has to be provided on a per-layer basis. Further, the scaling of the backup at the receiver is no more a multiplication by a scalar factor, but turns into a matrix multiplication for scaling each of the signal layers separately.

According to (2.20), the backup signals now contain inherent noise vectors $\bar{\mathbf{n}}_{\ell,1}, \ldots, \bar{\mathbf{n}}_{\ell,K}$, which have zero mean and are uncorrelated as a

result of the interleaving. Therefore we assume that

$$E\left[\bar{\mathbf{n}}_{\ell,k}\bar{\mathbf{n}}_{\ell,k}^{\dagger}\right] = \mathbf{\Omega}_{\ell}\vartheta_{\text{noise}}, \tag{3.40}$$

with $\mathbf{\Omega}_{\ell}$ a diagonal $m_{\ell} \times m_{\ell}$ matrix of the form

$$\mathbf{\Omega}_{\ell} = \begin{bmatrix} \frac{\gamma_{\text{inp}}}{\gamma_{\ell+1,1}^{(\text{ZF})}} & 0 & \dots & 0 \\ 0 & \frac{\gamma_{\text{inp}}}{\gamma_{\ell+1,2}^{(\text{ZF})}} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & \frac{\gamma_{\text{inp}}}{\gamma_{\ell+1,m_{\ell}}^{(\text{ZF})}} \end{bmatrix}, \tag{3.41}$$

where $\gamma_{\ell+1,m}^{(\text{ZF})}$ is the ZF constrained SNR of the mth substream in the subsequent $(\ell+1)$th frame, as denoted in (2.9), which provides a backup for the current $\ell$the frame. A scaling of the backup by $\mathbf{\Omega}_{\ell}^{-\frac{1}{2}}$ matches

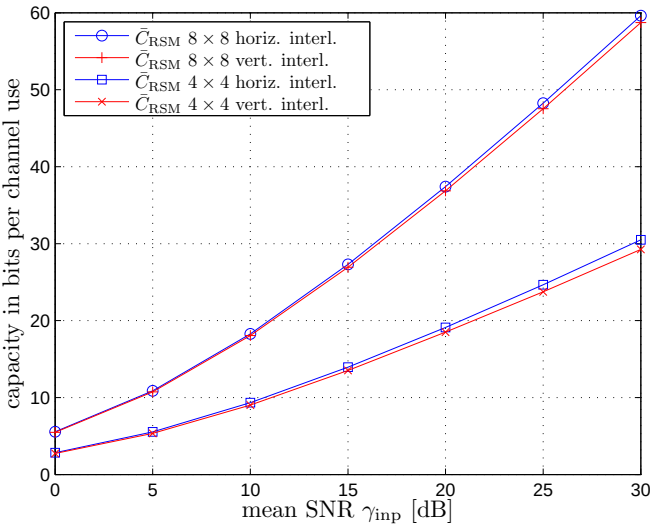

Figure 3.13 Comparing horizontal and vertical interleaving by means of the RSM constrained ergodic capacity for policy φ_2 in a Rayleigh flat-fading MIMO channel (4×4 and 8×8).

the second order properties of the noise in the backup to those of $\mathbf{n}_{\ell,k}$. Fig. 3.13 demonstrates the performance improvement of this approach compared to the previously described vertical interleaving for the case of an RSM system of size 4×4 and 8×8, respectively.

While horizontal interleaving can improve the RSM constrained capacity of a 4×4 MIMO system by up to 1 dB, the recognizable improvement in an 8×8 MIMO architecture is only marginal. In order to follow the objective of low receiver complexity, this work usually deals with vertical interleaving. A demonstration of horizontal interleaving is given in Sect. 3.5.

3.5. Precoded RSM

In Sect. 3.2, the transmitter/receiver-constrained capacity has been investigated. The ergodic MIMO capacity represents an upper limit which is the attainable capacity for a UP MIMO system (see Sect. 2.4.1) with full CSI at the transmitter[4]. CSI at the transmitter end can be used for precoding the multidimensional transmit signal such that the MIMO channel is turned into independent SISO channels [1]. As described in Sect. 2.4.1, each of the signal streams is attenuated by an individual SISO channel, which allows the use of off-the-shelf SISO encoders and decoders in the outer system and enables data rates close to the MIMO channel capacity while keeping the complexity limited. The weakness of this approach is the need for accurate CSI. The system performance is reduced significantly if the CSI is outdated due to a time-varying channel [3], [4] or inaccurate due to limited capacity of the feedback channel [5].

Imperfect CSI at the transmitter necessitates the receiver to recover the parallel signal streams by adequate methods, like ZF or LMMSE filtering at the receiver, as described in Sect. 2.3.1 and Sect. 2.3.2, respectively[5]. Although these techniques result in noise amplification over all signal layers in case of a badly conditioned channel matrix, it has been shown above that RSM enables the receiver to evade an excessive noise ampli-

[4] Clearly, a waterpouring based power allocation at the transmitter would even increase the capacity, but this is not considered in this thesis.

[5] As emphasized above, though there exists a bunch of methods with better performance, we focus on simple reconstruction techniques for complexity reasons.

fication in linear signal reconstruction. While the ergodic performance achievable by RSM is independent of the correlation value of the channel matrices associated with consecutive frame transmissions, the capacity of a MIMO architecture with feedback of CSI and precoding generally decreases for decreasing correlation values. The performance of perfect precoding based on *accurate* and *up-to-date* CSI is shown as $\bar{C}_{\mathrm{MIMO}}$ and $\bar{C}_{\mathrm{ARQ}}$ in Fig. 3.8, for instance. It can be seen that perfect precoding outperforms RSM, whereas in channels varying significantly from frame to frame – possibly because of long intervals between consecutive signal bursts – RSM is supposed to be the superior technique.

Motivated by the fact that similar information needs to be sent back from the receiver to the transmitter for facilitating precoding and backup signal identification, here, a closed-loop MIMO architecture incorporating both UP and RSM (UPRSM) is investigated. Therefore, the necessary system changes are described, and subsequently the performance improvement compared to each individual closed-loop architecture is analyzed. A basic block diagram of the considered system is shown in Fig. 3.14. The major difference according to the RSM system description is the precoding block at the transmitter and the feedback data, which now covers information for this very precoding and the backup signal addressing.

As mentioned above, unitary precoding takes advantage of CSI at the transmitter. Therefore, correlated channel matrices $\bar{\mathbf{H}}_1, \ldots, \bar{\mathbf{H}}_L$ are assumed in this section. According to the model introduced in [70], the channel matrices are generated such that $\left[\bar{\mathbf{H}}_1\right]_{n,m}, \left[\bar{\mathbf{H}}_2\right]_{n,m}, \ldots$ with

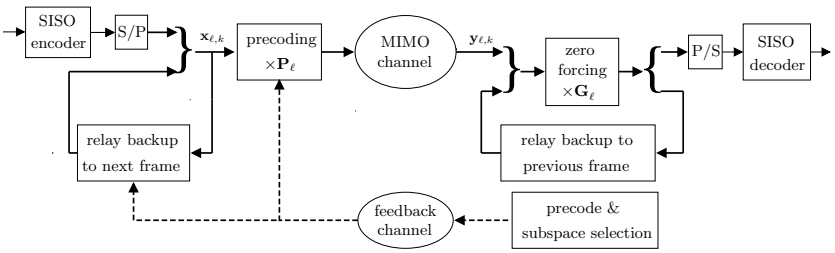

Figure 3.14 RSM transmission system extended by unitary precoding.

$1 \leq n \leq N$, $1 \leq m \leq M$ represent NM independent random sequences. For any given n, m, the sequence $\left[\tilde{\mathbf{H}}_1\right]_{n,m}, \left[\tilde{\mathbf{H}}_2\right]_{n,m}, \ldots$ represents a complex Gaussian random process with autocorrelation

$$E\left[\left[\tilde{\mathbf{H}}_\ell\right]^*_{n,m}\left[\tilde{\mathbf{H}}_{\ell+\delta}\right]_{n,m}\right] = \rho^{|\delta|} \tag{3.42}$$

with $\rho \in [0,1]$ the correlation coefficient of the elements of adjacent frames and $\delta \in \mathbb{Z}$ the lag of two considered frames.

We emphasize that the following comparison considers a *horizontal interleaving* (see Sect. 3.4) and, thus, the postprocessing SNR of each of the M signal layers is of interest. In the case of UP, a ZF at the receiver *without* any backup boosts the additive noise affecting the mth layer by a factor of $\left[(\underline{\mathbf{H}}_\ell^\dagger \underline{\mathbf{H}}_\ell)^{-1}\right]_{m,m}$ (cf. Sect. 2.4.1). Hence, using the eigenvalue decomposition of $\underline{\mathbf{H}}_\ell^\dagger \underline{\mathbf{H}}_\ell$ the postprocessing SNR at the mth signal layer becomes

$$\gamma_{\ell,m}^{(\mathrm{UP})} = \frac{\gamma_{\mathrm{inp}}}{\left[\underline{\mathbf{V}}_\ell \underline{\mathbf{S}}_\ell^{-2} \underline{\mathbf{V}}_\ell^\dagger\right]_{m,m}}. \tag{3.43}$$

Extending UP with RSM results in an improvement of the SNR by deriving the ZF estimator as

$$\mathbf{G}_\ell = \widetilde{\mathbf{H}}_\ell^+ = \begin{bmatrix} \underline{\mathbf{H}}_\ell \\ \boldsymbol{\Omega}_\ell^{-\frac{1}{2}} \mathbf{Q}_\ell^\dagger \end{bmatrix}^+. \tag{3.44}$$

Correspondingly to the factor in (3.43), the noise amplification for the mth signal layer is given by $\left[\widetilde{\mathbf{H}}_\ell^\dagger \widetilde{\mathbf{H}}_\ell\right]_{m,m}$. Again, using the eigenvalue decomposition, as expressed in (2.3), the postprocessing SNR at the mth signal layer results to

$$\gamma_{\ell,m} = \frac{\gamma_{\mathrm{inp}}}{\left[\left(\underline{\mathbf{V}}_\ell \underline{\mathbf{S}}_\ell^2 \underline{\mathbf{V}}_\ell^\dagger + \mathbf{Q}_\ell \boldsymbol{\Omega}_\ell^{-1} \mathbf{Q}_\ell^\dagger\right)^{-1}\right]_{m,m}}. \tag{3.45}$$

Since $\underline{\mathbf{V}}_\ell$ is a unitary matrix and $\mathbf{Q}_\ell$ is composed by the first m_ℓ columns

of $\underline{\mathbf{V}}_\ell$, factoring out $\underline{\mathbf{V}}_\ell$ leads to the postprocessing SNRs

$$\gamma_{\ell,m} = \frac{\gamma_{\text{inp}}}{\left[\underline{\mathbf{V}}_\ell \widetilde{\mathbf{S}}_\ell^{-2} \underline{\mathbf{V}}_\ell^\dagger\right]_{m,m}}, \quad m = 1, \ldots, M, \tag{3.46}$$

where $\widetilde{\mathbf{S}}_\ell^{-2}$ is a diagonal matrix with elements

$$\left[\widetilde{\mathbf{S}}_\ell^{-2}\right]_{m,m} = \begin{cases} \left([\underline{\mathbf{S}}_\ell]_{m,m} + \frac{\gamma_{\ell+1,m}^{(\text{ZF})}}{\gamma_{\text{inp}}}\right)^{-2} & \text{for } m \leq m_\ell \\ \left[\underline{\mathbf{S}}_\ell^{-2}\right]_{m,m} & \text{for } m > m_\ell. \end{cases} \tag{3.47}$$

The benefit from the backup has been shown above, for instance in Sect. 2.4.3 and Sect. 3.2. A corresponding comparison results from (3.43) and (3.46). While very small values in the diagonal of $\widetilde{\mathbf{S}}_\ell^2$ obviously boost the noise and thus decrease the SNRs after a ZF without any backup available, this effect is reduced by biasing the m_ℓ smallest eigenvalues in (3.47) by the backup.

Finally, the transmitter/receiver-constrained capacity is examined by performing Monte-Carlo simulations and the results of UPRSM are compared with a UP MIMO system [1] and the previously examined RSM. Therefore, sequences of 4×4 and 8×8 correlated MIMO channel matrices for different ρ with $0 \leq \rho \leq 1$ are generated as considered above. The UP constrained ergodic capacity with outdated CSI and ZF is given as

$$\bar{C}_{\text{UP}} = E\left[\sum_{m=1}^{M} \log_2\left(1 + \gamma_{\ell,m}^{(\text{UP})}\right)\right] \tag{3.48}$$

in bits per MIMO channel use. Different to (3.48), the calculations of $\bar{C}_{\text{RSM}}$ and $\bar{C}_{\text{UPRSM}}$, which are the RSM/UPRSM constrained ergodic capacities have to take into account the decreasing multiplexing gain caused by the retransmissions, such that

$$\bar{C}_{\text{UPRSM}} = E\left[\sum_{m=m_{\ell-1}+1}^{M} \log_2\left(1 + \gamma_{\ell,m}\right)\right]. \tag{3.49}$$

The simulation results are depicted in the following figures.

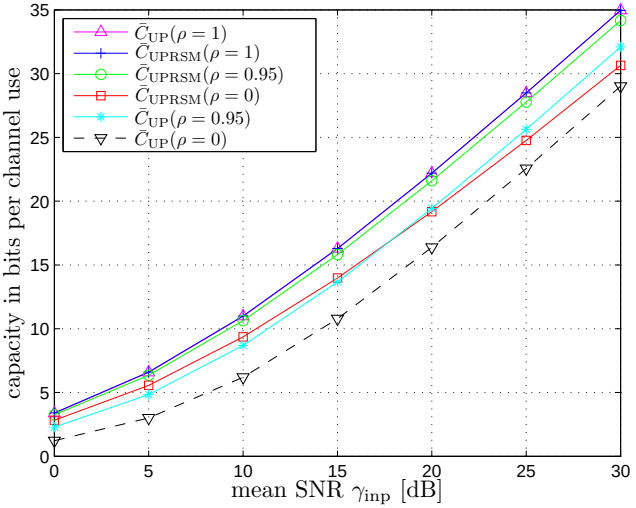

Figure 3.15 The ergodic capacity of precoding MIMO and UPRSM (both 4×4) versus SNR.

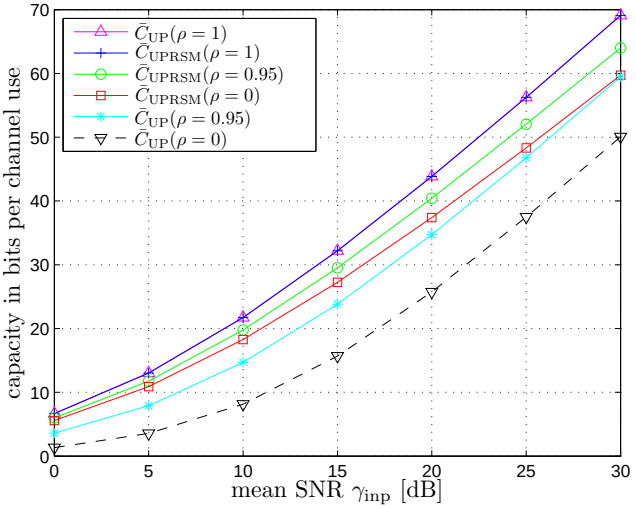

Figure 3.16 The ergodic capacity of precoding MIMO and UPRSM (both 8×8) versus SNR.

Fig. 3.15 and Fig. 3.16 show the comparison of UP, RSM and UPRSM for a 4×4 and an 8×8 MIMO channel, where $\bar{C}_{\mathrm{UP}}(\rho=0)$ and $\bar{C}_{\mathrm{UP}}(\rho=1)$ reflect the lower and upper performance limits, respectively. In between these limits, the achievable ergodic capacities of UPRSM for different correlation coefficients are plotted. In the case $\rho=0$, consecutive channels are uncorrelated and thus unitary precoding does not offer any benefit. In fact, the results of $\bar{C}_{\mathrm{UPRSM}}(\rho=0)$ are equal to the ergodic capacity of an RSM system (compare Sect. 3.4).

In the case of perfect CSI at the transmitter ($\rho=1$), the UPRSM constrained capacity is equal to that of UP, and UPRSM offers an improvement of approximately $2.5-3$ dB compared to RSM ($\bar{C}_{\mathrm{UPRSM}}(\rho=0)$). Fig. 3.17 and Fig. 3.18 show the comparison of the three different systems as the capacity is plotted versus ρ at a fixed SNR to demonstrate the different behavior of the systems in a correlated fading environment.

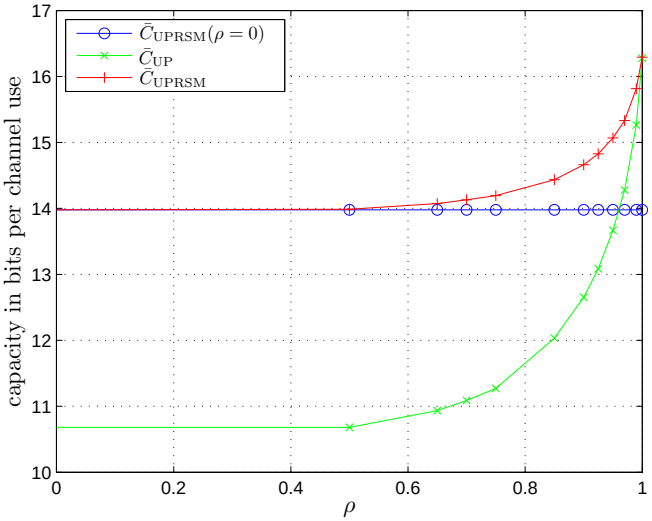

Figure 3.17 Comparison of the ergodic capacity of 4×4 MIMO transmission with UP, RSM and UPRSM for different channel correlation coefficients ρ and an SNR of $\gamma_{\mathrm{inp}}=15$ dB.

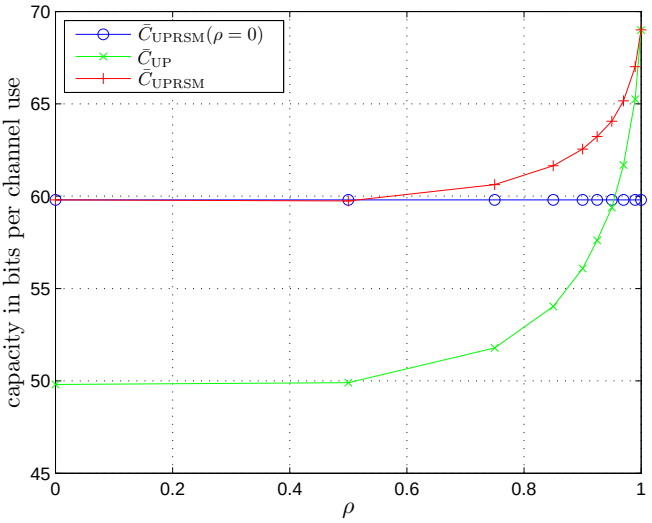

Figure 3.18 Comparison of the ergodic capacity of 8×8 MIMO transmission with UP, RSM and UPRSM for different channel correlation coefficients ρ and an SNR of $\gamma_{\mathrm{inp}} = 30$ dB.

While Fig. 3.17 gives the results for an SNR of $\gamma_{\mathrm{inp}} = 15$ dB in case of a (4×4) MIMO channel, the 8×8 MIMO channel in Fig. 3.18 is considered for an SNR of $\gamma_{\mathrm{inp}} = 30$ dB.

The leftmost value $\bar{C}_{\mathrm{UP}}(\rho = 0)$ is the transceiver constrained capacity per channel use without any CSI at the transmitter, since channel realizations of subsequent frames are considered to be uncorrelated. By increasing ρ, $\bar{C}_{\mathrm{UP}}$ is increasing because of the rising accuracy of the CSI at the transmitter. While $\bar{C}_{\mathrm{RSM}}$ is invariant to variations of ρ (see Sect. 2.4.3), UPRSM can exploit the CSI at the transmitter for increasing ρ to enhance the performance of the system. The relative impact of RSM decreases with increasing ρ as can be seen from the two curves $\bar{C}_{\mathrm{UP}}$ and $\bar{C}_{\mathrm{UPRSM}}$ approaching each other for $\rho \to 1$.

The latter aspect is also demonstrated in Fig. 3.19, where for a total

of L frames a plot of the average backup size

$$\bar{m} = \frac{1}{L} \sum_{\ell=1}^{L} m_\ell \tag{3.50}$$

versus the correlation coefficient ρ is drawn for an SNR of $\gamma_{\text{inp}} = 30$ dB. Obviously, the necessity of a retransmission decreases with increasing ρ. In case of perfect channel knowledge at the transmitter ($\rho = 1$), the average backup size $\bar{m} = 0$.

As a result, the combination of UP and RSM capitalizes on the advantages of both schemes for different values of ρ, while the amount of additional feedback is very low. For unitary precoding, the matrix $\underline{\mathbf{V}}_\ell$ is fed back from the receiver to the transmitter end comprising the eigenvectors of $\underline{\mathbf{H}}_\ell^\dagger \underline{\mathbf{H}}_\ell$, whereas in RSM several columns of this very matrix are used at the transmitter to form a backup signal. Hence, the required

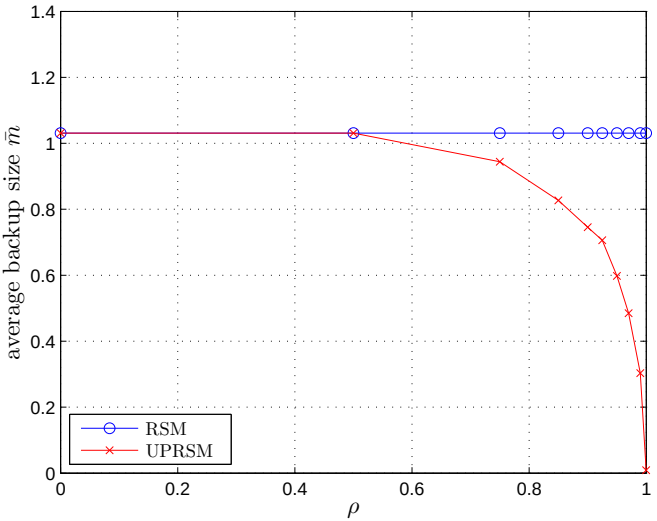

Figure 3.19 Comparing the average backup size $\bar{m}$ of UPRSM and RSM (both 8×8) for varying ρ and an SNR $\gamma_{\text{inp}} = 30$ dB.

feedback channel capacity is comparable to a conventional UP system. More precisely, the feedback channel needs additional capacity to signal $M+1$ states, namely for $m_\ell \in \{0, 1, \ldots, M\}$, along with the CSI for the precoding, noting the latter consumes considerable more bandwidth to be effective (cf. Sect. 3.3 and [5]).

3.6. Higher Order MIMO Systems

The performance analysis focuses on the maximum capacity per channel use for a certain policy at the receiver. The policies φ_1 and φ_2 introduced in Sect. 3.1, which depend on a certain threshold ξ, are assumed to optimize this parameter in order to maximize the ergodic capacity. However, it has not yet been clarified how to find such an optimal threshold.

Considering a MIMO channel matrix $\mathbf{H}_\ell$ with zero-mean complex Gaussian distributed elements, the random matrix $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ is known as a Wishart matrix [71]. The joint probability density function of the eigenvalues $\lambda_1^2 \leq \lambda_2^2 \leq \ldots \leq \lambda_M^2$ is given in [72] as

$$f(\lambda) = \frac{2^{-MN}\pi^{M(M-1)}}{\widetilde{\Gamma}_M(N)\,\widetilde{\Gamma}_M(M)} \exp\left(-\frac{1}{2}\sum_{i=1}^{M}\lambda_i\right) \prod_{i=1}^{M} \lambda_i^{N-M} \prod_{i>j} (\lambda_i - \lambda_j)^2 , \tag{3.51}$$

with the multivariate gamma function defined by

$$\widetilde{\Gamma}_N(a) = \pi^{N(N-1)/2} \prod_{i=1}^{N} \Gamma(a-i+1) \tag{3.52}$$

and $\Gamma(\cdot)$ the gamma function, [51]. Note that the expectation of the multiplexing gain $g = E[g_\ell]$ in (3.25) depends on (3.51) and the applied policy. In the following, we focus on policy φ_2, that is, the result m_ℓ of this policy depends on the random vector $\mathbf{z}_\ell$ and the choice of the threshold ξ. Thus, to point out this dependency, we introduce the notation $\varphi_2(\mathbf{z}_\ell, \xi)$. The instantaneous noise amplification in (3.19), where

$$W_\ell = \frac{1}{M}\left(\sum_{i=1}^{\varphi_2(\mathbf{z}_\ell,\xi)} \frac{1}{z_{\ell,i} + W_{\ell-1}^{-1}} + \sum_{i=\varphi_2(\mathbf{z}_\ell,\xi)+1}^{M} \frac{1}{z_{\ell,i}}\right), \tag{3.53}$$

illustrates the difficulties in finding a threshold value ξ that minimizes $E[W_\ell]$. The left hand side of (3.53) is a recursive function of random variables with a joint probability density function given in (3.51). The number of summands involved in the recursion again is calculated by the policy, for which the result in turn depends on the random eigenvalues and the threshold that shall be optimized. Instead of trying to find a generic analytical solution of this problem, applicable for different values of M, N, φ and γ_{inp}, a numerical assessment of ξ is introduced in the following.

The assessment presumes that the ergodic capacity $\bar{C}_{\mathrm{RSM}}$ for a given SNR γ_{inp} is a strictly concave function depending on the threshold ξ. This assumption is based on the fact that the multiplexing gain g is a monotonically decreasing function, while the term $\log_2(1+\gamma_{\mathrm{inp}}/S^{(\varphi)})$ is monotonically increasing for an increasing threshold ξ. Monte-Carlo simulations for a bunch of different values of ξ indicate that this assumption is valid (an analytical proof is missing, due to the aforementioned complexity). Fig. 3.20 depicts the result of such simulation, where $\bar{C}_{\mathrm{RSM}}$ is

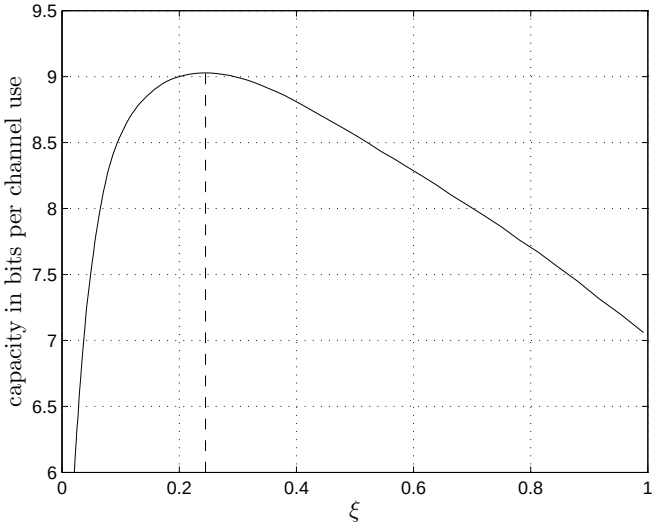

Figure 3.20 Assessment of threshold ξ_{opt} (dashed line) that maximizes $\bar{C}_{\mathrm{RSM}}$ in a 4×4 MIMO channel at an SNR $\gamma_{\mathrm{inp}} = 10\,\mathrm{dB}$.

shown as a function of ξ in a 4×4 MIMO channel at an SNR $\gamma_{\text{inp}} = 10\,\text{dB}$.

Following this approach for different SNRs and different MIMO channel sizes, a look-up table can be used at the receiver. Supposing that the receiver has adequate knowledge of the preprocessing SNR γ_{inp}, the look-up table provides a ξ_{opt} that maximizes the *ergodic* capacity per channel usage. A graphical representation of such a table is given in Fig. 3.21.

Obviously, as the number of transmit and receive antennas in the transmission system increases the optimal threshold changes due to different underlying joint probability density functions for the eigenvalues of $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$. Further, with increasing size of $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$, the assessed value of ξ_{opt} also increases. The results plotted in Fig. 3.21 raise the question, whether ξ_{opt} is bounded for a fixed γ_{inp}, since the difference of adjacent curves in Fig. 3.21 decreases for increasing MIMO channel dimensions.

Interestingly, the case of a MIMO channel with unlimited numbers of transmit/receive antennas can be observed by means of random matrix theory, e.g. [73]. A central result in random matrix theory states the Marcenko-Pastur law, i.e., if $\mathbf{H}_\ell^\dagger\mathbf{H}_\ell$ is a Wishart matrix, the empirical distribution of the eigenvalues $\lambda^2 = \left(\lambda_1^2(\mathbf{H}_\ell), \lambda_2^2(\mathbf{H}_\ell), \ldots, \lambda_M^2(\mathbf{H}_\ell)\right)$ con-

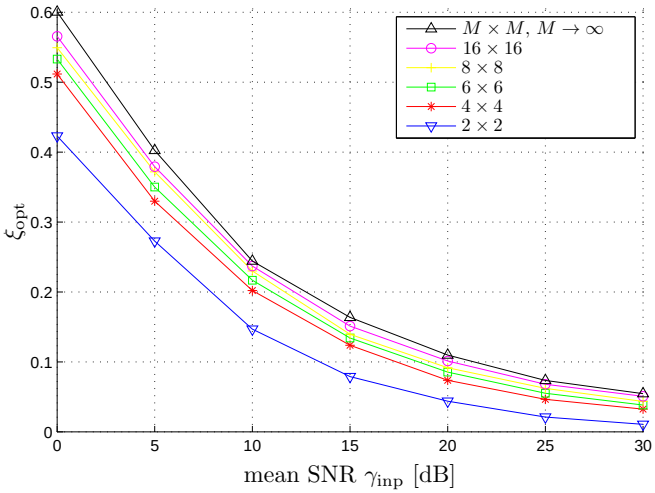

Figure 3.21 Numerical results of the optimal threshold for different number of antennas.

verges almost surely, as $M, N \to \infty$ with $\frac{M}{N} \to \beta$ to the density function

$$f_\beta(\lambda^2) = \left(1 - \frac{1}{\beta}\right)^+ \delta(\lambda^2) + \frac{\sqrt{(\lambda^2 - a)^+ (b - \lambda^2)^+}}{2\pi\beta\lambda^2} \tag{3.54}$$

with $(z)^+ = \max(0, z)$ and

$$a = (1 - \sqrt{\beta})^2 \qquad b = (1 + \sqrt{\beta})^2 \tag{3.55}$$

Assume $\mathbf{H}_\ell$ to be a square matrix with $N = M \to \infty$, i.e., there are as many receive antennas as transmit antennas for a communication link. In this case, $\beta = 1$ and (3.54) simplifies to

$$f_1(\lambda^2) = \frac{\sqrt{\lambda^2(4 - \lambda)}}{2\pi\lambda^2} = \frac{1}{\pi}\sqrt{\frac{1}{\lambda^2} - \frac{1}{4}}. \tag{3.56}$$

As it is shown for the case of finite MIMO dimensions, the noise amplification in (3.53) is a recursive function. In the case of $M, N \to \infty$, (3.56) simplifies the search for an optimal threshold. Of course, the maximal multiplexing gain for a MIMO channel with unlimited dimensions is equal to infinity. However, taking into account the backup signals "consume" a part of these infinite number of layers, the multiplexing gain *ratio* for the density function given in (3.56) can be expressed by

$$g^{(\varphi_2,\infty)} = \frac{1}{\pi}\int_\xi^4 f_1(\lambda^2)\mathrm{d}\lambda^2 = \frac{1}{\pi}\int_\xi^4 \sqrt{\frac{1}{\lambda^2} - \frac{1}{4}}\mathrm{d}\lambda^2 \tag{3.57}$$

where (3.57) equals one if the threshold $\xi = 0$. Using the tables in [74], it can be easily seen that (3.57) results to

$$g^{(\varphi_2,\infty)} = \frac{1}{2\pi}\left(\pi - \sqrt{(4 - \xi)\,\xi} + 2\arcsin\left(\frac{2 - \xi}{2}\right)\right). \tag{3.58}$$

Applying (3.56) to the recursive function of the noise amplification, we

obtain

$$W = \int_0^{\xi} \frac{f_1(\lambda^2)}{\lambda^2 + W^{-1}} \mathrm{d}\lambda^2 + \int_{\xi}^{4} \frac{f_1(\lambda^2)}{\lambda^2} \mathrm{d}\lambda^2 \tag{3.59}$$

$$= \frac{1}{\pi} \left(\int_0^{\xi} \frac{1}{\lambda^2 + W^{-1}} \sqrt{\frac{1}{\lambda^2} - \frac{1}{4}} \mathrm{d}\lambda^2 + \int_{\xi}^{4} \frac{1}{\lambda^2} \sqrt{\frac{1}{\lambda^2} - \frac{1}{4}} \mathrm{d}\lambda^2 \right) \tag{3.60}$$

$$= \frac{1}{\pi} \left(-\frac{\pi}{2} + \sqrt{\frac{4}{\xi} - 1} + \sqrt{1 + 4W} \arctan\left(\sqrt{\frac{\xi\,(1 + 4W)}{4 - \xi}} \right) \right) \tag{3.61}$$

while assuming W is stationary for an $M \times M$ matrix $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ with $M \to \infty$. The Banach fixpoint theorem [75] can be used to solve (3.61) numerically as shown in appendix A.1. Although the probability density function in (3.56) seems to offer the possibility of finding an optimal threshold in this case, the resulting terms in (3.58) and (3.61) illustrate that solving the equation

$$\xi_{\mathrm{opt}}(\gamma_{\mathrm{inp}}) = \arg\max_{\xi} g^{(\infty)}(\xi) \log_2 \left(1 + \frac{\gamma_{\mathrm{inp}}}{W(\xi)} \right) \tag{3.62}$$

is still a difficult task. For that reason, again a numerical assessment of $\xi_{\mathrm{opt}}(\gamma_{\mathrm{inp}})$ has been performed. The results shown in Fig. 3.21 suggest that the optimal threshold ξ_{opt} in order to maximize the ergodic capacity is upper bounded for a fixed γ_{inp} when the number of transmit and receive antennas increases.

Chapter 4.

Non-Ergodic RSM Performance

While in Chapt. 3 the ergodic capacity of RSM has been examined, this chapter focuses on the non-ergodic performance of the approach. We point out that the recursive nature of the RSM transmitter/receiver architecture, as described in Sect. 2.4.3, makes it impossible to guarantee a maximal delay. Since the instantaneous capacity at the time when transmitting the ℓth frame depends on the active as well as on the previous MIMO channel state, a measure of the performance per frame is unreasonable. This aspect becomes apparent recalling that the RSM constrained capacity gain is attained by supporting preceding frames through the retransmission of backup signals.

For this reason, the RSM constrained capacity for a *limited* number of frames is investigated. Further, for the purpose of assessing an appropriate frame group size, we estimate the mean delay expressed in number of frames of an RSM communication system with no frame limitation. Thereafter, optimal strategies for slowly time-varying MIMO channels are introduced, which have been found by means of reinforcement learning.

4.1. Limited Number of Frames

A simple yet efficient method to guarantee a maximal delay for transmitting the information of a specific frame is to send data with a limited number of frames. That is, the RSM architecture in Sect. 2.4.3 is changed in a way that the transmitter always sends information in groups of frames. Thus, the delay is limited to a maximum of the number of frames in a transmission group. In the following, we examine two different approaches on how to process the frames in a group.

One system approach to guarantee a maximal delay is to simply interrupt the receiver's requests for backup signals after a maximum number of frames have been involved in the recursive process. When the receiver processes the last frame of a group, no retransmission will be requested regardless of the current CSI and applied policy. Having received the last frame of a group, the receiver immediately starts to decode, i.e., the linear signal reconstruction. Doubtlessly, this potential violation of the policy at the last frame leads to a backward error propagation, because the SNR γ_{dec} at this very frame and with it the SNR of the backup signal for the previously received frame cannot be lower bounded.

Note that the previously described transmission in groups of frames does not exclude the occurrence of frames within this group with $m_\ell = 0$, i.e., the receiver might experience frames which start the decoding process (since $m_\ell = 0$) before the last frame of the group. Despite of these incidences, the recursive multiplexing starts every Lth frame, with L the number of frames in a group. Fig. 4.1 depicts the RSM constrained

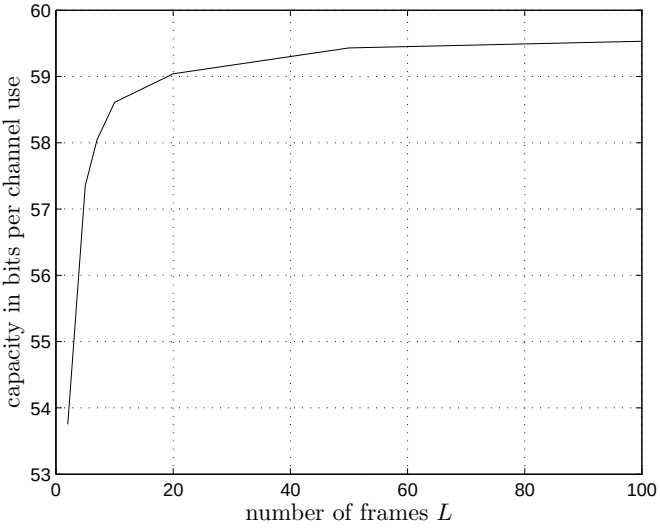

Figure 4.1 RSM constrained ergodic capacity in bits per channel use with limited number of frames in an 8×8 Rayleigh channel with $\gamma_{\mathrm{inp}} = 30\,\mathrm{dB}$.

capacity with limited number of frames L and the aforementioned potential violation as results of multiple Monte-Carlo simulations for different values L. The 8×8 MIMO channel is assumed to have zero-mean independent Gaussian distributed entries with $\rho = 0$ and $\gamma_{\text{inp}} = 30\,\text{dB}$.

It is apparent that the previously described strategy violates policy φ_3, since an SNR $\gamma_{\text{dec}} \geq \gamma_{\text{inp}}/\breve{\varrho}$ cannot be assured. However, a lower bound of the system capacity can be provided by a slight change in the aforementioned approach. This second method described below does not violate a given policy at the last frame of a group.

Instead of starting to decode and possibly neglecting a given policy after a maximum number of transmitted frames, the number of consecutive non-decodable frames F is considered as a criterion. According to policy φ_3, the decoding process only starts if no retransmission is requested. That is, to provide a guaranteed SNR at the decoder, outages may occur, if a frame still requires a backup while the maximum number of non-decodable frames is reached. Fig. 4.2 and Fig. 4.3 show the lower bound of the RSM constrained capacity for varying number of limited frames versus the SNR in a 4×4 and 8×8 MIMO channel, respectively.

It can be seen that a reduction of the allowed number of retransmissions leads to a dramatic decrease of the achievable capacity. Even for large values of possible retransmissions, e.g. $F = 50$ in a 4×4 MIMO channel, a clear gap of $\approx$ 2 dB to the performance of an unrestricted system ($F \to \infty$) remains in the low SNR region.

For increasing number of signal layers, the performance gap between transmissions with a small number of allowed retransmissions and transmissions with unlimited retransmission increases massively, as illustrated in Fig. 4.3 for an 8×8 MIMO channel. Comparing these results with the capacities depicted in Fig. 3.8 it becomes apparent that an LMMSE/ARQ based 8×8 MIMO system outperforms RSM with $F = 50$ and even RSM with $F = 150$ for $\gamma_{\text{inp}} \leq 13$ dB. In the following section a simple approximation of the transmission delay in an unrestricted RSM system is given. Thereafter, a possible solution is suggested to overcome the massive performance degradation when the delay/memory in an RSM architecture is limited.

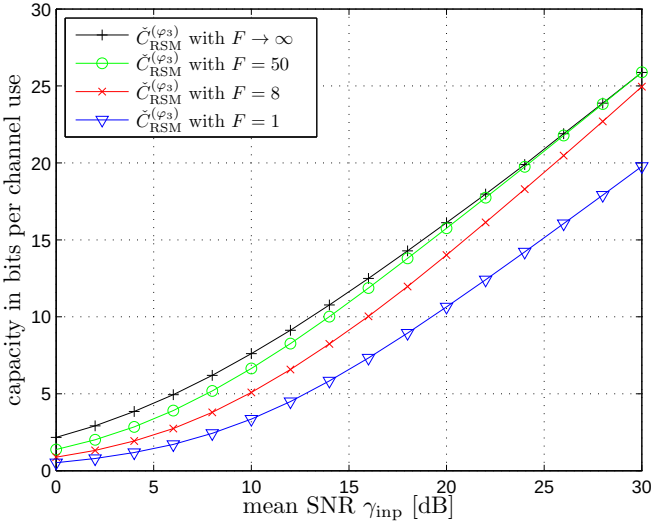

Figure 4.2 RSM constrained capacity per channel use with limited number of retransmissions in 4×4 MIMO Rayleigh channel.

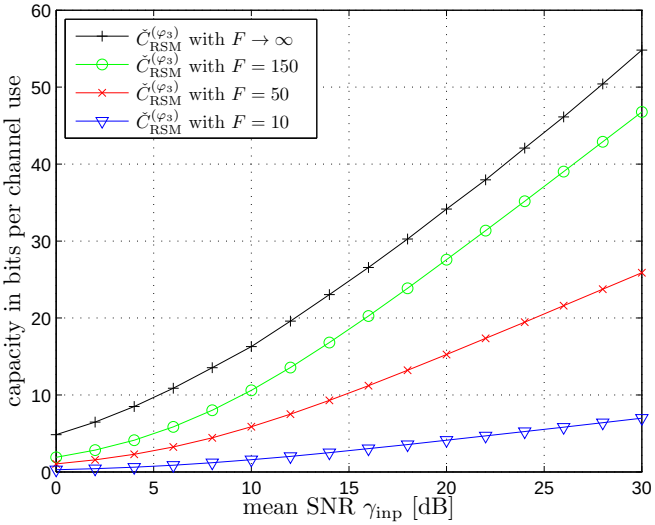

Figure 4.3 RSM constrained capacity per channel use with limited number of retransmissions in 8×8 MIMO Rayleigh channel.

4.2. Estimating the Delay

As explained in Sect. 2.4.3, the decoding process starts as soon as a received frame does not need a backup, which could be provided by a subsequent frame. Since it is assumed in the system approach that the channel is varying arbitrarily from frame to frame, the case of $m_\ell = 0$ is unpredictable. That is, the instance of starting the decoding process cannot be forseen and the transmission delay[1] is a random variable.

In case of a fixed threshold based policy, e.g. policy φ_2, an expectation of the mean delay can be calculated as follows. Still, the MIMO channel is considered as a non-zero random matrix with independent zero-mean complex Gaussian elements. The estimate of the mean delay is based on the PDF of the smallest eigenvalue $\lambda_1^2(\mathbf{H}_\ell)$ for a complex valued central Wishart matrix $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$ given in [72] as

$$f_{\lambda_1^2}(z) = \frac{M}{2}\, \mathrm{e}^{-\frac{M}{2}z}\,. \tag{4.1}$$

Again, we focus on policy φ_2 for simplicity. The probability of receiving a frame in a situation, where no retransmission is necessary, is equivalent to the probability that the smallest eigenvalue $\lambda_1^2(\mathbf{H}_\ell) \geq \xi$. This can be easily calculated by

$$\Pr\left[\lambda_1^2(\mathbf{H}_\ell) \geq \xi\right] = \int_\xi^\infty \frac{M}{2}\, \mathrm{e}^{-\frac{M}{2}\lambda_1^2}\, d\lambda_1^2 = \mathrm{e}^{-\frac{M}{2}\xi}\,. \tag{4.2}$$

The transmission delay depending on the threshold ξ to maximize the capacity and expressed in number of received frames can be calculated by applying the following Bernoulli experiment. Denoting X as the number of events when $\lambda_1^2(\mathbf{H}_\ell) \geq \xi$, the probability $\Pr[X \geq 1] = 1 - \Pr[X = 0]$ can be calculated with

$$\Pr[X \leq x] = \sum_{i \leq x} \binom{L}{i} p^i (1-p)^{L-i} \tag{4.3}$$

[1] The transmission delay is defined as the time from receiving a frame to the instance of decoding the content of this very frame.

denoting the binomial distribution and p the probability of each event $\lambda_1^2(\mathbf{H}_\ell) \geq \xi$ as denoted in (4.2). Referring to (4.2) the probability for at least one occurring event in a group of L frames can be defined as

$$\Pr[X \geq 1] = 1 - \Pr[X = 0] = 1 - (1 - \mathrm{e}^{-\frac{M}{2}\xi})^L. \tag{4.4}$$

The latter part $\Pr[X = 0]$ of (4.4) is the outage probability P_{out} for a given threshold ξ and a given frame group size L. Fig. 4.4 and 4.5 show contour plots of the three-dimensional function given in (4.4) for a 4×4 and 8×8 MIMO system, respectively. Each line represents an example for a required probability $\Pr[X \geq 1]$, that is, the probability to decode the message within a group of L frames. Referring to Fig. 4.4, for a given threshold $\xi = 1$ at least 21 frames have to be transmitted to ensure an outage probability $P_{\text{out}} \leq 0.05$.

Rearranging (4.4), different cost functions depending on the given parameters can be defined, where

$$\xi' \leq -\frac{2}{M} \ln\left(1 - \sqrt[L]{P_{\text{out}}}\right) \tag{4.5}$$

provides an upper border of the threshold for a chosen maximal outage probability P_{out} and a predefined maximum delay in number of transmitted frames. If ξ is fixed, e.g. it is only defined for maximizing the capacity and thus independent of the delay, the maximum delay L' can be estimated by

$$L' = \frac{\ln(P_{\text{out}})}{\ln\left(1 - \mathrm{e}^{-\frac{M}{2}\xi}\right)}. \tag{4.6}$$

Clearly, the choice of ξ determines the available postprocessing SNR $\gamma_{\text{dec}} = \gamma_{\text{inp}}\xi$ (cf. Sect. 3.1).

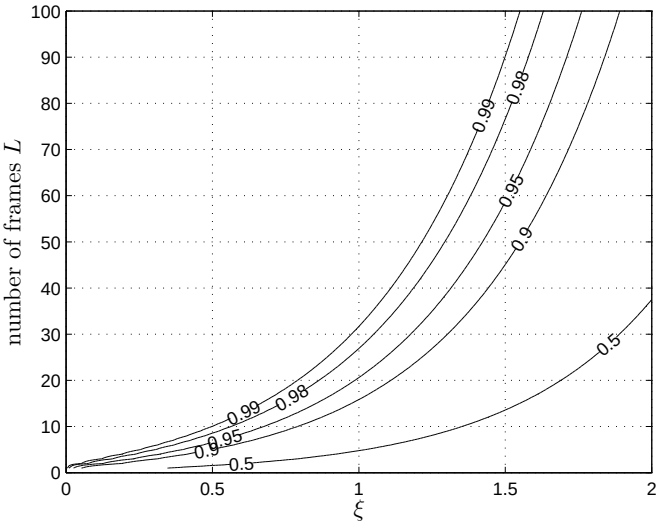

Figure 4.4 Expected number of frames versus threshold ξ for given outage probabilities P_{out} for a 4×4 MIMO channel.

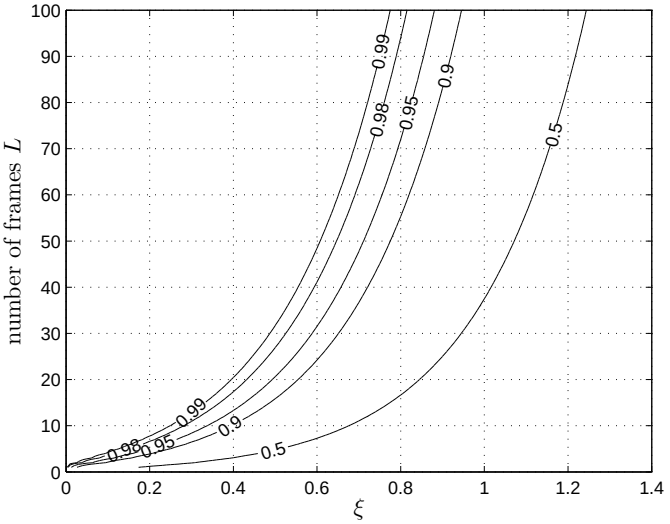

Figure 4.5 Expected number of frames versus threshold ξ for given outage probabilities P_{out} for an 8×8 MIMO channel.

4.3. Antenna Selection Strategies

The distribution of the smallest eigenvalue of a Wishart matrix is sensitive to the ratio of $(M/N) \leq 1$ [72]. The lower this ratio, the higher the probability that the smallest eigenvalue is greater than a considered threshold. In the case of RSM this yields $m_\ell = 0$, i.e., no retransmission in the subsequent frame is necessary. Exploiting this fact leads to an approach where the transmitter may use only a_ℓ out of M transmit antennas for the ℓth frame transmission as shown in Fig. 4.6. The rest of the transmitter/receiver architecture corresponds to the RSM system with limited number of frames as described in Sect. 4.1. Hence, $a_\ell - m_{\ell-1}$ layers are provided for serial/parallel converted signals from the SISO encoder, while $m_{\ell-1}$ dimensions are used for the transmission of the backup of the previous frame.

Since the MIMO channel is considered as an $N \times a_\ell$ matrix $\mathbf{H}_\ell$, the dimension of the backup signal addressing matrix $\mathbf{Q}_\ell^\dagger$ changes to $m_\ell \times a_\ell$. A vertical interleaving and deinterleaving (cf. Sect. 3.4) is applied, represented by $\mathbf{\Pi}$ and $\mathbf{\Pi}^{-1}$, respectively. Therefore, it is sufficient to multiply the backup by a *scalar* factor $\sqrt{\alpha_\ell}$, adapting the variance of the additive noise terms in the backup to ϑ_{noise}. Further, due to the transmit antenna selection, the backup size policy φ_3 described in (3.7) results to

$$m_\ell = \min\left\{m \in \{0, \ldots, a_\ell\} \;:\; \varrho_\ell\left(m, a\right) \leq \breve{\varrho}\right\} \tag{4.7}$$

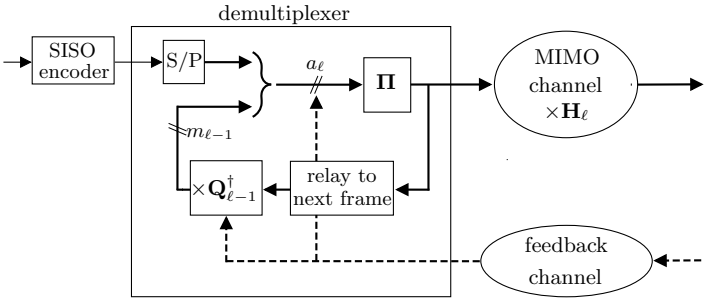

Figure 4.6 RSM transmitter with antenna selection strategies.

with the noise amplification $\varrho_\ell\,(m,a)$, also depending from the number of used transmit antennas a, defined as

$$\varrho_\ell\,(m,a) = \frac{1}{a_\ell}\left(\sum_{i=1}^{m}\frac{1}{\lambda_i\,(\mathbf{H}_\ell)+\alpha_\ell}+\sum_{i=m+1}^{a_\ell}\frac{1}{\lambda_i\,(\mathbf{H}_\ell)}\right). \tag{4.8}$$

The entries of $\mathbf{H}_\ell$ are considered as independent, zero-mean complex Gaussian distributed random values with variance $1/a_\ell$, which results in independent Rayleigh fading. The normalization by $1/a_\ell$ accounts for a uniform allotment of the fixed total transmit power to the a_ℓ active transmit antennas, with the result that γ_{inp} reflects the mean SNR at every receive antenna. Regarding $\mathbf{H}_1, \mathbf{H}_2, \ldots$ as independent random matrices, $p_{m_\ell}(m; a, \breve{\varrho})$ denotes the probability of $(m_\ell = m)$ given that $(a_\ell = a)$ transmit antennas are actively used, as determined by (4.7). As an example, Tab. 4.1 shows the incidence of the backup signal dimension m_ℓ for the case of $N = 8$ and $\breve{\varrho} = 2$, i.e., assuming the SISO decoder requires an SNR of $\gamma_{\text{dec}} \geq \gamma_{\text{inp}}/2$. The probabilities as found in Monte-Carlo simulations are rounded to two decimal places.

a_ℓ :	1	2	3	4	5	6	7	8
$p_{m_\ell}(0; a_\ell, \breve{\varrho})$:	1.00	1.00	1.00	0.98	0.78	0.26	0.01	0.00
$p_{m_\ell}(1; a_\ell, \breve{\varrho})$:	0.00	0.00	0.00	0.02	0.22	0.72	0.66	0.10
$p_{m_\ell}(2; a_\ell, \breve{\varrho})$:		0.00	0.00	0.00	0.00	0.02	0.33	0.81
$p_{m_\ell}(3; a_\ell, \breve{\varrho})$:			0.00	0.00	0.00	0.00	0.00	0.09
$p_{m_\ell}(4; a_\ell, \breve{\varrho})$:				0.00	0.00	0.00	0.00	0.00

Table 4.1. Distribution of the number of signal dimensions m_ℓ to be covered by the backup for limiting the noise amplification to a factor $\breve{\varrho} = 2$ in the case of Gaussian $8\times a_\ell$ MIMO channels.

The results in Tab. 4.1 suggest that the probability of achieving $\gamma_{\text{dec}} \geq \gamma_{\text{inp}}/\breve{\varrho}$ without backup (i.e. $m_\ell = 0$) with only up to three transmit antennas is close to 1, which is due to the diversity order. On the other hand, using all eight transmit antennas yields a request for a two-dimensional backup in most cases, while the chances for attaining the target SNR without backup is poor. Without delay and memory restric-

tions, the optimal policy with respect to the data rate is to always use all available transmit antennas (i.e. $a_\ell = M$), as examined in Sect. 3.2. However, the results in Sect. 4.2 suggest that in most applications the memory requirements of this policy are unacceptably large. Hence, the focus of this section is on finding adequate policies for choosing a_ℓ in RSM transmission systems with limited memory.

To this end, the triplet $s_\ell = (m_\ell, b_\ell, f_\ell)$ shall represent the state of the multiplexer after reception of the ℓth frame and the processing steps possible at this time (see also Fig. 2.7). The two latter variables of the triplet reflect parameters of the LIFO in Fig. 2.7. The accumulated amount of transmitted signals waiting to be decoded is denoted by b_ℓ. The number of frame signal observations stored in the memory is given by f_ℓ, which is also a gauge for the "charging level" of the LIFO memory. If no backup is required, the processing of the current frame and all stored signal observations can be completed, hence, $m_\ell = 0$ implies $b_\ell = 0$ and $f_\ell = 0$. Otherwise, if $m_\ell > 0$,

$$b_\ell = \begin{cases} b_{\ell-1} + a_\ell - m_{\ell-1} & \text{if } f_{\ell-1} < F \\ a_\ell - m_{\ell-1} & \text{if } f_{\ell-1} = F \end{cases}, \tag{4.9}$$

$$f_\ell = \begin{cases} f_{\ell-1} + 1 & \text{if } f_{\ell-1} < F \\ 1 & \text{if } f_{\ell-1} = F \end{cases}, \tag{4.10}$$

with F the number of frames that can be stored in the memory. In the case $f_{\ell-1} < F$, i.e., the memory still offers space for capturing the ℓth frame, the amount of signals waiting to be decoded increases by $a_\ell - m_{\ell-1}$ signal layers, corresponding to the number of layers transmitting new data from the SISO encoders. If the memory has reached its capacity, such that $f_{\ell-1} = F$, and $m_\ell > 0$, the memory is cleared *before* processing the ℓth frame, in which case the data encoded in all previously stored signals is lost and thus $b_\ell = a_\ell - m_\ell$ and $f_\ell = 1$. Fig. 4.7 illustrates the various states and possible transitions described above.

According to Tab. 4.1, a reasonable choice of the number of active transmit antennas a_ℓ depends on the state $s_{\ell-1}$. Thus, the information a_ℓ is fed back from the receiver to the transmitter based on the strategy

$$\phi : s_{\ell-1} \mapsto a_\ell = \phi(s_{\ell-1}) \in \{1, \ldots, M\}\,. \tag{4.11}$$

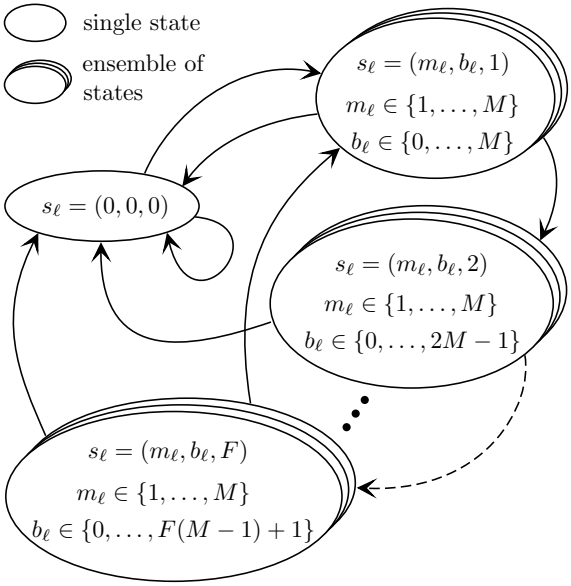

Figure 4.7 Finite-state machine representation of the Markov decision process. Some of the plotted state transitions, which represent the possible actions, may not appear under an optimal policy.

Defining such an antenna strategy ϕ yields the random process $s_1, s_2, \ldots$ which represents a *Markov chain*. The necessary state transition probabilities for this Markov chain can be found by Monte-Carlo simulations, as shown in Tab. 4.1. For instance, the probability of a state transition $s_{\ell-1} \mapsto s_\ell$ can be found in row $m_\ell + 1$ and column $\phi(s_{\ell-1})$ of Tab. 4.1. We emphasize that the Markov chain $s_1, s_2, \ldots$ is the result of a random process in combination with a controlling element ϕ. Finding an optimal antenna strategy, thus, equals solving an optimal control problem.

A promising concept for this problem is the theory of *Markov decision processes* (MDPs) introduced in [33]. The finite set of states, defined in (4.9) and (4.10), and the finite set of actions, given by (4.11) are two major components of a finite MDP. Further, the probability for a state transition from a current state s to a next state s' following the action

$\phi(s_\ell) = a$ is denoted by [32]

$$\mathcal{P}^a_{ss'} = \Pr\left[s_{\ell+1} = s' \,|s_\ell = s,\ \phi(s_\ell) = a\,\right]. \tag{4.12}$$

In the present RSM setup sketched in Fig. 4.6, the transition probability depends on the MIMO channel, more precisely on the occurrence of m_ℓ, thus $\mathcal{P}^a_{ss'} = p_{m_\ell}(m; a, \check{\varrho})$. The basic idea of an MDP is to define an *immediate reward* for any state transition, such that the expected value of the next reward is [32]

$$\mathcal{R}^a_{ss'} = E\left[R(s_\ell, s_{\ell+1}; a)\,|s_\ell = s,\ \phi(s_\ell) = a,\ s_{\ell+1} = s'\,\right], \tag{4.13}$$

where in the present case the reward is quantified by

$$R(s_{\ell-1}, s_\ell; a) = \begin{cases} b_{\ell-1} + a - m_{\ell-1} & \text{if } m_\ell = 0 \\ 0 & \text{if } m_\ell > 0 \end{cases} \tag{4.14}$$

and reflects the number of signal layers (in the LIFO) that can be decoded after the reception of the ℓth frame. If the current frame necessitates a backup, the reward is zero. Bellman's principle of optimality [76] provides a method for computing strategies with maximal expected *long term reward* denoted as a value function for each possible state. We define the expected reward $V^\phi(s)$ of state s for the strategy ϕ, such that

$$V(s) = E\left[\sum_{i=0}^{\infty} \eta^i R(s_{\ell+i}, s_{\ell+1+i}; a)\,\middle|\, s_\ell = s\right], \tag{4.15}$$

where η is the discount rate, with $0 \leq \eta \leq 1$, which determines the present value of future rewards. If $\eta = 0$ the controller tries to maximize the immediate reward $R(s, s', a_\ell)$ only. The closer η to one, the more farsighted the controller becomes and takes future rewards into account more intensely [32]. Finding an optimal strategy is equivalent to searching for a strategy that maximizes the long term reward in (4.15). The

corresponding strategy can be easily obtained as shown below. Since (4.15) is a recursive function, an iterative procedure helps solving it. The field of dynamic programming [76] provides standard iterative procedures for these problems. Here, the *value iteration* algorithm is applied by successively solving

$$V_{i+1}(s) = \max_a \sum_{s'} \mathcal{P}^a_{ss'} \left(\mathcal{R}^a_{ss'} + \eta V_i(s') \right) \tag{4.16}$$

for all s in the set defined by (4.9) and (4.10). It can be shown that the sequence $V_0, V_1, \ldots$ converges to an optimal value function for an arbitrary starting value V_0. Applied to the present RSM problem, for a required SNR $\gamma_{\mathrm{inp}}/\check{\varrho}$ at the decoder input and a given memory size F, the resulting optimal strategy $\phi^{(\check{\varrho},F)}_{\mathrm{opt}}$ maximizes the signal rate at the decoder input. Choosing η close to 1, the expected long term reward achieved by $\phi^{(\check{\varrho},F)}_{\mathrm{opt}}$ corresponds to the multiplexing gain of the RSM system, which is therefore denoted as $g(\phi^{(\check{\varrho},F)}_{\mathrm{opt}})$.

According to the incidences of backup dimensions displayed in Tab. 4.1, a subsequence of actions $a_\ell = \phi^{(\check{\varrho},F)}_{\mathrm{opt}}(s_{\ell-1})$ along with the corresponding states and rewards resulting from an optimal strategy is provided in Tab. 4.2.

ℓ :	1	2	3	4	5	6	7	8	9	10
$a_\ell = \phi^{(2,4)}_{\mathrm{opt}}(s_{\ell-1})$:	8	8	8	5	8	8	8	5	4	8
$s_\ell = \begin{cases} m_\ell : \\ b_\ell : \\ f_\ell : \end{cases}$	2	2	1	0	2	2	1	1	0	2
	8	14	20	0	8	14	20	24	0	8
	1	2	3	0	1	2	3	4	0	1
$R(s_{\ell-1}, s_\ell; a_\ell)$:	0	0	0	24	0	0	0	0	27	0

Table 4.2. Actions a_ℓ, states (m_ℓ, b_ℓ, f_ℓ), and immediate rewards $R(s_{\ell-1}, s_\ell; a_\ell)$ observed in a configuration with an incidence of m_ℓ as displayed in Tab. 4.1 and $F = 4$.

It can be seen that all transmit antennas are in use at the beginning of a group of frames. However, the chosen number of active transmit antennas decreases as the frame index ℓ approaches the memory limit F,

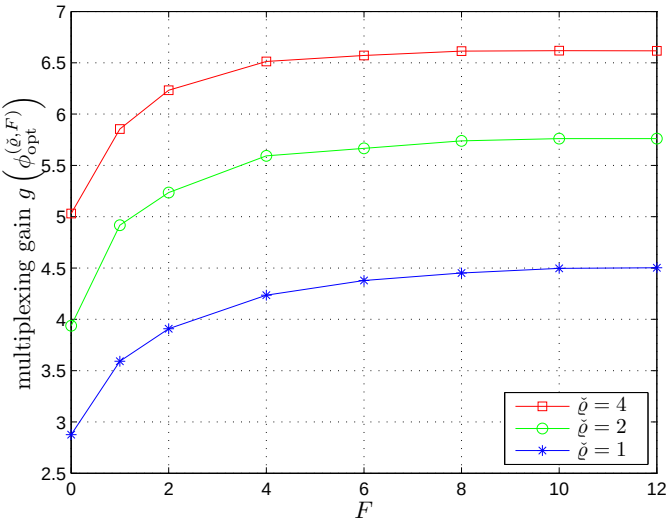

Figure 4.8 Achievable multiplexing gain over a Gaussian 8×8 MIMO channel versus the number F of frames that can be accommodated in the memory, for different noise amplifications $\check{\varrho}$.

in order to reduce the probability of losing information.

The performance of RSM when applying $\phi_{\text{opt}}^{(\check{\varrho},F)}$ is investigated by numerical simulations. The present RSM architecture is compared to a similar spatial multiplexing system relying on a conventional ARQ protocol, where the latter may be regarded as a special case of RSM with $F = 0$. In this case, the incidence of the noise power amplification

$$\frac{1}{a_\ell} \sum_{i=1}^{a_\ell} \left(\lambda_i \left(\mathbf{H}_\ell^\dagger \mathbf{H}_\ell \right) \right)^{-1} > \check{\varrho}, \tag{4.17}$$

yields an outage, i.e., the information of the frame is completely lost if the SNR after the ZF is below $\gamma_{\text{inp}}/\check{\varrho}$, due to the lack of a backup. As a consequence, the receiver requests a retransmission of the entire frame. Obviously, the probability of an outage in an ARQ-based MIMO architecture is determined by the first row of Tab. 4.1, considering a Gaussian MIMO channel as previously described. Hence, the optimal

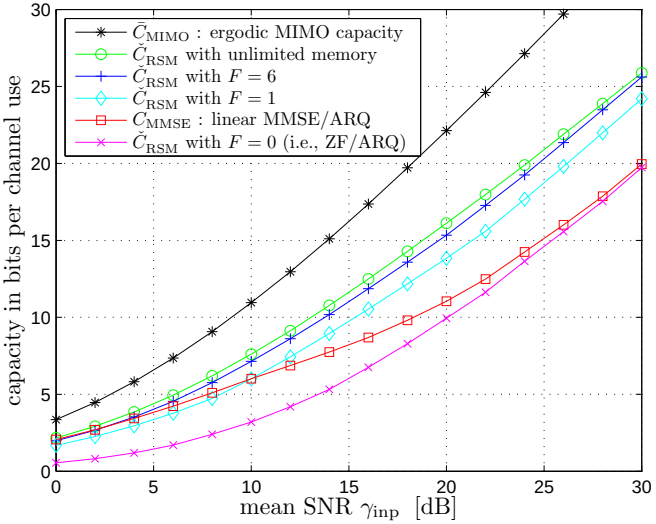

Figure 4.9 Achievable capacities with RSM, under optimal choice of noise amplification $\check{\varrho}$, over a Gaussian 4×4 MIMO channel as compared to conventional spatial multiplexing schemes with ZF/MMSE and ARQ, and theoretical unconstrained ergodic MIMO channel capacity.

number of transmit antennas for $F = 0$ leads to a multiplexing gain of

$$g\left(\phi_{\mathrm{opt}}^{(\check{\varrho},0)}\right) = \max_{a \in \{1,\ldots,M\}} ap_{m_\ell}(0; a, \check{\varrho}). \tag{4.18}$$

The multiplexing gain by the optimal policies versus F Gaussian 8×8 MIMO channel and various $\check{\varrho}$ considered above is depicted in Fig. 4.8. The results have been obtained by applying the optimal strategies in Monte-Carlo simulations. As easily seen, RSM increases the performance substantially even at a limited maximal number of recursions.

As shown for the ergodic performance examination in Sect. 3.2, the RSM constrained capacity depends on the choice of $\check{\varrho}$. Increasing $\check{\varrho}$ results in an increasing multiplexing gain, but reduces the SNR γ_{dec} at the SISO decoder input. Again using appropriate numerical methods, the

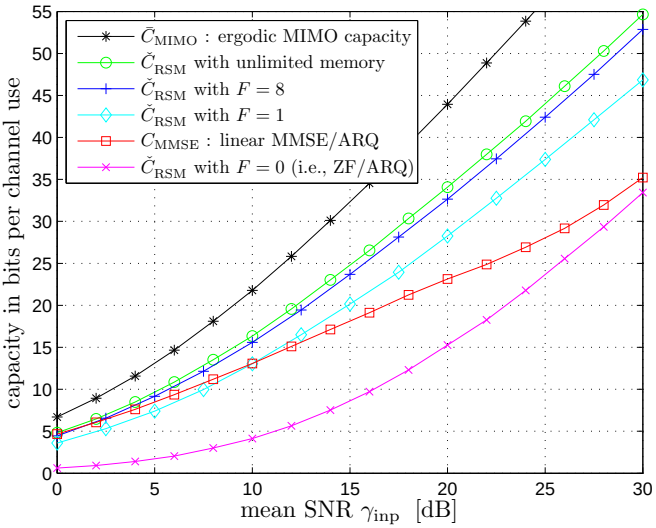

Figure 4.10 Achievable capacities with RSM, under optimal choice of noise amplification $\check{\varrho}$, over a Gaussian 8×8 MIMO channel as compared to conventional spatial multiplexing schemes with ZF/MMSE and ARQ, and theoretical unconstrained ergodic MIMO channel capacity.

achievable capacity per MIMO channel use for the proposed architecture can be calculated by

$$C_{\mathrm{RSM}}\left(\gamma_{\mathrm{inp}}\right) = \max_{\check{\varrho}>0} g\left(\phi_{\mathrm{opt}}^{(\check{\varrho},F)}\right) \log_2\left(1+\frac{\gamma_{\mathrm{inp}}}{\check{\varrho}}\right). \tag{4.19}$$

Considering the Gaussian 4×4 and 8×8 MIMO channel, Fig. 4.9 and Fig. 4.10 show the achievable capacities for different values F. Additionally, the constrained capacity of the aforementioned ARQ protocol for a ZF ($F = 0$) and an LMMSE ($C_{\mathrm{MMSE}}(\gamma_{\mathrm{inp}})$) based linear signal reconstruction are plotted. Further, the corresponding ergodic MIMO channel capacity $C_{\mathrm{MIMO}}(\gamma_{\mathrm{inp}})$ as derived in [11] and the RSM constrained ergodic capacity for unlimited memory ($F \to \infty$, ref. Sect. 3.2) are shown in the figures.

4.4. Strategy Adaptation

Finding an optimal strategy by means of the previously described dynamic programming method necessitates the accurate knowledge of the distribution of m_ℓ for all possible values a_ℓ. Usually, this assumption is unrealistic, in particular, in mobile transmitter/receiver scenarios, where the channel statistics are time-variant. Thus, the environment of the reinforcement learning problem, i.e., the distribution of m_ℓ, has to be modeled by suitable means. In the following, an *adaptive* RSM procedure building on a simple unsupervised learning of $\{p_{m_\ell}(m; a, \check{\varrho}) : m = 0, \ldots, M; a = 1, \ldots, M\}$ is introduced.

We assume that each frame contains a preamble which enables the receiver to estimate the $N \times M$ MIMO channel matrix. Further, this preamble (of the ℓth frame) facilitates the computation of $m_{\ell,1}, \ldots, m_{\ell,M}$, where $m_{\ell,a}$ denotes the necessary backup size for the case a active antennas are used for signal transmission. We point out that the preamble always uses *all* available transmit antennas regardless of the choice of a_ℓ. Let $\widehat{p}^{(\ell)}_{m_\ell}(m; a, \check{\varrho})$ represent an estimate of $p_{m_\ell}(m; a, \check{\varrho})$ computed by the receiver while processing the ℓth frame. Using a simple exponential moving average algorithm, the receiver updates the estimated probabilities according to

$$\begin{bmatrix} \widehat{p}^{(\ell)}_{m_\ell}(0; a, \check{\varrho}) \\ \vdots \\ \widehat{p}^{(\ell)}_{m_\ell}(M; a, \check{\varrho}) \end{bmatrix} = (1-\iota) \begin{bmatrix} \widehat{p}^{(\ell-1)}_{m_\ell}(0; a, \check{\varrho}) \\ \vdots \\ \widehat{p}^{(\ell-1)}_{m_\ell}(M; a, \check{\varrho}) \end{bmatrix} + \iota \begin{bmatrix} 1_{\{m_{\ell,a}=0\}} \\ \vdots \\ 1_{\{m_{\ell,a}=M\}} \end{bmatrix} \tag{4.20}$$

for $a = 1, \ldots, M$, where the parameter $\iota \in (0, 1)$ denotes the learning rate. The indicator function $1_{\{\cdot\}}$ equals 1 if the boolean expression in the subscript is true, and otherwise 0. For $\ell = 1$, the learning process starts from some properly chosen initial estimates $\{\widehat{p}^{(0)}_{m_\ell}(m; a, \check{\varrho}) : m = 0, \ldots, M; a = 1, \ldots, m\}$. The result of (4.20) is then used by the receiver to attempt computing an optimal policy, for instance by means of dynamic programming. Since the underlying distribution is only estimated, this may, of course, lead to suboptimal strategies.

The adaptability of this approach is tested in a scenario with time-variant MIMO channel statistics. We simulate the proposed adaptive RSM architecture, while we alternately change the transmission channel

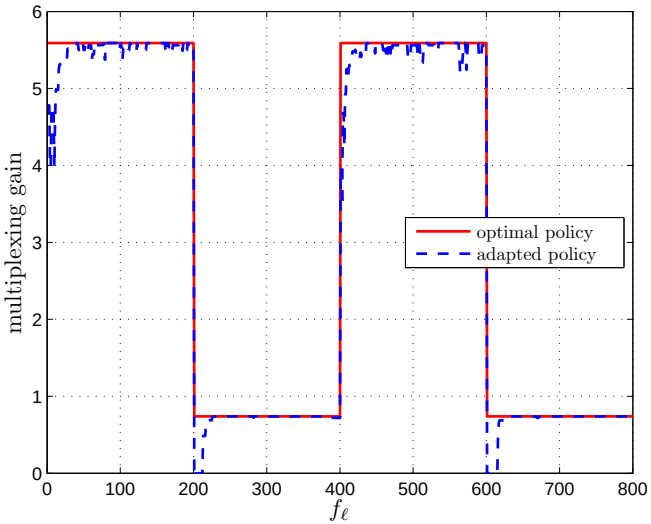

Figure 4.11 Multiplexing gain by suboptimal strategies obtained from the learning method proposed in (4.20), for a Gaussian 8×8 MIMO channel alternating between independent and maximally correlated fading at the receiver side, compared to the multiplexing gain by the respective optimal strategies.

model from the above considered Gaussian MIMO channel with *independent* fading to a second channel exhibiting fully correlated gain values at the receiver end. The latter MIMO channel is described by a random channel matrix $\mathbf{H}_\ell = \mathbf{1}_M \mathbf{h}_\ell^\dagger$, where $\mathbf{1}_M$ denotes the $M \times 1$ vector composed of 1's and $\mathbf{h}_\ell^\dagger$ a random $1 \times a_\ell$ vector containing independent zero-mean complex Gaussian distributed random variables with variance $1/a_\ell$. The channel statistics change from one model to the other every 200 frame transmissions. Fig. 4.11 shows the multiplexing gain achieved by the adapted strategies under the actual channel statistics over 800 frames. It can be seen, only about 20 frames after the switch from one to the other channel model the learning method adapts the system to the new conditions.

Chapter 5.

Verification

While the previously derived results are based on simulations considering a Rayleigh fading MIMO channel, this chapter deals with testing RSM in measured MIMO transmission channels. In the following, a novel flexible MIMO testbed is introduced, which facilitates a fast channel sounding while the impact of phase noise, usually experienced in similar systems [50], is reduced. The measured transmission scenarios are described and briefly characterized. Finally, several simulation results from the previous chapters are compared to the results when applying the according RSM architectures to the measured MIMO transmission channels.

5.1. Radio Channel Sounder

The radio channel sounding system (RaCS) provides a hardware and software solution for channel sounding in an indoor or outdoor environment. The system supports two different configurations: A SISO configuration with up to 4 transceiver stations, each consisting of a vector signal generator (VSG) and a vector signal analyzer (VSA), or a MIMO configuration, with one MIMO transmitter consisting of up to 4 VSGs, and one MIMO receiver embedding up to 4 VSAs. All stations can be synchronized while connected via cable and then moved apart before starting a measurement. The frequency response and time domain impulse response of the channels are calculated off-line by a software application.

5.1.1. Technical Description

RaCS basically consists of off-the-shelf vector signal generators *VSG2920* and vector signal analyzers *VSA2820A* from Keithley Instruments. The

VSG2920 contains an Arbitrary Waveform (ARB) generator, i.e., a fast memory for user-defined discrete baseband signal representations, which then modulate a radio frequency (RF) carrier. This enables a flexible generation of signals. The size of the occupied memory at the VSG and the adjusted sample rate define the cycle period of the signal, since the stored transmit baseband signal is periodically fed to the modulator and transmitted in an infinite loop.

The counterpart of this ARB generator is implemented in the VSA2820A. Having demodulated the received RF signal, a discrete baseband signal representation is stored in an ARB memory and can be accessed by the user after having finished the measurement. The size of the ARB memory in the analyzer and the appropriate sample rate define the maximum time duration of the captured receive baseband signal.

A special property of the VSG and the VSA is to interconnect these instruments to create a MIMO set-up. Therefore, the local oscillator (LO) of a so-called *master* instrument is shared by all *slave* instruments, e.g. the first VSG's LO is connected to the three remaining VSGs to create a 4×4 MIMO transmitter. Thus, the connected instruments operate *frequency-coherently*. Since this synchronization is restricted to the same type of instrument only, no frequency coherence between generators and analyzers can be assured. A frequency synchronization among different types of stations, namely VSGs and VSAs, is attempted by the usage of Rubidium frequency standards [77] providing a 10 MHz reference clock.

The initial phase of each VSG (or VSA) is arbitrary, but stable for a certain time period, depending on the stability of the 10 MHz Rubidium

	VSG2920	VSA2820A
frequency range	10 MHz - 6 GHz	400 MHz - 6 GHz
output power range	$-125 \ldots +13$ dBm	
measurement range		$-146 \ldots +35$ dBm
max. sample rate	100 MHz	50 MHz
max. bandwidth	80 MHz	40 MHz
max. ARB memory	100 Msamples	115 Msamples
max. ARB memory per sweep (CS)	32 ksamples	32 ksamples

Table 5.1. Technical key parameters of VSG2920 and VSA2820A.

standard. A *phase coherence* of the four carrier frequencies of the VSGs or VSAs, respectively, can be provided by a software controlled calibration routine, which has not been implemented yet. This will extend the transceiving properties to beamforming protocols.

Table 5.1 summarizes the basic key parameters of the VSG2920 and the VSA2820A.

5.1.2. MIMO Set-up

As mentioned before, the VSG2920 and VSA2820A instruments can be easily interconnected to build a MIMO transmission system. For this, the *synchronization unit 2895* (SU) from Keithley Instruments is needed. The SU is fed by the 10 MHz reference signal of the Rubidium standard and generates a 100 MHz reference signal for all connected instruments. The LO signal of the master instrument is split to up to four copies of this signal and then fed to the LO inputs of each connected instrument.

In a MIMO transmitter, the separate ARB generators have to be well synchronized. This is done by interconnecting the synchronization inputs

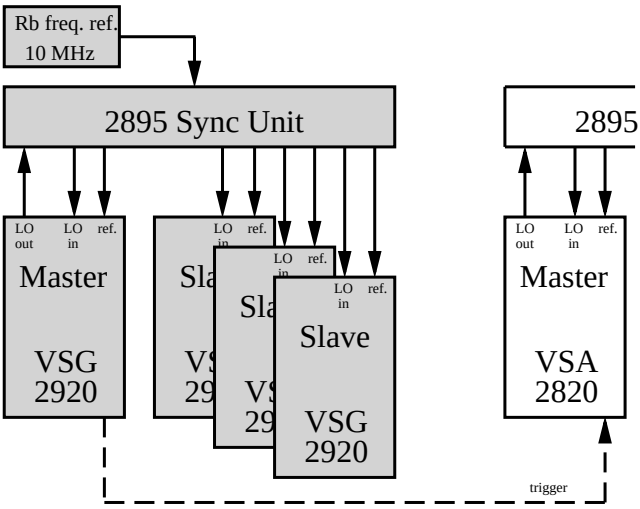

Figure 5.1 Hardware set-up for MIMO measurements.

and outputs *EvenSec-In* and *EvenSec-Out* of the instruments. The abbreviation of these connectors denotes Even Second Clock, i.e., the period of this rectangular synchronization signal is equal to two seconds. The same refers to the VSAs for synchronizing the sampling of the receive baseband signals.

The hardware set-up is shown in Fig. 5.1. MIMO signal transmissions can be performed by creating the ARB data files for the appropriate baseband signals for the signal generators. MATLAB is one possible tool to create and upload these files to the instruments. Further, MATLAB can also be used to access the VSAs after the transmission to copy the captured receive baseband signals to a PC. While the LABView based software *Sounder* (described in Sect. 5.1.3) also offers the possibility to manage these tasks, the MATLAB solution created for this thesis is the more flexible solution.

5.1.3. Channel Sounder Software Application

The user interface and the instruments' control for performing channel sounding campaigns are provided by a software application. This software is divided into a PC-based part and a firmware part of the connected instruments. These two parts interact to facilitate a real-time system control.

A measurement campaign can be summarized as follows. The ARB generator memory of the VSG contains a user-defined transmit baseband signal, which is then modulated to RF. Having started the ARB generator, the transmit baseband signal feeding the RF modulator is repeated periodically. At the analyzer side, the received RF signal is demodulated and the resulting receive baseband signal is saved in the VSA's ARB memory.

The off-the-shelf VSA features a memory access after each measurement. So, the recorded data has to be transfered to a PC before the next measurement is started. Otherwise the data is overwritten and lost. The data transfer from the VSAs to a PC is is performed via Ethernet connections, which turns out to be too slow for an appropriate real-time channel sounding. The effective data rate of approx. 10 ksamples/s would yield a maximum detectable Doppler shift of approx. 6 Hz, which is insufficient for the examination of fast time-variant channels. Further, testing adap-

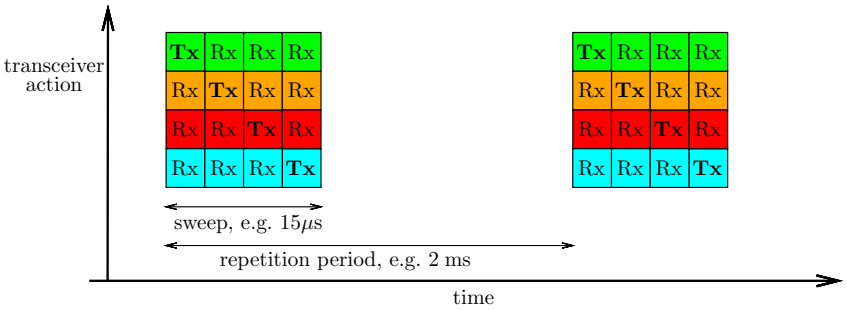

Figure 5.2 Channel sounding timing schedule.

tive transmission systems, which adapt to the present CSI in real-time for instance, requires a sufficiently fast testbed. However, high speed test equipment often lacks in flexibility. In order to verify a broad variety of even highly sophisticated algorithms and complex system architectures, RaCS follows a different approach.

In most cases the developed transmission schemes will be tested in a simulation environment (e.g. MATLAB). Therefore, the testbed is utilized as a *channel sounder*, which measures the channel properties of a certain environment or situation. The measured and analyzed CSI is then applied in the simulation tool. The previously applied channel model is then replaced by the measured CSI. The major advantage of this method is the irrelevant delay caused by the PC and its connection to the instruments. Furthermore, different transmission algorithms can be tested compared to each other for exactly the same CSI.

The channel sounding software controlled protocol is basically described in Fig. 5.2. The time period when all possible SISO transmission channels[1] are measured once is called a *sweep*. A sweep is divided in as many time intervals as VSGs exist in the setup, i.e., a time division multiple access (TDMA) scheme is used for the channel sounding. For each time interval, one VSG transmits a predefined training sequence while all VSAs are capturing this very signal. Each broadcast of the different generators is assigned to a distinct time slot to avoid interference. The

[1] The SISO transmission channels from each VSG to each VSA form the MIMO channel.

active capture duration of each VSA is printed in different colors. The sweep is repeated after a specified time called *repetition interval* to investigate the time variance of the channel. The choice of the repetition interval depends on the expected time variations of the examined channel. Note that the VSGs are also transmitting during the break between two sweeps, while the VSAs capture the received signal only during a sweep to save ARB memory space.

The channel sounder software application collects the measured data in the VSAs' memory until the measurement campaign is finished, that is, the considered number of sweeps has been performed. Then, the content of the memory is transfered to the PC at once. This change in the VSA's firmware allows to massively decrease the measurement interval. Depending on the chosen signal bandwidth and the sweep length, the measurement interval can be as small as 250 μs, i.e., a Doppler shift up to 2 kHz (considering the Nyquist theorem) can be detected.

The maximum sweep length is limited to 32 ksamples, but the total ARB memory size of the VSA is approx. 115 Msamples. For instance, considering a bandwidth of 40 MHz (this yields a 50 MHz sample rate, due to the instruments constraints), a sweep length of 15 $\mu\,s$ (750 samples) and a repetition interval of 250 μs leads to a maximum of 155,000 sweeps, or an equivalent measurement time of approx. 38 s, respectively.

The channel characterization is performed off-line using the transmit training sequences and the captured data of each VSA, when these have been downloaded after the measurement campaign.

5.2. Transmission Scenarios

Many different measurements have been taken to evaluate the usability of RSM in real environments. Although, RaCS can also provide results for outdoor measurements, we focus on indoor scenarios in this work. In the following, two examples of these transmission scenarios are introduced which can be considered typical application environments for e.g. wireless LAN, Bluetooth or Zigbee wireless transmission systems.

5.2.1. System Set-up

The measurements in both scenarios have been taken by using the 4×4 MIMO equipment described above. The center frequency $f_c = 2.45$ GHz is chosen according to an industrial scientific and medical (ISM) band used for Wireless LAN, Bluetooth or Zigbee. Transmitter and receiver front-ends are equipped by standard antennas of the type RD2458-5-SMA from Laird Technologies. The antenna spacing given by the physical measurement set-up is about $50 \ldots 60$ cm. In the case of $f_c = 2.45$ GHz, this distance corresponds a factor of $4 \ldots 4.8$ of the wavelength of the electromagnetic wave in free space.

The output signal power of the VSGs is set to -20 dBm each. Each of the four VSG outputs is connected to an amplifier of the type ZVE-8G+ amplifying the transmit signal by approx. 30 dB. This set-up shall enable the transmitter to apply a large signal power not exceeding the allowed output power for the ISM band while minimizing the non-linear impact of the external and internal amplifiers[2]. The remaining set-up parameters are summarized in Tab. 5.2.

parameter	**scenario A**	**scenario B**
measurement interval	10 ms	10 ms
bandwidth	40 MHz	40 MHz
sample rate	50 MHz	50 MHz
FFT size	128	128
active carriers	102	102
sweep length	15 μs	15 μs
VSA reference level	-40 dBm	-50 dBm
VSG transmit power	-20 dBm	-200 dBm
measurement duration	5 min.	5 min.

Table 5.2. RaCS setup parameters for the considered scenarios

[2] The directional pattern of the antennas and the cable attenuation is considered here, too. Thus, the equivalent isotropically radiated power does not exceed +20 dBm.

5.2.2. Scenario Description

In both considered transmission scenarios, the two parts of RaCS, namely the transmitter and the receiver, are located in an office environment. Walls, floor and ceiling consist of reinforced concrete or bricks, respectively. Furniture and office tools, like computers, monitors and printers, remain in the room. The blinds, made of aluminium, are closed during a measurement. Fig. 5.3 and Fig. 5.4 show a blue print of the part of the building where the measurements take place including the contents of the rooms. The RaCS transmitter and receiver are marked in orange.

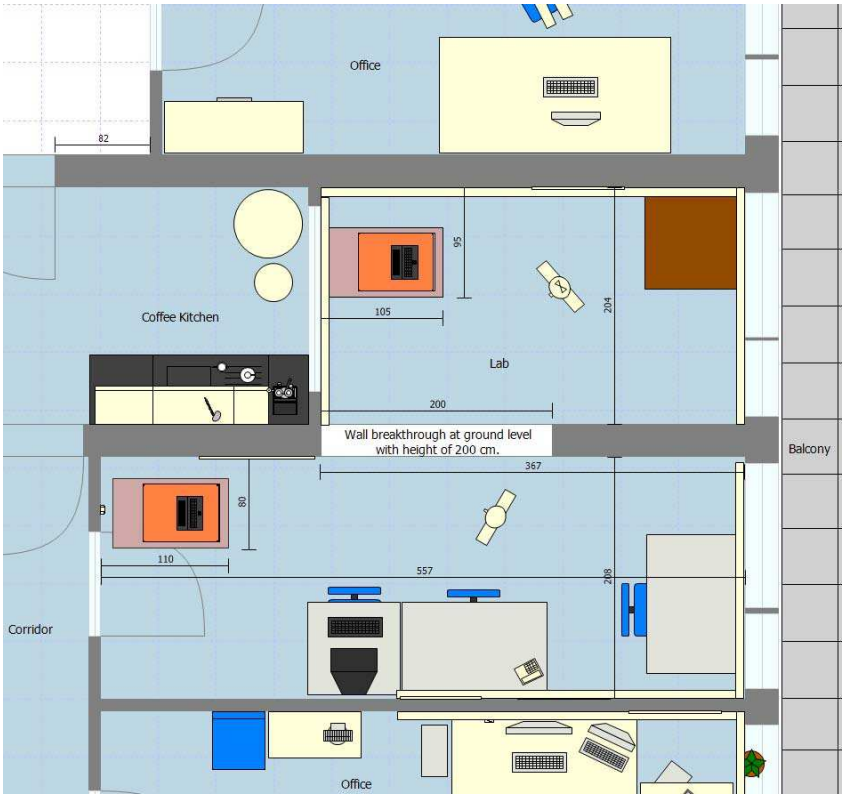

Figure 5.3 Floor plan for measurement A. MIMO transmitter and receiver are marked in orange.

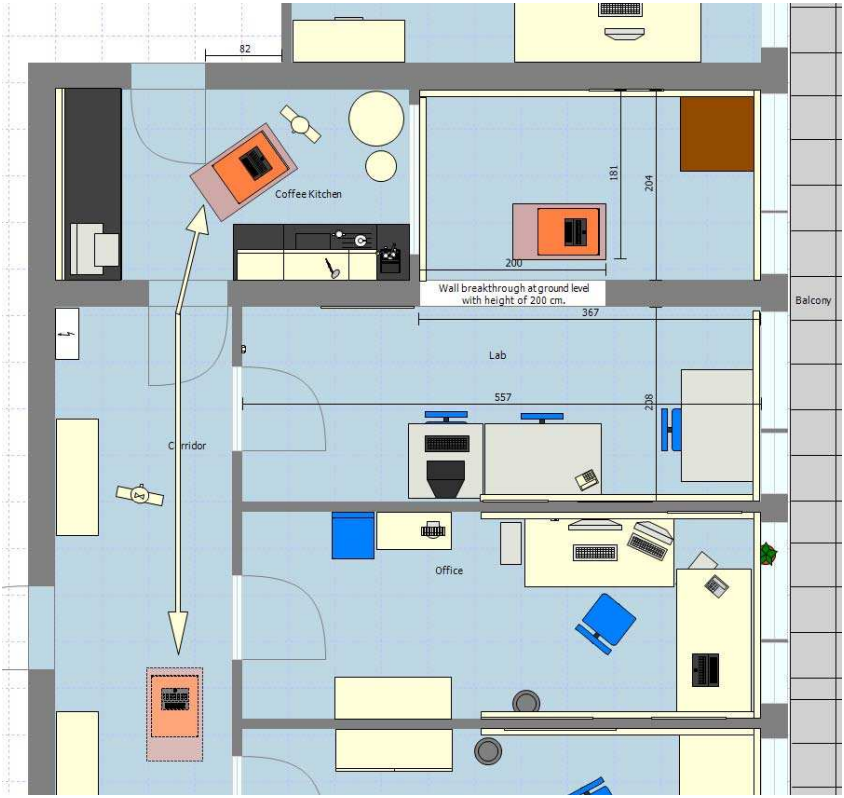

Figure 5.4 Floor plan for measurement B. MIMO transmitter and receiver are marked in orange. The transmitter location is fixed at the lab, while the receiver is permanently moving in the coffee kitchen and in the corridor, respectively.

Scenario A considers the location of RaCS transmitter and receiver as shown in Fig. 5.3, while in the same room two persons arbitrarily move around at usual pedestrian velocity. Note that this scenario represents a so-called non-line-of-sight transmission.

Scenario B is depicted in Fig. 5.4. As in scenario A, the MIMO transmitter is located at the lab, but at a different exact position. The receiver is permanently moving from the corridor to the coffee kitchen, as indicated by the arrows in the blueprint. Again two adult persons present in

the environment of the receiver can be considered as additional arbitrary moving scatterers. As can be seen from Fig. 5.4, the movement along the corridor includes both line-of-sight and non-line-of-sight transmissions.

5.2.3. Brief Scenario Characterization

Although the different scenarios have been described in the previous section, it remains difficult to compare the results of the measurement campaigns with those of other measurements. Among other reasons, the main problem is the lack of a quantifying characterization of the measurement results. Following the theory of *Wide-Sense Stationary Uncorrelated Scattering* (WSSUS) channels [78], this section tries to provide nominal parameters of the captured data for each considered scenario.

RaCS periodically captures snapshots of the channel transfer function at equidistant time intervals. This can be interpreted as the *time-variant transfer function*, one of four system functions described in [79]. Due to the high complexity of real environments, the system functions of practical channels are commonly considered as random processes. An exact statistical characterization of the channel requires the knowledge of the multidimensional joint probability density functions of all the system functions [80]. Since this knowledge is not available in a practical system, a statistical description based on the correlation functions of the various system functions is suggested in [78]. In many cases, a jointly Gaussian distribution is a sufficiently accurate model for the considered stochastic baseband processes including, for instance, the aforementioned time-variant transfer function $T(f,t)$. Furthermore, normally distributed stochastic processes are fully characterized by their first and second order moments.

The autocorrelation function of $T(f,t)$ is defined by

$$\Psi_T(f,f';t,t') = E\left[T^*(f,t)\,T(f',t')\right], \tag{5.1}$$

where f, f' denote two different frequencies and t, t' are two time instances. Supposing a WSSUS property of the channel, the correlation function of the time-variant transfer function changes in the following way. Considering the channel as a weak (or wide-sense) stationary stochastic process, the system correlation functions are invariant under a translation in time [78]. That is, it depends only on the time difference

$\Delta t = t' - t$ for all t, t'. The US property of the channel is defined by the assumption that the gain coefficients at different delays are uncorrelated. Referring to the duality of WSS and US system functions [81] and their relations by Fourier transforms, as shown in [78], [79], [82], the US property (in delay domain) corresponds to a WSS stochastic process in frequency domain, i.e., the system correlation function depends on the frequency difference $\Delta f = f' - f$ and, thus, the autocorrelation function in (5.1) results to

$$\Psi_T\left(f, f'; t, t'\right) = \Psi_T\left(f, f + \Delta f; t, t + \Delta t\right) = \Psi\left(\Delta f, \Delta t\right) \tag{5.2}$$

when modeling the channel as a WSSUS system. From the simplified time-frequency correlation function $\Psi\left(\Delta f, \Delta t\right)$, the frequency correlation function $\Psi\left(\Delta f\right)$ and the time correlation function $\Psi\left(\Delta t\right)$ can be easily derived by setting $\Delta t = 0$ or $\Delta f = 0$, respectively.

For the estimation of the RMS delay spread σ_τ, the *delay power spectral density* has to be generated first by performing the inverse Fourier transform of $\Psi\left(\Delta f\right)$ with respect to Δf according to

$$P_h\left(\tau\right) = \int_{-\infty}^{\infty} \Psi\left(\Delta f\right) \mathrm{e}^{\jmath 2\pi \Delta f \tau}\, d\Delta f, \tag{5.3}$$

where τ denotes the delay. Similarly, the definition of the Doppler spread σ_ν requires the *Doppler power spectral density* to be evaluated by the Fourier transform of $\Psi\left(\Delta t\right)$ with respect to Δt according to

$$P_H\left(\nu\right) = \int_{-\infty}^{\infty} \Psi\left(\Delta t\right) \mathrm{e}^{-\jmath 2\pi \Delta t \nu}\, d\Delta t, \tag{5.4}$$

with ν denoting the Doppler frequency. When normalizing (5.3) and (5.4) by the average power of the process, the result can be interpreted as a probability density function[3] for τ and ν. The nominal characterization of the taken measurements shall be expressed by the "width" of the normalized power spectral densities defined in (5.3) and (5.4), which can

[3] In the case of discrete values for τ and ν, normalizing (5.3) and (5.4) represent probability mass functions.

be determined by the square root of the central second order moment. Hence, the delay spread is defined as [79]

$$\sigma_\tau = \left(\frac{\int_{-\infty}^{\infty} (\tau - \theta_\tau)^2 P_h(\tau)\, d\tau}{\int_{-\infty}^{\infty} P_h(\tau)\, d\tau} \right)^{\frac{1}{2}}, \tag{5.5}$$

where

$$\theta_\tau = \frac{\int_{-\infty}^{\infty} \tau P_h(\tau)\, d\tau}{\int_{-\infty}^{\infty} P_h(\tau)\, d\tau} \tag{5.6}$$

denotes the first order moment of the normalized delay power spectral density, the so-called mean delay. Accordingly, the Doppler spread is given by [79]

$$\sigma_\nu = \left(\frac{\int_{-\infty}^{\infty} (\nu - \theta_\nu)^2 P_H(\nu)\, d\nu}{\int_{-\infty}^{\infty} P_H(\nu)\, d\nu} \right)^{\frac{1}{2}} \tag{5.7}$$

with

$$\theta_\nu = \frac{\int_{-\infty}^{\infty} \nu P_H(\nu)\, d\nu}{\int_{-\infty}^{\infty} P_H(\nu)\, d\tau}, \tag{5.8}$$

which is the first order moment of the normalized Doppler power spectral density, the so-called mean Doppler frequency. The measurement campaign for scenarios A and B are divided into appropriate groups of small time intervals to improve the estimate of the power density functions given in (5.3) and (5.4). We assume that the channel is WSS for the duration of the appropriate measurement campaign. For each of the aforementioned intervals the delay power spectral density and the Doppler power spectral density are estimated. Averaging these density functions for different groups is done using the Bartlett method [83], which helps to lower the variance of the resulting estimate.

The Doppler spread and the RMS delay spread estimates for the two

considered scenarios are given in Tab. 5.3, while the mean delay and the mean Doppler frequency are not considered below. We emphasize that the results in Tab. 5.3 shall be considered as a very rough estimate, which is due to the resolution of captured data. Especially the limited bandwidth of approximately 40 MHz is too small for an accurately resolved delay power density function in indoor environments. The biased estimation of the Doppler power density function helps reducing the variance, but leads to inaccurate results for slowly time-varying channels, due to the remaining bias.

parameter	**scenario A**	**scenario B**
Doppler spread	2.1 Hz	3.3 Hz
delay spread	28 ns	25 ns

Table 5.3. Estimated channel parameters for the considered scenarios.

5.3. Results and Comparison

Several ways exist to use the captured data of the channel sounding measurements described above. As an intuitive approach, we simulate an RSM MIMO system based on the recorded and calculated MIMO channels. That is, the measured channel attenuations are considered unchanged. The major challenge of this approach is to estimate the present SNR.

In order to enable a full comparison of the numerical performance results with the performance achievable in the measured scenarios, a method is used focusing on the time-variance of each channel and the spatial correlation of the MIMO channel. Here, the measured MIMO channel data is normalized to a fixed total average transmit power of one, i.e., the average attenuation experienced for each single channel (of the MN channels) equals $1/\sqrt{M}$. Furthermore, in view of the strong receive signals and the absence of an automatic gain control at the receiver which could increase the receiver front-end noise level in cases of a weak receive signal, thermal noise is disregarded in the estimation.

Thus, the results of the RSM constrained capacity simulations shown in previous chapters can be compared by the achievable capacity for a

given (measured) scenario. In the following, the measured channel data are used to investigate the performance of RSM in comparison to other closed-loop MIMO systems. Since RSM is considered as a narrowband transmission architecture, the wide band channel measures are transformed to narrowband channel representations.

5.3.1. RSM vs. ARQ

First, we estimate the average capacity per MIMO channel use by applying RSM to the measured channel data. Further, the unconstrained ergodic MIMO capacity is estimated based on the measured MIMO channel realizations. The capacity using a conventional ARQ mechanism is examined for an optimal MIMO coding system, an LMMSE and a ZF receiver, as explained in Sect. 3.2.1. Fig. 5.5 shows the results for scenario A.

Obviously, the large number of performed sweeps and the variety of different transmission conditions facilitate an average capacity per chan-

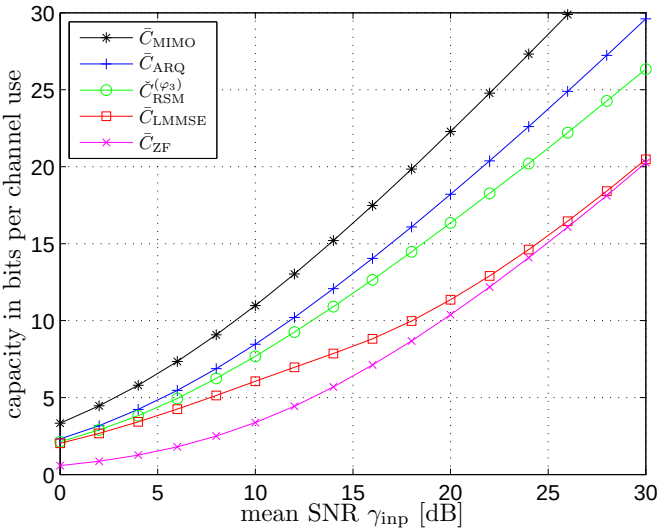

Figure 5.5 Capacity by the RSM-enhanced and standard ARQ-based scheme over a 4×4 MIMO channel measured in scenario A described in Sect. 5.2.2.

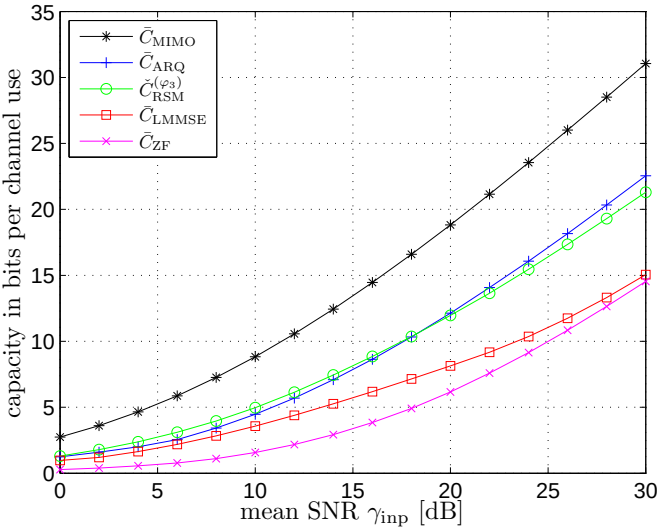

Figure 5.6 Capacity by the RSM-enhanced and standard ARQ-based scheme over a 4×4 MIMO channel measured in scenario B described in Sect. 5.2.2.

nel use similar to the simulation results denoted in Fig. 3.7. Applying the measured channel data of scenario B leads to the results depicted in Fig. 5.6. Due to the large variance of the experienced SNR in scenario B the average capacity decreases, according to Jensen's law [69].

5.3.2. RSM vs. PC

Since we investigated time-variant environments, the comparison of RSM with precoding closed-loop MIMO architectures is of great interest. Therefore, the performance of the three MIMO transmission systems RSM, UP (cf. Sect. 2.4.1) and UPRSM (cf. Sect. 3.5) are examined for the two scenarios. The transmission interval, i.e., the time between two consecutive frames, is varied by considering the measured CSI at appropriate time instances. Fig. 5.7 shows the comparison for a transmission interval of 10 ms. It can be seen that UP outperforms RSM, which is due to the accurate knowledge of CSI at the transmitter, i.e., the time-variant channel changes only slightly within 10 ms.

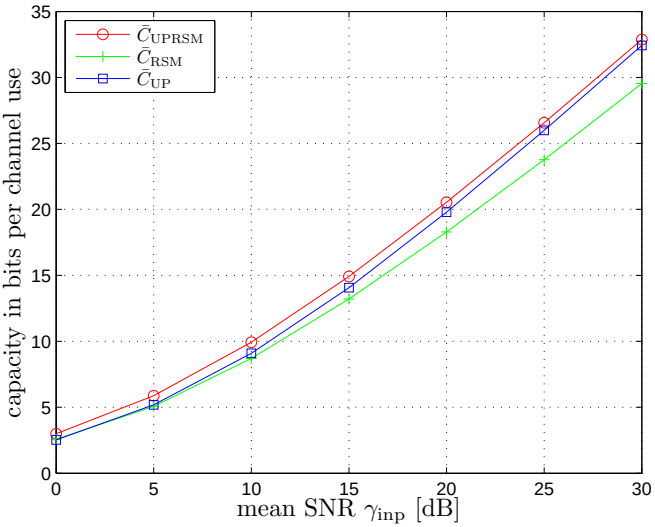

Figure 5.7 Constrained ergodic capacities with RSM, UPRSM and UP transmission schemes for the measured scenario A and a transmission interval of 10 ms.

When the transmission interval increases, e.g. due to a given TDMA schedule that does not allow shorter intervals, it is expected that the performance of UP decreases. Therefore, we consider the same measured data of scenario A, but for a transmission interval of 50 ms, i.e., only the CSI of every fifth captured MIMO channel is taken into account for the performance simulation. The result is printed in Fig. 5.8. In this case, RSM and UPRSM offer a higher capacity than UP, since UP uses outdated CSI.

Scenario B provides similar results, in particular the instantaneous capacity improvement of UPRSM versus UP. This is shown in Fig. 5.9, where the measurements of scenario B are subdivided into groups of 3 sec. time length each, i.e., a group comprises 60 transmitted frames.

Fig. 5.9 depicts the capacity in bits per channel use of each group versus the total measurement duration. It can be seen that the moving MIMO receiver results in large variations of the instantaneous capacity, as already mentioned in Sect. 5.3.1. Further, the permanent improvement facilitated by UPRSM is apparent.

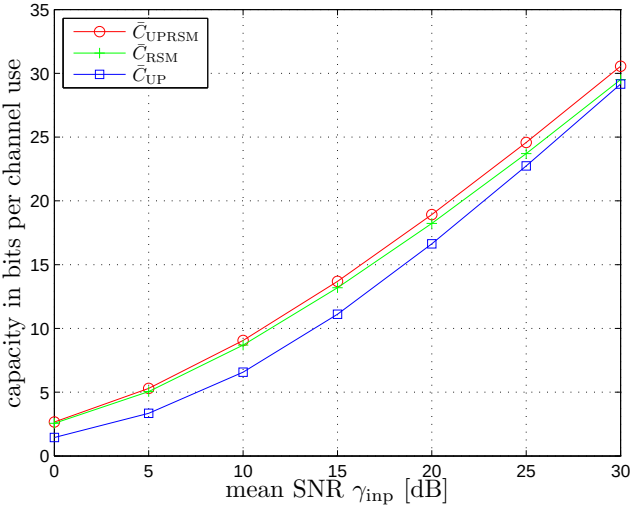

Figure 5.8 Constrained ergodic capacities with RSM, UPRSM and UP transmission schemes for the measured scenario A and a transmission interval of 50 ms.

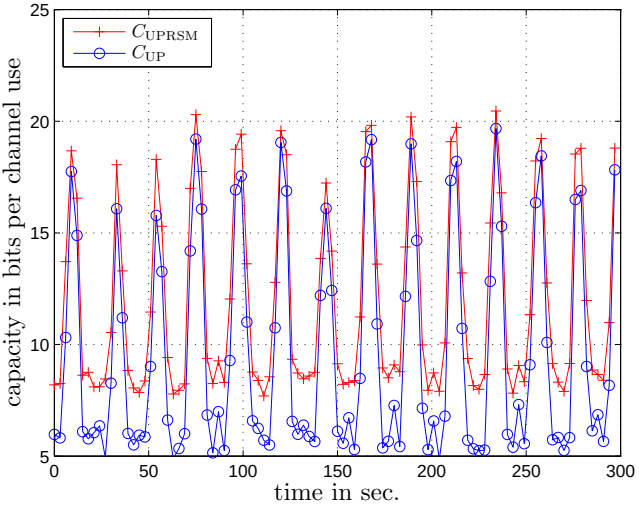

Figure 5.9 Instantaneous capacity of a UP and a UPRSM MIMO architecture in scenario B with a fixed transmitter, while the receiver is moving occasionally.

5.3.3. Limited Delay

The two measured scenarios are considered with respect to a limited number of possible retransmissions and, thus, providing an upper bound for the transmission delay or the receiver memory space, respectively. Therefore, we apply antenna selection strategies, as derived in Sect. 4.3.

Fig. 5.10 shows the results for scenario A. Similarly to the results in Sect. 4.3, even a considerably small number of frames F that can be stored in the memory enables a good performance for RSM. Fig. 5.11 depicts the performance of RSM in scenario B for a limited memory space of the receiver. It is apparent that larger values of F are required for the capacities to approach the performance of RSM with unlimited memory. As shown in Fig. 5.9, the short-term average SNR of the received signal in scenario B is substantially time-variant. For that reason, larger values of F are assumed to offer a possible averaging of CSI in good and bad conditions, which facilitates a performance improvement.

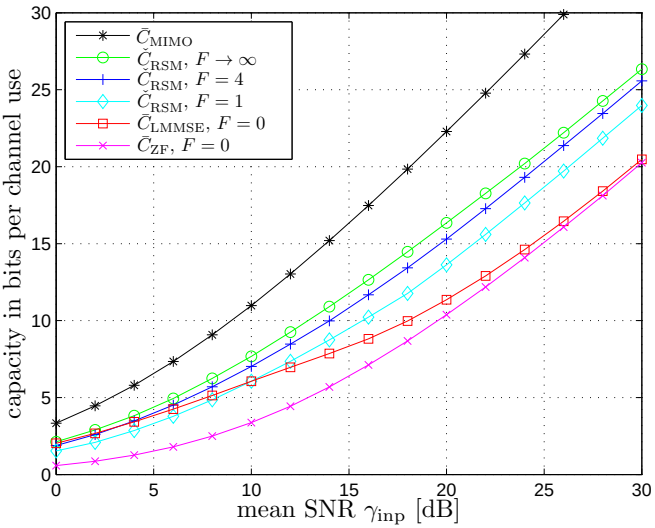

Figure 5.10 Assessed capacities in scenario A with antenna selection strategies derived from Gaussian channel model compared to the ergodic MIMO capacity.

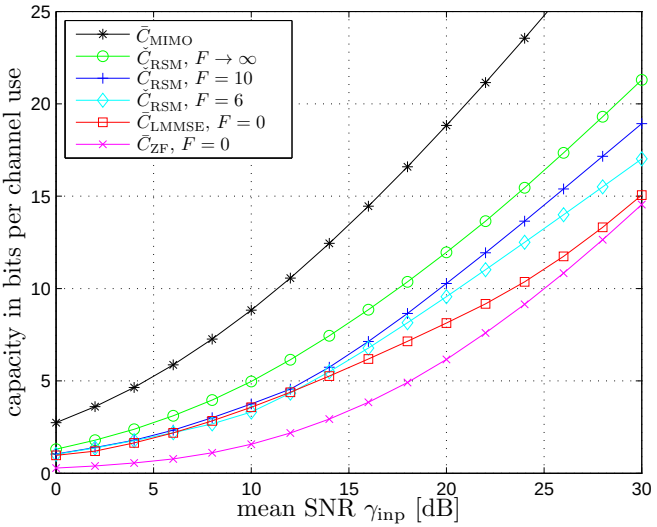

Figure 5.11 Assessed capacities in scenario B with antenna selection strategies derived from Gaussian channel model compared to the ergodic MIMO capacity.

The antenna selection strategies used in scenario A and B have been derived on the basis of transition probabilities of a random Gaussian $4 \times a_\ell$ MIMO channel (ref. Sect. 4.3). Interestingly, the performance difference is only marginal, if the antenna selection strategies are derived from the measured channel data.

Chapter 6.

Conclusions and Outlook

This chapter concludes the thesis and summarizes tasks left open for future investigation.

6.1. Conclusions

Motivated by the need for simple but efficient MIMO transmission architectures RSM has been suggested as a new approach. RSM is based on linear signal processing techniques, such as LMMSE or ZF receivers. The usage of SISO encoders and decoders at the transmitter and receiver, respectively, punctuate the simplicity of this new approach. Like other closed-loop MIMO systems an RSM receiver returns information to the transmitter. In contrast, though, to commonly used precoding MIMO architectures, the returned information focuses on the last received signal frame and not on the CSI anticipated for the coming frame. The RSM receiver requests a retransmission of signal layers that are in bad channel conditions. The RSM transmitter addresses these requested signal layers from the last transmitted frame and embeds this backup signal in the present frame waiting to be sent. Due to the combination of new data and backup signal in one frame, an economic exploitation of the MIMO resources can be considered.

Several retransmission policies have been studied in this thesis. It has turned out that a simple retransmission of the signal layer according to the weakest eigenvalue (policy φ_1) is an efficient scheme for MIMO systems with few transmit/receive antennas. However, aiming at a certain target SNR as pursued by policy φ_3 with a flexible number of retransmitted layers is convenient for MIMO architectures of greater dimension and outperforms policy φ_1. In contrast to policy φ_3, the threshold-based

policy φ_2 has low applicability in a practical scenario, but offers the same RSM constrained ergodic capacity as policy φ_3.

RSM has been compared to convenient MIMO ARQ and precoding systems in terms of the constrained throughput in Rayleigh fading MIMO channels. Optimal MIMO precoding outperforms RSM, but it necessitates accurate knowledge of CSI, while the capacity of RSM is independent of this information. It has been shown that precoding, as for instance a UP MIMO architecture, can be easily extended by RSM features with a minimal effort. Both systems take advantage of this symbiosis, since RSM only changes the transmitted signal if a signal layer is in bad condition. For practical considerations, the performance degradation caused by a limited feedback channel has been studied in this thesis. It could be shown that even feedback channels with low capacity decrease the performance of RSM only slightly.

Although the aforementioned advantages indicate the potential of RSM, still some drawbacks of the system approach remain. The trade-off of lowering the multiplexing gain for limiting the noise amplification usually caused by linear signal reconstruction is not easy to find. Optimal policy thresholds or target SNR, depending on the MIMO dimension and the SNR at the receiver input, were derived numerically.

A significant penalty of RSM is its potentially unlimited recursion depth. Since the decoding of a frame, which is in need for a retransmission, is postponed to the arrival of the backup signal (which in turn can be embedded in a frame that again requires a backup), the delay of the transmitted information data can increase extensively. Further, memory restrictions at the receiver prohibit the storage of many frames for a later decoding. A simple limitation of the maximum number of retransmitted backup signals has turned out to significantly decrease the performance of RSM.

Since the probability for the need of a retransmission can be considerably lowered by reducing the number of transmit antennas, while the quantity of receive antennas stays constant, an antenna selection strategy has proposed in this thesis. By means of machine learning optimal strategies for slowly time-variant MIMO channels have been derived. These strategies facilitate to decrease the transmission delay substantially, while the performance degradation is limited only slightly.

Different aspects of RSM have been verified in practical transmission

scenarios. Therefore, a new MIMO testbed has been introduced in this thesis. The flexible test equipment offers MIMO and wireless network channel sounding, but does not allow for real-time signal processing in highly sophisticated transmission protocols. Thus, the testbed capabilities focus on an accurate channel sounding, with limited influence of phase noise compared to switching-antenna solutions in this field. Complex system architectures are then applied by means of simulations performed off-line and the usage of the captured CSI. Several properties of RSM schemes expected from qualitative considerations of considered scenarios could be quantified for two scenarios of indoor wireless communication systems.

6.2. Outlook

Due to the property that the performance of RSM does not depend on accurate CSI, RSM is a candidate for MIMO systems operating in ISM bands and, thus, experiencing unpredictable interference from other users. An extensive consideration of applying RSM as an interference mitigating technique is necessary, even though a first examination has been introduced in [84]. Especially the verification by means of the existing testbed is a challenge, since interference has to be identified in the measurement results.

The existing transmission policies shall be investigated for their performance sensitivity when applying non-optimal thresholds or target SNRs, respectively. The search for additional retransmission policies seems to be a promising task. Different approaches from wideband signaling, like the famous water pouring, can be applied to the backup signal to even increase the efficiency of the backup. The same holds for strategies for limiting the transmission delay. For instance, lowering the target SNR and in turn the throughput after a maximum number of retransmissions can also decrease the need for a backup signal.

In the field of MDPs, an open task is how to deal with fast fading MIMO channels, since the suggested learning method for adapting the antenna selection strategy requires a sufficiently slow channel variation. Reinforcement learning offers several solution methods, which have to be analyzed and verified for their practicability in RSM systems. Compared to the introduced dynamic programming approach, reinforcement

learning methods can provide solutions with a substantial decrease of the computational effort.

RSM has been introduced as a narrowband system in this thesis. Considering each subcarrier of an OFDM system with appropriate coding as an (independent) narrowband system, RSM is also applicable in the wideband regime. A subcarrier interleaving over different frames could induce an artificial time-variance. However, practical implementations of such an approach are not available, in particular an efficient exploitation of the feedback channel in this case.

Clearly, regarding the verification of RSM and related approaches, a real-time testing of the protocol shall be considered and the simulation results for larger MIMO dimensions shall be investigated by extending the existing testbed to a MIMO system larger than 4×4.

Nomenclature

Operators

$\mathbf{\Pi}$ ..... interleaving, p. 22

Symbols

a ..... MDP action, p. 80

a_ℓ ..... number of used transmit antennas for antenna selection in frame ℓ, p. 76

b_ℓ ..... accumulated amount of transmitted signals waiting to be decoded, p. 78

$\widehat{\mathbf{b}}_{\ell,k}$ ..... kth reconstructed signal vector of the ℓth received frame fed to SISO decoder, p. 24

$\mathbf{b}_{\ell,k}$ ..... kth signal vector of the ℓth frame generated by SISO encoder, p. 21

$\mathcal{C}_m$ ..... codebook of m-dimensional subspaces of $\mathbb{C}^M$, p. 47

$\bar{C}_{\mathrm{ARQ}}$ ..... ARQ constrained ergodic MIMO channel capacity, p. 37

C_{LMMSE} ..... LMMSE constrained capacity, p. 38

$\bar{C}_{\mathrm{LMMSE}}$ ..... LMMSE constrained ergodic capacity, p. 38

C_{MIMO} ..... MIMO channel capacity, p. 36

$\bar{C}_{\mathrm{MIMO}}$ ..... ergodic MIMO channel capacity, p. 37

$\bar{C}_{\mathrm{RSM}}$ ..... RSM constrained ergodic capacity, p. 40

$\bar{C}_{\mathrm{UP}}$ ..... UP constrained ergodic capacity, p. 58

$\bar{C}_{\mathrm{UPRSM}}$ ..... UPRSM constrained ergodic capacity, p. 58

C_{ZF} ..... ZF constrained capacity, p. 37

$\bar{C}_{\mathrm{ZF}}$ ..... ZF constrained ergodic capacity, p. 37

d_{chord} ..... chordal distance, p. 47

$d_{\max,\ell}$ ..... maximal distance between $\bar{\mathbf{Q}}_\ell$ and its nearest neighbor $\widehat{\mathbf{Q}}_\ell \in \mathcal{C}_m$, p. 48

$\mathbf{D}_\ell$ ..... eigenvalue deviation vector of the $(L-\ell)$th frame caused by $\widehat{\mathbf{Q}}_{L-\ell}$ in Markov chain, p. 48

$D_{\ell,m}$ ..... mth eigenvalue deviation of the $(L-\ell)$th frame caused by $\widehat{\mathbf{Q}}_{L-\ell}$ in Markov chain, p. 48
$\mathbf{E}_\ell$ ..... error introduced by chosing $\widehat{\mathbf{Q}}_\ell$, p. 47
f ..... frequency, p. 96
f_c ..... center frequency, p. 93
f_ℓ ..... frame signal observations in memory, p. 78
F ..... number of frames that can be stored in memory, p. 71
g_ℓ ..... multiplexing gain at the ℓth frame, p. 32
$\mathbf{G}_\ell$ ..... linear signal estimator matrix of the ℓth frame, p. 14
$\mathbf{G}_\ell^{(\mathrm{LMMSE})}$ ... LMMSE based signal estimator matrix of the ℓth frame, p. 16
$\mathbf{H}_\ell$ ..... $N \times M$ MIMO channel matrix of the ℓth frame transmission, p. 11
$\underline{\mathbf{H}}_\ell$ ..... composite MIMO channel matrix including precoding, p. 18
$\widetilde{\mathbf{H}}_\ell$ ..... $(N + m_\ell) \times M$ MIMO channel matrix of the ℓth frame transmission extended by backup signal dimensions, p. 23
$\bar{\mathbf{H}}_\ell$ ..... time-correlated MIMO channel matrix of the ℓth frame transmission, p. 56
K ..... number of signal vectors per frame, p. 10
ℓ ..... frame index, p. 10
L ..... number of transmitted frames, p. 10
$\mathcal{M}$ ..... family of retransmission policies, p. 40
m_ℓ ..... number of layers requested for retransmission in frame ℓ, p. 21
$\bar{m}$ ..... average number of layers requested for retransmission, p. 62
M ..... number of transmit antennas, p. 10
$\mathbf{n}_{\ell,k}$ ..... kth noise vector of the ℓth received frame, p. 11
$\mathbf{n}'_{\ell,k}$ ..... kth noise vector of the ℓth backup signal, p. 23
$\widetilde{\mathbf{n}}_{\ell,k}$ ..... kth noise vector of the ℓth frame at the input of the linear estimator, p. 23
N ..... number of receive antennas, p. 10
$\mathcal{P}^a_{ss'}$ ..... MDP transition probability, p. 80
$p_{m_\ell}(m; a, \alpha)$. probability of $(m_\ell = m)$ given that $(a_\ell = a)$, p. 77
$\widehat{p}^{(\ell)}_{m_\ell}(m; a, \alpha)$. estimate of $p_{m_\ell}(m; a, \alpha)$, p. 85
$\mathbf{P}_\ell$ ..... unitary precoding matrix of the ℓth frame, p. 17

P_h ..... delay power spectral density, p. 97
P_H ..... Doppler power spectral density, p. 97
$\mathbf{q}_\ell$ ..... any column vector of $\mathbf{Q}_\ell$, p. 27
$\mathbf{Q}_\ell$ ..... retransmission subspace selector of the ℓth frame, p. 22
$\bar{\mathbf{Q}}_\ell$ ..... optimal retransmission subspace selector of the ℓth frame, p. 27
$\widehat{\mathbf{Q}}_\ell$ ..... retransmission subspace selector from finite codebook, p. 47
$\mathbf{r}_{\ell,k}$ ..... kth backup signal vector of the ℓth frame, p. 23
R ..... MDP reward, p. 80
$\mathcal{R}^a_{ss'}$ ..... MDP next expected reward, p. 80
s ..... MDP current state, p. 79
s' ..... MDP next state, p. 79
s_ℓ ..... MDP state of multiplexer after reception of ℓth frame, p. 78
$\mathcal{S}(M,m)$ ..... Stiefel manifold of all $(M\times m)$ matrices with orthonormal columns, p. 47
$S^{(\varphi)}$ ..... average noise amplification in Markov chain, p. 39
$\mathbf{S}_\ell$ ..... diagonal singular value matrix of $\mathbf{H}_\ell$, p. 11
$\widetilde{\mathbf{S}}_\ell$ ..... diagonal singular value matrix of $\widetilde{\mathbf{H}}_\ell$, p. 58
$\underline{\mathbf{S}}_\ell$ ..... diagonal singular value matrix of $\underline{\mathbf{H}}_\ell$, p. 19
t ..... time, p. 96
T ..... transfer function, p. 96
$\mathbf{U}_\ell$ ..... orthonormal output basis of $\mathbf{H}_\ell$, p. 12
$\underline{\mathbf{U}}_\ell$ ..... orthonormal output basis of $\underline{\mathbf{H}}_\ell$, p. 19
V ..... MDP state constrained expected reward, p. 80
V_i ..... MDP iterative state constrained expected reward, p. 81
$\mathbf{V}_\ell$ ..... orthonormal input basis of $\mathbf{H}_\ell$, p. 12
$\underline{\mathbf{V}}_\ell$ ..... orthonormal input basis of $\underline{\mathbf{H}}_\ell$, p. 19
$\mathbf{w}_{\ell,k}$ ..... kth noise vector of the ℓth frame at the output of the linear estimator, p. 24
W_ℓ ..... noise amplification of the $(L-\ell)$th frame in Markov chain, p. 39
$\widetilde{W}_\ell$ ..... noise power propagation of the $(L-\ell)$th frame caused by $\widehat{\mathbf{Q}}_{L-\ell}$ in Markov chain, p. 48
$\mathbf{x}_{\ell,k}$ ..... kth transmitted signal vector of the ℓth frame, p. 10
$\widehat{\mathbf{x}}_{\ell,k}$ ..... kth signal vector of the ℓth frame at the output of the

linear estimator, p. 14

$\widehat{\mathbf{x}}_{\ell,k}^{(\mathrm{ZF})}$ ..... kth signal vector of the ℓth frame at the output of the linear zero forcing estimator, p. 14

$\mathbf{y}_{\ell,k}$ ..... kth signal vector of the ℓth frame at the receiver input, p. 10

$\widetilde{\mathbf{y}}_{\ell,k}$ ..... kth signal vector of the ℓth frame at the input of the linear estimator, p. 23

$\mathbf{z}_\ell$ ..... vector of eigenvalues of $\mathbf{H}_{L-\ell}$ in Markov chain, p. 39

$z_{\ell,i}$ ..... ith eigenvalue of $\mathbf{H}_{L-\ell}$ in Markov chain, p. 39

α_ℓ ..... backup scaling factor for the ℓth frame, p. 23

δ ..... lag of two frames, p. 57

Δf ..... frequency lag, p. 97

Δt ..... time lag, p. 97

ϵ ..... outage probability, p. 37

$\varepsilon_{\mathrm{sig}}$ ..... variance of the transmit signal $\mathbf{x}_{\ell,k}$, p. 11

$\phi_{\mathrm{opt}}^{(\alpha,F)}$ ..... optimal antenna selection strategy, p. 81

ϕ ..... antenna selection strategy, p. 78

φ ..... retransmission policy, p. 32

γ_{dec} ..... RSM constrained postprocessing SNR at the SISO decoder input, p. 33

γ_{inp} ..... SNR at receiver input, p. 11

γ_{target} ..... target postprocessing SNR at the SISO decoder input for RSM policies φ_2 or φ_3, p. 44

$\gamma_{\ell,m}$ ..... UPRSM constrained postprocessing SNR of the mth layer in the ℓth frame, p. 57

$\gamma_{\ell,m}^{(\mathrm{UP})}$ ..... UP constrained postprocessing SNR of the mth layer in the ℓth frame, p. 57

$\gamma_{\ell,m}^{(\mathrm{ZF})}$ ..... ZF constrained postprocessing SNR of the mth layer in the ℓth frame, p. 15

η ..... MDP discount rate, p. 80

ι ..... learning rate, p. 85

$\kappa_{\ell,m}$ ..... LMMSE constrained postprocessing SINR of the mth layer in the ℓth frame, p. 16

λ_m^2 ..... eigenvalue of the mth layer, p. 12

Λ_ℓ ..... column vector constituted by the eigenvalues of $\mathbf{H}_\ell^\dagger \mathbf{H}_\ell$, p. 32

μ ..... MSE of $\mathbf{w}_{\ell,k}$, p. 27

ν ..... Doppler frequency, p. 97
θ_ν ..... mean Doppler frequency, p. 98
θ_τ ..... mean delay, p. 98
ϑ_{noise} ..... variance of the received noise vector $\mathbf{n}_{\ell,k}$, p. 11
$\vartheta^{(\ell)}_{\text{backup}}$ ..... variance of the backup noise vector $\mathbf{n}_{\ell,k}$, p. 23
$\mathbf{\Theta}^{(\text{ZF})}_{\text{noise}}$ ..... variance of the noise vector $\mathbf{w}_{\ell,k}$ when using ZF, p. 15
ϱ_ℓ ..... noise amplification caused by linear estimation in the ℓth frame, p. 32
$\check{\varrho}$ ..... noise amplification limit as applied for policy φ_3, p. 32
ρ ..... correlation coefficient of the elements of adjacent frames, p. 57
σ_ν ..... RMS Doppler spread, p. 97
σ_τ ..... RMS delay spread, p. 97
τ ..... delay, p. 97
$\mathbf{\Omega}_\ell$ ..... backup scaling matrix for the ℓth frame, p. 54
ξ ..... threshold used in policy φ_1 and φ_2, p. 33
Ψ ..... Ψ_T in a WSSUS system, p. 97
Ψ_T ..... time-frequency autocorrelation of the time-variant transfer function, p. 96
$\Psi^{(\varphi)}$ ..... probability measure in Markov chain, p. 39
$\zeta_{\ell,m}$ ..... mth eigenvalue deviation introduced by chosing $\widehat{\mathbf{Q}}_\ell$, p. 48

Abbreviations

ARQ ..... automatic repeat request, p. 4
BLAST ..... Bell Lab Layered Space-Time, p. 2
BPSK ..... binary phase-shift keying, p. 12
CSI ..... channel state information, p. 1
D-BLAST .... Diagonal Bell Lab Layered Space-Time, p. 3
DSP ..... digital signal processor, p. 5
FEC ..... forward error correction, p. 19
FPGA ..... field programmable gate array, p. 5
HARQ ..... Hybrid-ARQ, p. 4
ISI ..... inter symbol interference, p. 14
ISM ..... industrial scientific and medical, p. 93
LIFO ..... last-in first-out, p. 24
LMMSE ..... linear minimum mean-square error, p. 2
LO ..... local oscillator, p. 88

LTE ..... Long Term Evolution, p. 5
MDP ..... Markov decision process, p. 4
MIMO ..... multiple-input multiple-output, p. 1
ML ..... maximum likelihood, p. 3
MSE ..... mean squared error, p. 16
MSI ..... multistream interference, p. 14
OFDM ..... Orthogonal Frequency Division Multiplex, p. 5
QAM ..... quadrature amplitude modulation, p. 12
RaCS ..... Radio applications and Channel Sounding, p. 87
RF ..... radio frequency, p. 88
RSM ..... recursive spatial multiplexing, p. 6
SIC ..... successive interference cancellation, p. 2
SINR ..... signal to interference and noise ratio, p. 16
SISO ..... single input single output, p. 1
SNR ..... signal-to-noise ratio, p. 11
T-BLAST .... Turbo Bell Lab Layered Space-Time, p. 3
UP ..... unitary precoding, p. 17
UPRSM ..... UP supported by RSM, p. 56
US ..... Uncorrelated Scattering, p. 97
V-BLAST .... Vertical Bell Lab Layered Space-Time, p. 2
VE ..... vertical encoding, p. 2
VoD ..... Video on Demand, p. 1
VoIP ..... Voice over Internet Protocol, p. 1
WSS ..... Wide-Sense Stationary, p. 97
WSSUS ..... Wide-Sense Stationary Uncorrelated Scattering, p. 96
ZF ..... zero forcing, p. 2

Appendix A.

Appendix

A.1. Banach Fixed Point Theorem

Solving (3.61) and, hence, deriving a unique function $W(\xi)$ offers a method to estimate $\xi_{\mathrm{opt}}(\gamma_{\mathrm{inp}})$, which in turn maximizes the RSM constrained ergodic capacity for the case that $M, N \to \infty$ with $M = N$. There are several numeric methods to cope with this problem, the Banach fixed point theorem [75] is applied here, since (3.61) equals the structure of a contraction mapping [85].

First, the considered interval is limited to $0 < W < \infty$ supposing it represents a complete metric space (B, d_{B}) and $f_{\mathrm{B}} : B \to B$ is the contraction map corresponding to the right hand side of (3.61). To fulfil the theorem, for each $0 < \xi < 4$ there is a nonnegative real number $\kappa < 1$ such that the Lipschitz condition

$$d_{\mathrm{B}}\left(f_{\mathrm{B}}(b), f_{\mathrm{B}}(b')\right) \leq \kappa \cdot d_{\mathrm{B}}\left(b, b'\right) \tag{A.1}$$

holds for all $b, b' \in B$. Then the function f_{B} allows only one fixed point $\widehat{b} \in B$, i.e., $f_{\mathrm{B}}(\widehat{b}) = \widehat{b}$.

In the present case with $b = W \in \mathbb{R}^+$, a suitable metric is given by the derivative of f_{B}

$$\frac{df_{\mathrm{B}}}{dW} = \frac{2}{\pi}\left(\frac{1}{4}\sqrt{\frac{\xi(4-\xi)}{(1+\xi W)^2}} + \frac{1}{\sqrt{1+4W}}\arctan\left(\sqrt{\frac{\xi(1+4W)}{4-\xi}}\right)\right). \tag{A.2}$$

It can be easily seen that (A.2) decreases for increasing W. If $W \to 0$ (while this is already outside the considered interval), the maximum value of $df_{\mathrm{B}}/dW \to 1$ if $\xi = 4$, as can be seen in Fig. A.1 for the interval

$0 < W < 4$.

Thus, the condition in (A.1) is satisfied and, therefore, a unique fixed point can be found in the given interval by interpreting (3.61) as an iterative sequence

$$b_n = f_{\mathrm{B}}(b_{n-1}) \quad \text{for } n = 1, 2, 3, \ldots, \tag{A.3}$$

where this sequence converges to $\widehat{b}$ for any start value $b_0 \in B$. The result of this numerical iteration can be demonstrated by a graphical interpretation of (3.61) shown in Fig. A.2. While f_{B} is depicted by a "bended" plane depending on W and ξ, the left hand side of (3.61) is represented by a "straight plane", which is independent of ξ and its slope in W direction equals 1. The computational result $\widehat{b} = W(\xi)$ is marked by the crossing line of the two previously described planes.

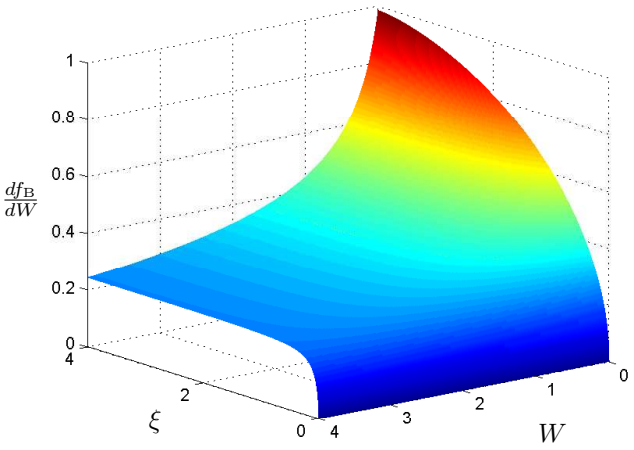

Figure A.1 $\frac{df_{\mathrm{B}}}{dW}$ for $0 < \xi < 4$ and $0 < W < 4$.

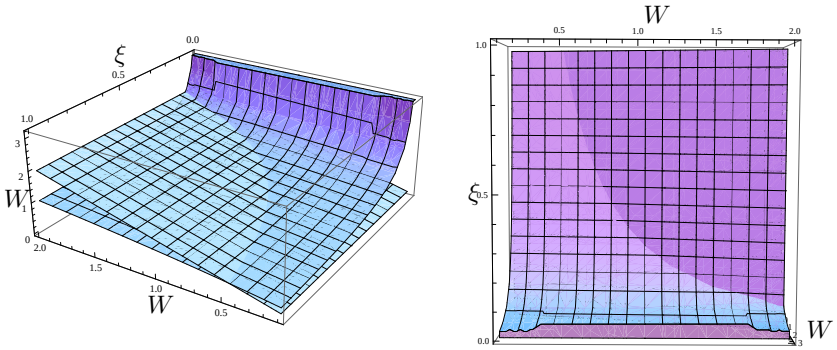

Figure A.2 Graphical interpretation of (3.61). Both figures depict the same functions in different views. The intersection of the two planes represents the values of W solving (3.61) for different ξ.

Bibliography

[1] M. Vu and A. Paulraj, "MIMO wireless linear precoding," *IEEE Sig. Processing Mag.*, vol. 24, no. 5, pp. 86–105, Sep. 2007.

[2] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[3] G. Lebrun, J. Gao, and M. Faulkner, "MIMO transmission over a time-varying channel using SVD," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 757–764, March 2005.

[4] A. Khrwat, B. Sharif, C. Tsimenidis, S. Boussakta, and S. Le Goff, "Feedback delay in precoded spatial multiplexing MIMO systems," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2008*, Sept. 2008, pp. 1–4.

[5] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[6] G. Jongren and M. Skoglund, "Quantized feedback information in orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2473–2486, Oct. 2004.

[7] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. New York, NY, USA: Cambridge University Press, 2003.

[8] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *URSI International Symposium on Signals, Systems and Electronics, ISSSE 98*, Sep. 1998, pp. 295–300.

[9] J. Winters, J. Salz, and R. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1740–1751, Feb. 1994.

[10] A. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception (DTDR)," United States Patent 5345599, 1994.

[11] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. /Dec. 1999.

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

[13] G. J. Foschini, "Layered space–time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[14] M. Sellathurai and S. Haykin, "TURBO-BLAST for high-speed wireless communications," in *IEEE Wireless Communications and Networking Conference, WCNC 2000*, vol. 1, 2000, pp. 315–320.

[15] M. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[16] B. Hassibi and H. Vikalo, "On the expected complexity of sphere decoding," in *Conference Record of the Thirty-Fifth Asilomar Conference on Signals, Systems and Computers*, vol. 2, 2001, pp. 1051–1055.

[17] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

[18] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics of Computation*, vol. 44, no. 170, pp. 463–471, 1985.

[19] S. Yoon and S. K. Lee, "A detection algorithm for multi-input multi-output (MIMO) transmission using poly-diagonalization and trellis decoding," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 993–1002, Aug. 2008.

[20] W. Santipach and M. Honig, "Asymptotic performance of MIMO wireless channels with limited feedback," in *IEEE Military Communications Conference, MILCOM 2003*, vol. 1, Oct. 2003, pp. 141–146.

[21] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, no. 2, pp. 139–159, 1996.

[22] L. Zheng and D. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[23] A. Barg and D. Nogin, "Bounds on packings of spheres in the grassmann manifold," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2450–2454, Sep. 2002.

[24] J. Choi, B. Mondal, and R. Heath, "Interpolation based unitary precoding for spatial multiplexing MIMO-OFDM with limited feedback," *IEEE Trans. Signal Processing*, vol. 54, no. 12, pp. 4730–4740, Dec. 2006.

[25] M. Vu and A. Paulraj, "Optimal linear precoders for MIMO wireless correlated channels with nonzero mean in space-time coded systems," *IEEE Trans. Signal Processing*, vol. 54, no. 6, pp. 2318–2332, Jun. 2006.

[26] S. Lin, D. Costello, and M. Miller, "Automatic-repeat-request error-control schemes," *IEEE Commun. Mag.*, vol. 22, no. 12, pp. 5–17, Dec. 1984.

[27] S. Lin and P. S. Yu, "A hybrid ARQ scheme with parity retransmission for error control of satellite channels," *IEEE Trans. Commun.*, vol. 30, no. 7, pp. 1701–1719, July 1982.

[28] D. Mandelbaum, "An adaptive-feedback coding scheme using incremental redundancy (corresp.)," *IEEE Trans. Inform. Theory*, vol. 20, no. 3, pp. 388–389, May 1974.

[29] D. Chase, "Code combining – a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," *IEEE Trans. Commun.*, vol. 33, pp. 385–393, May 1985.

[30] E. Onggosanusi, A. Dabak, Y. Hui, and G. Jeong, "Hybrid ARQ transmission and combining for MIMO systems," in *IEEE International Conference on Communications, ICC '03*, vol. 5, May 2003, pp. 3205–3209.

[31] A. Dekorsy, "A cutoff rate based cross-layer metric for MIMO-HARQ transmission," in *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2005*, vol. 4, Sept. 2005, pp. 2166–2170.

[32] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. The MIT Press, Mar. 1998.

[33] R. E. Bellman, "A Markovian decision process," *Indiana Univ. Math. Journal*, vol. 6, pp. 679–684, 1957.

[34] D. Djonin and V. Krishnamurthy, "V-BLAST power and rate control under delay constraints in Markovian fading channels - optimality of monotonic policies," in *IEEE International Symposium on Information Theory*, Jul. 2006, pp. 2099–2103.

[35] ——, "MIMO transmission control in fading channels – a constrained Markov decision process formulation with monotone randomized policies," *IEEE Trans. Signal Processing*, vol. 55, no. 10, pp. 5069–5083, Oct. 2007.

[36] M. Luccini, A. Shami, and S. Primak, "Cross-layer optimisation of network performance over multiple-input multiple-output wireless mobile channels," *Communications, IET*, vol. 4, no. 6, pp. 683–696, Apr. 2010.

[37] X. Liao, X. Sun, L. Kong, and J. Jiang, "Optimal cross-layer performance over MIMO fading channel with adaptive transmission," in *5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCom '09*, Sep. 2009, pp. 1–4.

[38] V. Lau and Y. Chen, "Delay-optimal power and precoder adaptation for multi-stream MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3104–3111, Jun. 2009.

[39] S. Yun and C. Caramanis, "Reinforcement learning for link adaptation in MIMO-OFDM wireless systems," in *IEEE Global Telecommunications Conference, GLOBECOM 2010*, Dec. 2010, pp. 1–5.

[40] E. Aschbacher, S. Caban, C. Mehlfuhrer, G. Maier, and M. Rupp, "Design of a flexible and scalable 4x4 MIMO testbed," in *11th IEEE Digital Signal Processing Workshop and the 3rd IEEE Signal Processing Education Workshop*, Aug. 2004, pp. 178–181.

[41] F. Azami, A. Ghorssi, H. Hemesi, A. Mohammadi, and A. Abdipour, "Design and implementation of a flexible 4x4 MIMO testbed," in *International Symposium on Telecommunications, IST 2008*, Aug. 2008, pp. 268–272.

[42] D. Bates, S. Henriksen, B. Ninness, and S. Weller, "A 4x4 fpga-based wireless testbed for lte applications," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2008*, Sep. 2008, pp. 1–5.

[43] D. Perels, S. Haene, P. Luethi, A. Burg, N. Felber, W. Fichtner, and H. Bölcskei, "ASIC implementation of a MIMO-OFDM transceiver for 192 Mbps WLANs," in *Proceedings of the 31st European Solid-State Circuits Conference, ESSCIRC 2005*, Sep. 2005, pp. 215–218.

[44] M. Wenk, P. Luethi, T. Koch, P. Maechler, N. Felber, W. Fichtner, and M. Lerjen, "Hardware platform and implementation of a real-time multi-user MIMO-OFDM testbed," in *IEEE International Symposium on Circuits and Systems, ISCAS 2009*, May 2009, pp. 789–792.

[45] K. Nishimori, R. Kudo, Y. Takatoti, A. Ohta, and K. Tsunekawa, "Performance evaluation of 8x8 multi-user MIMO-OFDM testbed in an actual indoor environment," in *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2006, pp. 1–5.

[46] J.-M. Molina-Garcia-Pardo, J.-V. Rodriguez, and L. Juan-Llacer, "MIMO channel sounder based on two network analyzers," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 9, pp. 2052–2058, Sept. 2008.

[47] M. Steinbauer, D. Hampicke, G. Sommerkorn, A. Schneider, A. Molisch, R. Thoma, and E. Bonek, "Array measurement of the double-directional mobile radio channel," in *IEEE 51st Vehicular Technology Conference Proceedings, VTC 2000-Spring Tokyo*, vol. 3, 2000, pp. 1656–1662.

[48] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, and A. Paulraj, "Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2253–2265, Jun. 2010.

[49] K. Mammasis, R. Stewart, and E. Pfann, "3-dimensional channel modeling using spherical statistics for multiple input multiple output systems," in *IEEE Wireless Communications and Networking Conference, WCNC 2008*, Apr. 2008, pp. 769–774.

[50] D. Baum and H. Bölcskei, "Impact of phase noise on MIMO channel measurement accuracy," in *IEEE 60th Vehicular Technology Conference, VTC2004-Fall*, vol. 3, Sep. 2004, pp. 1614–1618.

[51] J. G. Proakis and M. Salehi, *Digital communications*, 5th ed. Boston: McGraw-Hill, 2008.

[52] E. H. Moore, "On the reciprocal of the general algebraic matrix," *Bulletin of the American Mathematical Society*, vol. 26, pp. 394–395, 1920.

[53] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*, 4th ed., ser. McGraw-Hill in electrical and computer engineering. Boston: McGraw-Hill, 2002.

[54] J. Cioffi, "EE379A - Digital Communication: Signal Processing," 2008. [Online]. Available: http://stanford.edu/group/cioffi/ee379a/

[55] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *SIAM Journal of Applied Mathematics*, vol. 8, pp. 300–304, 1960.

[56] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution : HSPA and LTE for Mobile Broadband*, 2nd ed. Oxford: Elsevier Ltd., 2008.

[57] A. Maaref and S. Aissa, "A cross-layer design for MIMO rayleigh fading channels," in *Canadian Conference on Electrical and Computer Engineering*, vol. 4, May 2004, pp. 2247–2250.

[58] J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix," *Annals of Mathematical Statistics*, vol. 20, p. 621, 1949.

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

[60] H. Lee and I. Lee, "Channel capacity of BLAST based on the zero-forcing criterion," in *IEEE 63rd Vehicular Technology Conference, VTC 2006-Spring*, vol. 4, May 2006, pp. 1615–1619.

[61] T. Hunziker, T. Edlich, and D. Dahlhaus, "On the ergodic performance of a simple closed-loop spatial multiplexing architecture," in *International Symposium on Information Theory and Its Applications, ISITA 2008*, Dec. 2008, pp. 1–6.

[62] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inform. Theory*, vol. 42, no. 5, pp. 1520–1529, Sep. 1996.

[63] N. J. A. Sloane, "Packings in Grassmannian Spaces." [Online]. Available: http://www2.research.att.com/ njas/grass/

[64] A. J. Hoffman and H. Wielandt, "The variation of the spectrum of a normal matrix," *Duke Math. J.*, vol. 20, no. 1, pp. 37–39, 1953.

[65] D. Agrawal, T. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2618–2626, Sep. 2001.

[66] I. Dhillon, R. Heath, T. Strohmer, and J. Tropp, "Constructing packings in Grassmannian manifolds via alternating projection," *Experiment. Math.*, vol. 17, pp. 9–35, 2008.

[67] B. Hochwald, T. Marzetta, T. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1962–1973, Sep. 2000.

[68] D. J. Love, "Tables of complex Grassmannian packings." 2004. [Online]. Available: http://www.ece.purdue.edu/∼djlove/grass.html

[69] J. Jensen, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Mathematica*, vol. 30, pp. 175–193, 1906.

[70] C. Gutierrez Diaz de Leon, M. Bean, and J. Garcia, "Generation of Correlated Rayleigh-Fading Envelopes for Simulating the Variant Behavior of Indoor Radio Propagation Channels," in *IEEE 60th Vehicular Technology Conference*, vol. 6, Sept. 2004, pp. 4245–4249.

[71] A. Edelman and N. Rao, "Random matrix theory," *Acta Numerica*, vol. 14, pp. 233—297, 2005.

[72] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*. PhD. thesis, Massachusetts Inst. of Technol., 1989.

[73] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Commun. Inf. Theory*, vol. 1, pp. 1–182, June 2004.

[74] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed. New York: Dover, 1964.

[75] S. Banach, "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales," *Fund. Math.*, vol. 3, pp. 133–181, 1922.

[76] R. E. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.

[77] Stanford Research Systems, "PRS10 - Rubidium frequency standard with low phase noise," 2011. [Online]. Available: http://www.thinksrs.com/downloads/PDFs/Catalog/PRS10c.pdf

[78] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Transactions on Communications Systems*, vol. 11, no. 4, pp. 360–393, Dec. 1963.

[79] B. Fleury, *Charakterisierung von Mobil- und Richtfunkkanälen mit schwach stationären Fluktuationen und unkorrelierter Streuung (WSSUS).* Dissertation ETH Zürich Nr. 9030, 1990.

[80] J. D. Parsons, *The Mobile Radio Propagation Channel*, 2nd ed. New York: John Wiley & Sons, 2000.

[81] P. A. Bello, "Time-frequency duality," *IEEE Trans. Inform. Theory*, vol. 10, no. 1, pp. 18–33, Jan. 1964.

[82] R. Kattenbach, *Charakterisierung zeitvarianter Indoor-Funkkanäle anhand ihrer System- und Korrelationsfunktionen.* Shaker Verlag, 1997.

[83] M. S. Bartlett, "Smoothing periodograms from time series with continuous spectra," *Nature*, vol. 161, pp. 686–687, May 1948.

[84] I. Shah, T. Hunziker, T. Edlich, and D. Dahlhaus, "Recursive spatial multiplexing in the presence of time-variant co-channel interference," in *17th IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, Nov. 2010, pp. 1–5.

[85] V. I. Istratescu, *Fixed Point Theory, An Introduction.* Dordrecht, Holland: D. Reidel Publishing Company, 1981.

In wireless transmissions great capacity gains can be realized by using multiple-input multiple-output (MIMO) systems. We introduce a closed-loop MIMO architecture with linear array signal processing. The novel architecture uses the feedback channel to request a retransmission of signal parts in critical subspaces. The requested backup subsequently helps the receiver to evade excessive noise amplification in the linear signal reconstruction. Since the backup signal is embedded in the successive transmission frame, which may require yet another backup, the procedure results in a recursive spatial multiplexing (RSM) scheme. We compare the performance of RSM with other closed-loop MIMO systems by means of the system constrained ergodic capacities. The recursive nature of RSM makes it impossible to guarantee a maximum delay. Exploiting receive diversity in MIMO systems, a reduction of the transmit signal dimensions for a constant number of receive antennas leads to less frequent retransmission requests. For that reason, we introduce optimal transmit antenna selection strategies, which are found by means of machine learning methods.

ISBN 978-3-86219-608-1