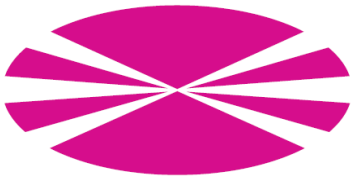
Department of Electronics and Systems
University of A Coruña, Spain



PhD Thesis

Design of Limited Feedback for Robust MMSE Precoding in Multiuser MISO Systems

Paula María Castro Castro

October 2009

PhD Advisors:
Luis Castedo Ribas
Michael Joham

Design of Limited Feedback for Robust MMSE Precoding in Multiuser MISO Systems

Paula María Castro Castro

Department of Electronics and Systems
University of A Coruña, Spain

D. Luis Castedo Ribas y D. Michael Joham

CERTIFICAN:

Que la memoria titulada “Design of Limited Feedback for Robust MMSE Precoding in Multiuser MISO systems”, ha sido realizada por Dña. Paula María Castro Castro bajo nuestra dirección en el Departamento de Electrónica y Sistemas de la Universidad de A Coruña y concluye la Tesis que presenta para optar al grado de Doctor con la mención de Doctor Europeo.

A Coruña, 29 de junio de 2009

Fdo: Dr. Luis Castedo Ribas
Director de la Tesis Doctoral
Catedrático de Universidad
Dpto. de Electrónica y Sistemas
Universidad de A Coruña

Fdo: Dr. Michael Joham
Director de la Tesis Doctoral
Senior Researcher
Associate Institute for Signal Processing
Technische Universität München

Tesis Doctoral:	Design of Limited Feedback for Robust MMSE Precoding in Multiuser MISO Systems
Autor:	Dña. Paula María Castro Castro
Directores:	Dr. Luis Castedo Ribas y Dr. Michael Joham
Fecha:	9 de octubre de 2009

Tribunal

Presidente:

Vocal 1:

Vocal 2:

Vocal 3:

Secretario:

A Juan

Agradecimientos

En primer lugar quiero expresar mi más profundo y sincero agradecimiento a mi tutor y director de tesis, el Dr. Luis Castedo Ribas, por su apoyo y ayuda durante estos años. Gracias por haberme motivado a realizar la tesis doctoral y por la inestimable ayuda en todos y cada uno de los artículos que hemos realizado juntos. Gracias por todo lo que me has enseñado y por haber compartido conmigo tantas horas de trabajo durante todos estos años. Y, sobre todo, gracias por haber sido más que un jefe, un amigo. Gracias.

Quiero darle especialmente las gracias al Dr. Wolfgang Utschick de la Universidad Técnica de Múnich, por haberme permitido trabajar con él y con su gran equipo de investigadores, y, sobre todo, por haber apoyado esta colaboración durante todos estos años.

También me gustaría agradecer al Dr. Michael Joham de la Universidad Técnica de Múnich y codirector también de esta tesis, su enorme aportación a la misma, con sus ideas y su inestimable ayuda y comprensión desde el año 2006, año en el que realicé mi primera estancia en dicha universidad. Le agradezo el haberme enseñado tantas cosas de Álgebra matemática, la extraordinaria paciencia que ha tenido conmigo y las maravillosas estancias en Múnich de las que he podido disfrutar los últimos años.

A los miembros del tribunal de evaluación de mi tesis, los profesores doctores Ana I. Pérez–Neira, Wolfgang Utschick, John Thompson, Roberto López Valcarce y Miguel González López, por el trabajo realizado y su excelente disposición en todo momento.

A mi compañero de doctorado y laboratorio, Fran, por haber compartido tan buenos ratos juntos. Y a Jose, siempre un buen amigo. También quisiera agradecerles especialmente a Miguel y a Adriana su apoyo y amistad durante todos estos años, y sobre todo su interés y sus ánimos mientras escribía la tesis. A Manuel, Héctor y Tiago por las charlas mantenidas sobre temas relacionados con mi tesis, y a veces, no tanto. Y en general, a todos los becarios que han ido pasando por el laboratorio y al resto de la gente del departamento, por haberme hecho el día a día más agradable. No quiero olvidarme

tampoco de Dani, Carlos, Lamas y Paradela, ni, claro está, de nuestra secretaria y una más del grupo, Cris.

Gracias a mi mejor amiga, Sonia, por los buenos ratos pasados durante la carrera en Vigo y por los buenos momentos que aún nos quedan por compartir. Gracias por estar siempre ahí.

A mis amigos de esgrima, en especial Jose, Carmen, Mari y Cris, por las bonitas experiencias que hemos vivido y por su amistad durante estos años. Por haber dejado tan alto el pabellón coruñés en tantas ocasiones y por su meritoria lucha por un deporte minoritario.

Por suposto, quero agradecerlles ós meus pais a súa vida enteira adicada ós seus fillos e o seu apoio e cariño dende que comencei este longo camiño na universidade. Por estar aí nos momentos difíciles e por saber que sempre estarán aí.

A Sara, a pequena, por ser como é, e para que siga sendo sempre así. A Susana e Ángel.

Ó meu avó, Darío, polo cariño que demostrou sempre polos seus netos e o orgulloso que falou sempre de nós.

A Roberto, o mellor xornalista deportivo da TvG, por ser sempre máis ca un curmán, case un irmán. A Pilar e José, polas vacacións de cando éramos cativos. A meus tíos, polas súas entrañables e numerosas comidas familiares ó longo de cada ano. Á miña familia, á casa Ancarés. A Lugo e Cospeito, por ser os lugares que máis quero e onde pasei algúns dos mellores momentos da miña vida.

A Juanita y Pepe, por ser como son con ochenta años. A Jose, Manolo, Toni, Fran, Pablo, Luis, Esther y Mari, por ser unos “cuñaos” tan estupendos. Y especialmente a Carmen y María Jesús, por su cariño y apoyo en la reciente situación personal por la que desgraciadamente tuvimos que pasar.

Quisiera expresar mi más profundo agradecimiento al Prof. José María Domínguez, por su gran confianza en mí durante todos estos años, por su enorme apoyo en el duro camino de la docencia y por ser más que un compañero, por ser un amigo en todo momento. Gracias por haberme enseñado tanto sobre Electrónica y por lo que me gustaría aún poder aprender de él. A ti te dedico este trabajo, por todo lo que hemos compartido en Náutica durante estos seis años.

Y por último, y no por eso el menos importante, quiero agradecer esta tesis a mi marido, Juan, por su gran apoyo durante todos estos años y por los ánimos que durante los ocho últimos meses me ha dado para lograr escribirla. Gracias por ser como eres y por la vida que hemos empezado juntos un mayo de hace más de dos años. No podría enumerar todas las cosas por las que me gustaría dedicarte esta tesis, así que, sinceramente, gracias por todo y a ti te la dedico.

Resumen

En este trabajo consideramos un sistema multiusuario con múltiples antenas en transmisión y una única antena en cada uno de los usuarios receptores y que se denota por brevedad como MU-MISO, del inglés *Multi–User Multiple–Input/Single–Output*. Este modelo MU–MISO se ajusta perfectamente al enlace descendente de un sistema de comunicaciones móviles, donde múltiples antenas situadas en la estación base envían información a varios usuarios dentro de su zona de cobertura y cuyos terminales móviles disponen generalmente de una única antena. Este canal descendente se denomina también canal de difusión (BC, del inglés *Broadcast Channel*). Cuando se considera un canal de difusión, el transmisor centralizado tiene claramente más grados de libertad que cada uno de los receptores descentralizados, por lo que es más apropiado separar las señales aplicando *precodificación* en transmisión. Para poder realizar el diseño de los parámetros del precodificador, el transmisor necesita conocer la información de canal (CSI, en inglés *Channel State Information*) correspondiente a los distintos usuarios receptores. En el caso de sistemas FDD (del inglés, *Frequency Division Duplex*), esta información puede obtenerse (al menos parcialmente) mediante realimentación, siempre tras haber aplicado un proceso de cuantificación de la información enviada con el objetivo de adaptarse a las condiciones de ancho de banda limitado del canal de retorno.

La asunción estándar para el diseño del retorno es CSI libre de errores en los usuarios receptores (e.g., [1–5]), pero los receptores consiguen su CSI mediante estimación, por lo que, evidentemente, ésta contiene errores. Así, para optimizar la CSI realimentada será necesario obtener una adecuada caracterización estadística de los errores. A lo largo de esta trabajo se considerarán las siguientes fuentes de error: estimación de canal, truncamiento (reducción de rango), cuantificación, y retardo inherente al envío de la información por el canal de retorno. Consideramos, sin embargo, que el canal de retorno no sufre errores durante la transmisión.

Como primera aproximación, planteamos un diseño basado en una métrica CSI-MSE, es decir, los parámetros de la realimentación se van a obtener mediante la minimización del error cuadrático medio (MSE, del inglés *Mean Squared Error*) entre el canal verdadero y el canal erróneo o ruidoso enviado desde cada uno de los usuarios receptores al transmisor. Los filtros del precodificador, sin embargo, se obtienen a partir de una optimización MSE independiente de la anterior. Se propone, por lo tanto, una optimización conjunta de la estimación, la reducción de rango y los parámetros de la librería, disponible tanto en el transmisor como en el receptor. Con ello tendremos el interesante resultado de que tanto la estimación como la reducción de rango obtenidas de esta formulación son independientes de la librería, y que ésta va a poder computarse *off-line* mediante el algoritmo de Lloyd. El rendimiento final en términos de BER (del inglés, *Bit Error Rate*) puede mejorarse, como veremos, mediante el algoritmo propuesto

para la asignación dinámica de los bits asociados al proceso de cuantificación evaluando de forma sencilla su impacto en el MSE obtenido.

Como segunda aproximación, presentamos el diseño conjunto de los estimadores de canal y los cuantificadores junto con el precodificador basado en un único criterio orientado al precodificador en lugar del criterio CSI–MSE aplicado en los primeros esquemas. Por lo tanto, la optimización planteada consiste en minimizar el MSE entre los símbolos transmitidos y los símbolos recuperados en recepción. Las entradas de la librería son ahora los posibles filtros de precodificación, de forma que cada usuario realimenta el índice correspondiente a un conjunto de precodificadores y la intersección de estos conjuntos realizada en transmisión va a proporcionar el precodificador óptimo empleado mientras no varíen los estadísticos del canal.

Las simulaciones realizadas con MATLAB nos muestran que la precodificación robusta basada en CSI imperfecta enviada desde los usuarios receptores presenta un claro mejor rendimiento que la precodificación convencional que no tiene en cuenta esos errores en la CSI. También se observa que estos diseños robustos son especialmente cruciales en sistemas que emplean precodificación no lineal con un canal de retorno limitado, puesto que son más sensibles a errores en la CSI. Si efectuamos una comparación entre las dos aproximaciones propuestas, claramente un diseño orientado al precodificador lleva a mejores resultados en términos de BER a costa de incrementar notablemente la complejidad computacional del algoritmo robusto.

La metodología de trabajo seguida en el desarrollo de la presente Tesis Doctoral ha consistido fundamentalmente en definir una lista de tareas, teniendo en cuenta tanto los trabajos previos como los recursos disponibles; determinar a continuación su secuencia u orden de ejecución, estableciendo una duración aproximada; organizar estas tareas por bloques de cierta entidad que definan etapas; y, finalmente, fijar los objetivos concretos de cada etapa y la metodología de trabajo a emplear para alcanzarlos. En la Tesis Doctoral se ha realizado una exhaustiva revisión bibliográfica y, tras evaluar las aportaciones realizadas durante los últimos años por la autora de la misma en congresos y revistas del ámbito de conocimiento, sólo aquellas más relevantes han sido incluidas finalmente en este trabajo. Tal y como se ha mencionado antes, los resultados de simulación por ordenador realizados sobre un clúster de 10 PCs del Laboratorio de Cálculo del grupo GTEC de la Universidad de A Coruña han sido obtenidos utilizando el lenguaje de programación técnica de alto nivel MATLAB.

Summary

In this work, we consider a multiuser system with a transmitter equipped with multiple antennas and only one antenna at each receiver user. This system, which is termed MU-MISO (*Multi–User Multiple–Input/Single–Output*), is of use to model the downlink of a wireless communications system, where multiple antennas at the base station transmit to several users with usually only one antenna at each receiving unit. This downlink channel is also called *Broadcast Channel* (BC). When considering this broadcast channel, the centralized transmitter clearly has more degrees of freedom than each of the receivers. Therefore, it is appropriate to separate the signals by applying precoding at the transmitter. To be able to design precoding, the transmitter needs knowledge about the channel states of the different receivers. In the case of *Frequency Division Duplex* (FDD) systems, this knowledge can be obtained by feedback (at least partially), where the *Channel State Information* (CSI) of the receivers is quantized to adapt to the limited rate conditions of the feedback channel.

The standard assumption for feedback design is error-free CSI at the receivers (e.g. [1–5]), but the receivers get their CSI by estimation. Thus, it contains errors. In order to properly design the limited feedback, it is necessary to obtain an adequate statistical characterization of the CSI errors. The following sources of error are considered throughout this work: channel estimation, truncation (rank reduction), quantization, and feedback channel delay. It is assumed, however, that the feedback channel does not suffer from errors during the transmission.

As a first approach, we propose a design based on a CSI-MSE metric, i.e. the feedback parameters are found by means of the minimization of the *Mean Squared Error* (MSE) between the true channel and the erroneous channel sent from the receiver side to the transmitter. The precoder filters, however, are obtained by means of a different minimum squared error optimization. In other words, we propose a joint optimization of the estimation, the rank reduction, and the codebook used for the feedback, available at both the transmitter and the receiver side. Interestingly, the estimator and the rank reduction resulting from this formulation are independent of the codebook used, which can be computed off–line with the generalized Lloyd algorithm. As we will see, the results in terms of *Bit Error Rate* (BER) can be improved by the algorithm proposed to dynamically allocate the bits associated to the quantization process by means of easily computing the obtained MSE.

As a second approach, we jointly design the channel estimators and the quantizers at the receivers together with the precoder at the transmitter based on a single criterion oriented to the precoder instead of the CSI–MSE criterion applied for the first approach. Therefore, this optimization consists of minimizing the MSE between the symbols

transmitted and recovered by each user. The codebook entries are now the possible precoder filters so that each receiver feeds back the index corresponding to a set of precoders and the intersection of the sets gives the optimum precoder to be used while channel statistics remain unchanged.

Several simulations carried out using MATLAB show that robust precoding based on fed–back information clearly outperforms conventional precoding that does not take into account the errors in the CSI. Additionally, we observe that a robust design is especially crucial for systems employing non-linear precoders with scarce feedback rate. Some comparisons between the above–mentioned approaches show that a limited feedback design involving the precoder in the MSE optimization exhibits better performance compared to the isolated precoder optimization, although the computational complexity is much higher.

Index

List of figures

List of tables

Chapter 1

Introduction

This work focuses on *Multiuser MISO* (MU–MISO) systems where a centralized transmitter equipped with multiple antennas communicates with several single antenna decentralized receivers. MU–MISO systems typically arise in the downlink of cellular communication systems. The transmitter is the *Access Point* (AP) or *Base Station* (BS), which admits more complexity and can support several antennas. The receivers are *Mobile Stations* (MS) with limited power consumption, size, and processing capabilities, and they will support a single antenna at the most. Recently, it has been shown that the *Dirty Paper Coding* (DPC) [6] signaling technique designed according to the *Signal-to-Interference-plus-Noise Ratio* (SINR) criteria is able to approach the sum capacity of a broadcast channel [7,8]. These contributions, however, only consider the ideal case where the CSI at the transmitter is perfectly known, similar to [9–11]. In the more practical case, where only an estimate of the CSI is available, the capacity region of the broadcast channel has not yet been found. Furthermore, the application of DPC is questionable, since it is unclear up to now how to systematically include the uncertainties in the SINR criterion (see the discussion in [12] and the attempt in [13] for the case of statistical CSI).

As shown in [14], the SINR and the MSE achievable regions for MU-MISO systems are closely related. Additionally, *minimum MSE* (MMSE) allows for a robust precoder design by considering a conditional expectation of the cost function [15–18]. Hence, we concentrate on MMSE precoder design. Based on the MMSE design for linear precoding as in [19, 20], for THP in [11, 21], and for VP in [22, 23], we develop robust linear precoding, robust THP, and robust VP, taking the expectation of MSE conditional on the available CSI.

Most of the work on precoding with erroneous CSI was motivated by a *Time Division Duplex* (TDD) setup, where the transmitter can estimate the CSI during the transmission in the opposite direction (e.g. [17, 18]). This approach, however, is difficult due to the need for very good calibration (e.g. [24]). Contrarily, we focus on the more difficult case, where the CSI is obtained by the receivers and fed back to the transmitter. In this case,

calibration errors are estimated as being part of the CSI, and therefore no special problems arise from calibration. Additionally, the feedback of CSI enables precoding in *Frequency Division Duplex* (FDD) systems, where the transmitter is unable to obtain the CSI during reception, because the channels are not reciprocal.

Since the data rate of the feedback channels is limited (e.g. [25]), the CSI must be compressed to ensure that the tight scheduling constraints are satisfied. Moreover, when the CSI is not perfectly known, it is a matter of discussion what kind of information has to be sent from the receiver to the transmitter and the way it is recovered at the transmitter side.

In the limited feedback systems proposed in this work, we start by estimating the channel at the receivers using the observations of pilot symbols sent from all transmit antennas. This enables the receivers to estimate their respective vector channels. Then, we reduce the estimates to a low-dimensional representation by projecting them onto a basis, which depends only on the channel statistics. We assume that the channel statistics are also known to the transmitter. The coefficients are quantized prior to transmission over the feedback channel which also introduces a delay. We restrict ourselves to scalar quantization (uniform and non-uniform quantization) in order to obtain closed-form solutions for the robust designs. However, in order to illustrate the trade-off between performance and complexity achieved with scalar quantizers, we also show how vector quantization can be applied in our limited feedback design.

Basically, we consider two types of limited feedback systems, namely, those systems that are based on minimizing the MSE between the true channel and the erroneous channel available at the transmitter, i.e. based on a CSI metric, and those systems that are based on a metric oriented to the precoder, i.e. that minimize the MSE between the transmitted symbols and the symbols recovered by the users and that therefore include the precoder in the MSE optimization. The idea of the limited feedback based on CSI is to jointly optimize the estimator and the quantizer parameters (i.e. codebook entries and partition cells), although the precoders must be obtained by means of a separate MSE optimization. Contrary to this idea, we find the second type based on a new metric that is not derived from the MSE of the CSI, but from the MSE of the data transmission. We derive expressions for the optimum estimators, quantizer parameters, and precoders obtained from this joint optimization that clearly outperform the previous approaches based on the MSE of the CSI. We also develop a strategy to optimally allocate the bits of each user in the sense of minimizing the MSE that results from each scheme.

On the other hand, in order to properly design robust precoders, it is necessary to obtain an adequate statistical characterization of the errors in the fed–back CSI. The following sources of error are considered: channel estimation, truncation (rank reduction), quantization, and feedback channel delay. Channel estimation and truncation errors are Gaussian and their analysis follows a conventional MSE approach (e.g. [26]). Since the

delayed channel versions fed back to the transmitter after estimation and truncation are also Gaussian, we can also easily obtain their statistical properties. Taking into account an initial codebook designed according to the Lloyd algorithm, we obtain an expression for the probability density function of the channel vector according to a Bayesian framework, i.e. conditional on the delayed, truncated, and quantized channel estimate. The expression found for this conditional PDF of the channel enables us to find closed-form expressions for the robust precoders.

1.1 Thesis Overview

This thesis is organized as follows:

In Chapter 2, we introduce the concepts of multipath and fading useful for understanding the correlated channel model described in this chapter, and which will be used throughout this work: the *Spatial Channel Model* (SCM). The signal model for the downlink of a multiuser system with multiple antennas at the transmitter is also presented in this chapter.

We review different types of receive and transmit processing in Chapter 3 where we assume that perfect *Channel State Information* (CSI) is available at the transmitter for precoding and at the receiver for detection.

However, this assumption is not realistic since the transmitter has no full channel knowledge. In Chapter 4, we describe the error sources appearing as a result of the estimation and CSI compression performed by each user to limit the overhead of the feedback channel.

In Chapter 5, we derive the MMSE robust precoder design to compensate the mismatch between the true channel and the erroneous channel available at the transmitter, in order to construct the precoder filters. Additionally, the MMSE receivers used instead of the common weights obtained from the optimizations are derived in this chapter for each type of precoder.

In Chapter 6, we investigate the design of the limited feedback, i.e. how to take into account the estimation, truncation, quantization, and feedback delay processes, in order to minimize the MSE between the true channel and the erroneous channel available at the transmitter.

Chapter 7 includes the precoder design in the MMSE joint feedback optimization, so that now the MSE between the transmitted symbols and the recovered symbols at each user is minimized.

Finally, Chapter 8 is dedicated to the conclusions and future work.

1.2 Assumptions and Notation

All derivations are based on the assumption of perfect knowledge of the second-order statistics of the noise, the symbols, and the channels. However, these parameters have to be estimated in practice, although we will not deal with this problem in this work. Finally, we assume that all random variables are zero–mean and stationary.

Vectors and matrices are denoted by lower case bold and capital bold letters, respectively. The $K \times K$ identity matrix is denoted by $\mathbf{I}_K$ and $\mathbf{0}_K$ is a K-dimensional zero vector. We use $\mathrm{E}[\bullet]$, $\Re(\bullet)$, $\Im(\bullet)$, $\mathrm{tr}(\bullet)$, $(\bullet)^*$, $(\bullet)^{\mathrm{T}}$, $(\bullet)^{\mathrm{H}}$, $\det(\bullet)$, $\otimes$, $*$, $\| \bullet \|_2$, and $\| \bullet \|_{\mathrm{F}}$ for expectation, real and imaginary part of the argument, trace of a matrix, complex conjugation, transposition, conjugate transposition, determinant of a matrix, Kronecker product, convolution, Euclidean norm, and Frobenius norm, respectively. The i-th element of a vector $\boldsymbol{x}$ is x_i. With $f_{\mathrm{G}}(\boldsymbol{x}, \boldsymbol{\mu}_x, \boldsymbol{C}_x)$, we refer to a circularly symmetric complex Gaussian distribution of $\boldsymbol{x} \in \mathbb{C}^m$ with the mean $\boldsymbol{\mu}_{\boldsymbol{x}} \in \mathbb{C}^m$ and the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}} \in \mathbb{C}^{m \times m}$, i.e.

$$f_{\mathrm{G}}\left(\boldsymbol{x}, \boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{C}_{\boldsymbol{x}}\right) = \frac{\exp\left(-\left(\boldsymbol{x} - \boldsymbol{\mu}_x\right)^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \left(\boldsymbol{x} - \boldsymbol{\mu}_x\right)\right)}{\pi^m \det(\boldsymbol{C}_{\boldsymbol{x}})}.$$

Chapter 2

Signal Model

In wireless communications systems the channel is time-variant, and it is thus very hard to find out how to predict future variations. This does not happen in wired communications where the channel remains almost unchanged. In this work, we focus on outdoor channels, whose analysis is no easy matter. The task of channel modeling is one of the most difficult parts in the design of wireless systems. The channel can be statistically modeled based on experimental measurements that are performed adapted to a given propagation environment. Since we exploit spatial and time correlations of the channel to design the optimum limited feedback, we introduce in this chapter characteristics common to most of the radio propagation environments, so we can talk about some general channel features.

A signal propagating through a wireless channel arrives at the destination along different paths. This phenomenon is known as multipath effect. The different paths arise from scattering, reflection, and diffraction of the radiated energy of objects in the environment or refraction in the medium. Multipath propagation results in the spreading of the signal over the different dimensions: time, frequency, and arrival angle. Correspondingly we have delay spread, Doppler spread and angle spread.

Additionally, the received signal level exhibits fluctuations, termed *fading*. Variations in the signal are due to three effects: mean path loss, macroscopic fading, and microscopic fading. The mean path loss depends on the distance between the transmitter and the receiver; on the antenna characteristics; and on the average attenuation introduced by the channel. Macroscopic fading, also termed *long-term channel variations* or *shadowing*, results from the type of scenario between the transmitter and the receiver, while microscopic fading results from destructive and constructive combination of the different paths, and is also known as *short-term fading*.

Mean path loss, macroscopic fading, microscopic fading, delay spread, Doppler spread, and angle spread are the main channel effects, and are described below.

2.1 Multipath Effects

In wireless communication systems, the transmitted signal typically propagates via several different paths from the transmitter to the receiver. This effect, termed *multipath propagation*, is caused by reflections of the radio waves from the surrounding obstacles. Let the transmitted signal in continuous time domain be given by [27–29]

$$s(t) = \Re\left\{u(t)\mathrm{e}^{\mathrm{j}\,2\pi f_{\mathrm{c}}t}\right\} = \Re\left\{u(t)\right\}\cos\left(2\pi f_{\mathrm{c}}t\right) - \Im\left\{u(t)\right\}\sin\left(2\pi f_{\mathrm{c}}t\right) \tag{2.1}$$

where t is time in seconds, $u(t)$ is the equivalent lowpass signal for $s(t)$, and f_{c} is the carrier frequency in Hz. $\Re(\bullet)$ and $\Im(\bullet)$ denote, respectively, the real and imaginary part of its argument. If we neglect the noise, the received signal is obtained by convolving the equivalent lowpass input signal $u(t)$ with the equivalent lowpass time-varying channel response to an impulse at time t, $h(t)$, and then upconverting to the carrier frequency:

$$r(t) = \Re\left\{\left[u(t) * h\left(t\right)\right]\mathrm{e}^{\mathrm{j}\,2\pi f_{\mathrm{c}}t}\right\}. \tag{2.2}$$

The equivalent lowpass time-varying channel $h(t)$ is modeled as the sum of the *Line-Of-Sight* (LOS) path and each of the multipath components, i.e.

$$h\left(t\right) = \sum_{m=0}^{M} \alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}\delta(t-\tau_m(t)) \tag{2.3}$$

where the phase shift $\psi_m(t)$ is given by

$$\psi_m\left(t\right) = 2\pi f_{\mathrm{c}}\tau_m(t) - \psi_{\mathrm{D},m}. \tag{2.4}$$

Hence, $\psi_{\mathrm{D},m}$ is the *Doppler phase shift* for each multipath component obtained as

$$\psi_{\mathrm{D},m} = \int_t 2\pi f_{\mathrm{D},m}\left(t\right)\mathrm{d}t$$

with $f_{\mathrm{D},m}(t)$ known as *Doppler frequency shift* and expressed as follows

$$f_{\mathrm{D},m}\left(t\right) = \frac{v\cos\left(\theta_m(t)\right)}{\lambda} \tag{2.5}$$

where v is the velocity of the mobile and λ is the wavelength. $\theta_m(t)$ is the angle of arrival of each multipath component relative to the direction of motion. Note that the component $m = 0$ in Eq. (2.3) corresponds to the LOS path. The number of multipath components is given by M. In general, each path has different relative propagation delays (given by $\tau_m(t)$ in Eq. (2.3)), different amplitudes or attenuations for each path ($\alpha_m(t)$) and different phases (given by $\psi_m(t)$ in Eq. (2.3)). We assume that $\alpha_m(t)$, $\tau_m(t)$, and

$\psi_m(t)$ are stationary and ergodic. Thus, the received signal will also be a stationary and ergodic random process.

Remember that the convolution of two functions f and g is defined as $f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\mathrm{d}\tau$. Substituting Eq. (2.3) into Eq. (2.2), we obtain the received signal

$$\begin{aligned} r(t) &= \Re\left\{\left[\int_{-\infty}^{\infty} h(\tau,t)u\left(t-\tau\right)\mathrm{d}\tau\right]\mathrm{e}^{\mathrm{j}\,2\pi f_c t}\right\} \\ &= \Re\left\{\left[\sum_{m=0}^{M}\alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}\delta(\tau-\tau_m(t))u\left(t-\tau\right)\mathrm{d}\tau\right]\mathrm{e}^{j2\pi f_c t}\right\} \\ &= \Re\left\{\left[\sum_{m=0}^{M}\alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}\left(\int_{-\infty}^{\infty}\delta(\tau-\tau_m(t))u\left(t-\tau\right)\mathrm{d}\tau\right)\right]\mathrm{e}^{\mathrm{j}\,2\pi f_c t}\right\} \\ &= \Re\left\{\left[\sum_{m=0}^{M}\alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}u\left(t-\tau_m(t)\right)\right]\mathrm{e}^{\mathrm{j}\,2\pi f_c t}\right\} \end{aligned} \tag{2.6}$$

where τ is the variable of delay and $h(\tau,t)$ represents the equivalent lowpass response of the channel at time t to an impulse at time $t-\tau$ as follows

$$h\left(\tau,t\right) = h\left(t\right) * \delta\left(t-\tau\right) = \sum_{m=0}^{M}\alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}\delta(\tau-\tau_m(t)). \tag{2.7}$$

Last equality in Eq. (2.6) is obtained from the shift property of the Dirac distribution

$$\int_{-\infty}^{\infty}\delta(\tau-\tau_m(t))u\left(t-\tau\right)\mathrm{d}\tau = \delta(t-\tau_m(t)) * u\left(t\right) = u\left(t-\tau_m(t)\right).$$

Multiple antennas at the transmitter and/or the receiver is becoming a common feature of wireless systems since diversity and capacity benefits increase with the number of antennas. Systems with multiple antennas require channel models to characterize both spatial and temporal correlations of the channel. Therefore, we now consider a multipath environment in which the receiver or transmitter has an antenna array with P elements. And also assuming that the *Angle-of-Arrival* (AoA, given by $\theta_m(t)$) is stationary and identically distributed for all multipath components, and denoting this random AoA with respect to the origin of the array by θ, we can introduce the angle dimension in Eq. (2.7) as follows

$$h\left(\tau,t,\theta\right) = \sum_{m=0}^{M}\alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}a_p\left(\theta\left(t\right)\right)\delta(\tau-\tau_m(t)) \tag{2.8}$$

where $a_p(\theta(t)) \in \mathbb{C}$ is the p–th element of the antenna array vector expressed as

$$a_p\left(\theta\left(t\right)\right) = \mathrm{e}^{-\mathrm{j}\,\frac{2\pi}{\lambda}(x_p\cos\theta(t)+y_p\sin\theta(t))} \tag{2.9}$$

for (x_p, y_p) indicating the antenna location relative to the origin of the array.

Then, the received signal in Eq. (2.2) can be expressed as

$$\begin{aligned}\boldsymbol{r}(t) &= \Re\left\{\left[\int_{-\infty}^{\infty}\sum_{m=0}^{M}\alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}\boldsymbol{a}\left(\theta\left(t\right)\right)\delta(\tau-\tau_m(t))u\left(t-\tau\right)\mathrm{d}\tau\right]\mathrm{e}^{j2\pi f_c t}\right\}\\ &= \Re\left\{\left[\sum_{m=0}^{M}\alpha_m(t)\mathrm{e}^{-\mathrm{j}\,\psi_m(t)}\boldsymbol{a}\left(\theta\left(t\right)\right)u\left(t-\tau_m(t)\right)\right]\mathrm{e}^{\mathrm{j}\,2\pi f_c t}\right\}. \qquad (2.10)\end{aligned}$$

We assume some conditions about the channel impulse response $h(\tau, t, \theta)$ in Eq. (2.8). First, we consider that the channel is *Wide Sense Stationary* (WSS), i.e. the temporal channel autocorrelation depends only on time difference, i.e.

$$R_{\mathrm{h}}\left(\tau, t, \theta\right) = \mathrm{E}\left[h\left(\tau, t_0, \theta\right)h^*\left(\tau, t_0+t, \theta\right)\right] = \mathrm{E}\left[h\left(\tau, 0, \theta\right)h^*\left(\tau, t, \theta\right)\right]$$

for all τ and θ. We also assume that fading corresponding to different obstacles is uncorrelated (this is called *Uncorrelated Scattering*, US), i.e.

$$\mathrm{E}\left[h\left(\tau_1, t, \theta\right)h^*\left(\tau_2, t, \theta\right)\right] = 0 \quad \text{if } \tau_1 \neq \tau_2$$

for all t and θ. When the channel satisfies both conditions, it is termed *Wide Sense Stationary Uncorrelated Scattering* (WSSUS).

We can define the Fourier transform of the time autocorrelation of the channel response $h(\tau, t, \theta)$ as the function given by

$$\phi\left(\tau, f, \theta\right) = \int_{\infty}^{\infty} R_{\mathrm{h}}\left(\tau, t, \theta\right)\mathrm{e}^{-\mathrm{j}\,2\pi f t}\mathrm{d}t.$$

This function $\phi(\tau, f, \theta)$ is the channel description in the frequency, time, and angle domain. In this context, the variable f is termed *Doppler frequency*. The average channel power as a function of the Doppler frequency is obtained as

$$\phi_{\mathrm{D}}\left(f\right) = \int_{-\pi}^{\pi}\int_{0}^{\tau_{\max}}\phi\left(\tau, f, \theta\right)\mathrm{d}\tau\mathrm{d}\theta \qquad (2.11)$$

which is called *Doppler power spectrum*. The time interval between the instant of arrival of the first multipath component and that of the last one is denoted by $\tau_{\max}$. θ is regarded as a uniform variable on $[-\pi, \pi]$. The spectral spreading covers the range $f \in [f_{\mathrm{c}} - f_{\mathrm{D,\,max}}, f_{\mathrm{c}} + f_{\mathrm{D,\,max}}]$, where the maximum Doppler frequency, $f_{\mathrm{D,\,max}}$, is related to the relative velocity between the transmitter and the receiver and is obtained when $\cos(\theta_m(t)) = 1$ in Eq. (2.5) which leads to

$$f_{\mathrm{D,\,max}} = f_{\mathrm{c}}\frac{v}{c} \qquad (2.12)$$

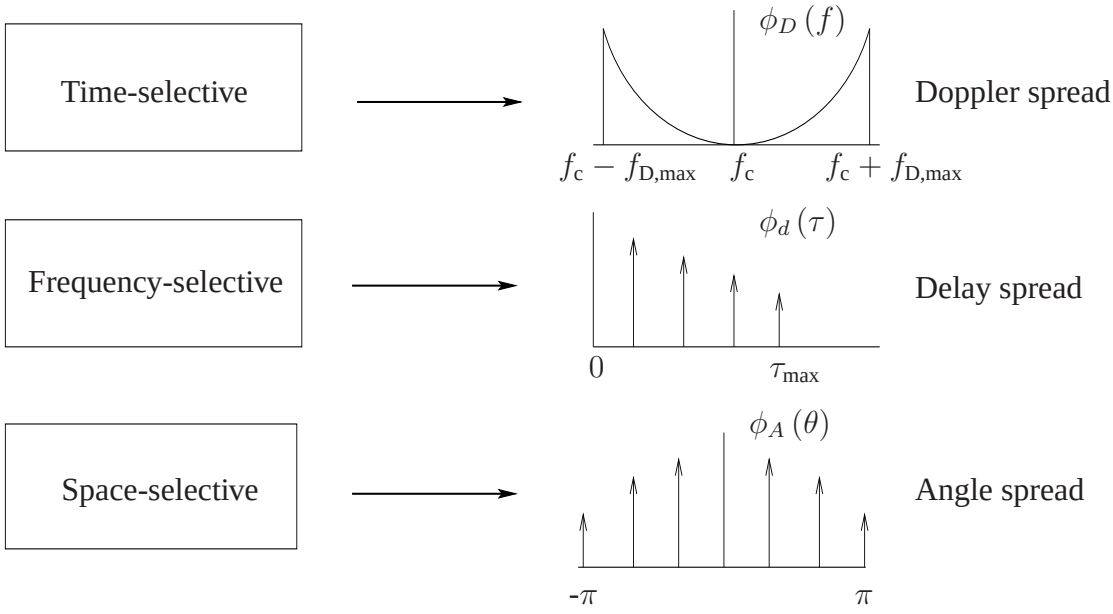


Figure 2.1: Space-, Frequency- and Time-Selective Fading.

where c is the constant for the speed of light.

Similarly, we obtain the *delay power spectrum* or average channel power as a function of the delay (τ) taking the marginal integral over the other two variables, i.e.

$$\phi_{\mathrm{d}}(\tau) = \int_{-\pi}^{\pi} \int_{f_{\mathrm{c}}-f_{\max}}^{f_{\mathrm{c}}+f_{\max}} \phi(\tau, f, \theta)\, \mathrm{d}f \mathrm{d}\theta. \tag{2.13}$$

The delay power spectrum is also commonly referred to as multipath intensity profile.

Finally, the average power as a function of the angle of arrival is obtained as follows,

$$\phi_{\mathrm{A}}(\theta) = \int_{f_{\mathrm{c}}-f_{\max}}^{f_{\mathrm{c}}+f_{\max}} \int_{0}^{\tau_{\max}} \phi(\tau, f, \theta)\, \mathrm{d}\tau \mathrm{d}f \tag{2.14}$$

which denotes the *angle power spectrum*.

2.1.1 Delay Spread and Frequency-Selective Fading

In a multipath propagation environment, the receiver gets some scaled and delayed versions of the transmitted signal. If the signal only suffers from attenuation (i.e. there are no delayed components), all frequency components of the signal will experience the same magnitude of fading. This effect is termed *flat fading*. If several delayed signals arrive at the receiver, then different frequency components of the signal experience decorrelated

fading, which is typically termed *frequency-selective fading* (see Fig. 2.1). The *delay spread*, τ_{RMS}, is defined as the *Root Mean Square* (RMS) delay of $\phi_{\text{d}}(\tau)$ (Fig. 2.1), i.e.

$$\tau_{\text{RMS}} = \sqrt{\text{E}\left[(\tau - \bar{\tau})^2\right]} = \sqrt{\frac{\int_0^{\tau_{\max}} (\tau - \bar{\tau})^2 \, \phi_{\text{d}}(\tau) \, \text{d}\tau}{\int_0^{\tau_{\max}} \phi_{\text{d}}(\tau) \, d\tau}}$$

where $\bar{\tau}$ is the average delay defined as

$$\bar{\tau} = \frac{\int_0^{\tau_{\max}} \tau \phi_{\text{d}}(\tau) \, \text{d}\tau}{\int_0^{\tau_{\max}} \phi_{\text{d}}(\tau) \, \text{d}\tau}$$

and $\phi_{\text{d}}(\tau)$ is the delay power spectrum defined in Eq. (2.13). When the distance between two frequencies is greater than the inverse of the delay spread, these two frequencies experience a totally different attenuation by the channel. If the separation is less than the inverse of the delay spread, then they suffer from similar attenuation. Therefore, the frequency-selective channel characteristic depends on the bandwidth of the transmit signal compared to the inverse of the delay spread, the so called *channel coherence bandwidth*, B_c, i.e.

$$B_{\text{c}} \approx \frac{1}{\tau_{\text{RMS}}}.$$

Signals with bandwidth smaller than the channel coherence bandwidth suffer from flat frequency attenuation. Signals with bandwidth greater than the channel coherence bandwidth experience different attenuations according to the frequency, i.e. they experience frequency-selective fading.

2.1.2 Doppler Spread and Time-Selective Fading

Another important channel characteristic is concerned with the relative mobility between the transmitter and the receiver. When a user (or scatterers in the surroundings) is in motion, the user's velocity causes a shift in the frequency of the signal transmitted along each signal path.

Signals traveling along different paths can have different Doppler frequency shifts and, therefore, different Doppler phase shifts. The difference in Doppler shifts between different channel components is known as *Doppler spread*. The Doppler spread is a result of the mobile terminal movement during the communication. To be precise, the Doppler spread, similarly to delay spread, is defined as the RMS bandwidth of $\phi_{\text{D}}(f)$ (see Fig. 2.1), i.e.

$$f_{\text{RMS}} = \sqrt{\text{E}\left[\left(f - \bar{f}\right)^2\right]} = \sqrt{\frac{\int_{f_{\text{c}}-f_{\max}}^{f_{\text{c}}+f_{\max}} \left(f - \bar{f}\right)^2 \phi_{\text{D}}(f) \, \text{d}f}{\int_{f_{\text{c}}-f_{\max}}^{f_{\text{c}}-f_{\max}} \phi_{\text{D}}(f) \, \text{d}f}}$$

where $\bar{f}$ is the average frequency of the Doppler spectrum defined as

$$\bar{f} = \frac{\int_{f_c - f_{max}}^{f_c + f_{max}} f \phi_D(f)\, df}{\int_{f_c - f_{max}}^{f_c + f_{max}} \phi_D(f)\, df}$$

and $\phi_D(f)$ is the Doppler power spectrum given by Eq. (2.11). This causes the overall radio channel to be time-variant, i.e. with time-varying delays and attenuations for the individual multipath components. This phenomenon is generally termed *time-varying* or *time-selective fading*. The *coherence time*, denoted by T_c, is the time during which the channel behavior remains approximately unchanged. The coherence time T_c is inversely proportional to the Doppler spread, i.e.

$$T_c \approx \frac{1}{f_{RMS}}.$$

T_c measures the minimum time required for the channel magnitude to become decorrelated from its previous value. According to its time-selectivity, the channel is said to be *slow fading* if its coherence time is much greater than the frame duration. Otherwise, the channel is said to be *fast fading*, meaning that the channel changes considerably from one transmission frame to another.

The frequency- and time-selective nature of mobile wireless channels is one of the most critical elements from the point of view of overall wireless link quality. Various transmitter and/or receiver signal processing techniques are utilized in practice to overcome the time- and frequency-selective fading effects in practical communications systems, including, for example, various channel equalization, coding, and diversity transmission schemes.

2.1.3 Angle Spread and Space-Selective Fading

Angle spread at the receiver refers to the spread in *Angles of Arrival* (AoA) of the multipath components at the receive antenna array. Similarly, angle spread at the transmitter refers to the spread in *Angles of Departure* (AoD) for those multipath signals that finally reach the receiver (see Fig. 2.1). Note that we are talking only about AoAs, and not AoDs (Angles of Departure), since the downlink of a wireless communication system is considered. We define the RMS angle spread, θ_{RMS}, as

$$\theta_{RMS} = \sqrt{\mathrm{E}\left[\left(\theta - \bar{\theta}\right)^2\right]} = \sqrt{\frac{\int_{-\pi}^{\pi} \left(\theta - \bar{\theta}\right)^2 \phi_A(\theta)\, d\theta}{\int_{-\pi}^{\pi} \phi_A(\theta)\, d\theta}}$$

where $\phi_{\mathrm{A}}(\theta)$ is the angle power spectrum defined according to Eq. (2.14) and $\bar{\theta}$ is the average angle defined as

$$\bar{\theta} = \frac{\int_{-\pi}^{\pi} \theta \phi_{\mathrm{A}}(\theta)\, \mathrm{d}\theta}{\int_{-\pi}^{\pi} \phi_{\mathrm{A}}(\theta)\, \mathrm{d}\theta}.$$

Angle spread causes *space selective fading* which means that the received signal amplitudes depend on the antennas' spatial location. Space selective fading is characterized by the *coherence distance*, D_{c}, which is inversely proportional to the angle spread, i.e.

$$D_{\mathrm{c}} \propto \frac{1}{\theta_{\mathrm{RMS}}}.$$

Larger angle spreads imply shorter coherence distances. If the separation among the antenna elements is higher than the coherence time, the signal amplitude depends on the antenna location, and vice versa, i.e. if the separation is smaller than the coherence distance, signals arriving at the different antennas suffer from similar attenuations.

Although space-selectivity has not been as widely studied as time- or frequency-selectivity, this topic has achieved greater prominence in recent years due to the increasing number of antennas at both the transmitter and the receiver side.

2.2 Mean Path Loss

The path loss is the ratio between the transmitted power and the received power (see Fig. 2.2) given by

$$\frac{P_{\mathrm{r}}}{P_{\mathrm{t}}} = G_{\mathrm{t}} G_{\mathrm{r}} \frac{\lambda^2}{(4\pi)^2 d^{\gamma} L} \tag{2.15}$$

where P_{t} and P_{r} are the transmitted and received powers, respectively, d is the distance between the transmitter and the receiver, and G_{t} and G_{r} are the power gains for the transmit and receive antenna, respectively. L is related to the loss due to the antenna characteristics and the average channel attenuation. γ is the slope index from a value of 2 for free space to 6 depending on the environment. Some values for γ depending on the environment are shown in Table 2.1. Several empirical path loss models have been developed for microcellular and macrocellular systems, such as Okumura, Hata or Cost-231 models [30].

2.3 Fading

In Section 2.1 we explained that fading is the fluctuation in the received signal level caused by multipath propagation. Fading is due to two multiplicative phenomenons: microscopic

Environment	γ
Free space	2
Flat rural	3
Rolling rural	3.5
Suburban, low rise	4
Dense urban, skyscrapers	4.5

Table 2.1: Path Loss Exponents.

and macroscopic effects (see Fig. 2.2). They are also referred to as *short term* and *long term* channel variations, respectively, which are described in the following subsections.

2.3.1 Macroscopic Fading

Macroscopic fading is caused by changes in the scenario, i.e. alterations in the surrounding environment (rural, suburban, urban...) or as a result of the terrain configuration (open, flat, hilly, mountain...). The deviation of macroscopic fading about the mean propagation loss is treated as a random variable that is considered to be lognormal. Its probability density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.16}$$

where x is the random variable expressed in decibels (dB) that represents the long-term signal power level fluctuation over the mean path loss. The variables μ and σ are the mean and standard deviation of x, respectively. Both, μ and σ, are expressed in dB. The mean value, μ, is equal to the mean propagation loss discussed in the previous section. The standard deviation, σ, may have values around 8 dB for some environments (see the parameter σ_{SF} in Tables A.3, A.4, and A.5).

2.3.2 Microscopic Fading

In many practical situations the transmitter and the receiver are not within direct sight of each other. This situation is referred to as *Non-Line-Of-Sight* (NLOS) propagation. The received signal is the sum of multiple signals produced by reflection from the elements that surround the transmitter and the receiver. This produces rapid fluctuations over the mean of the received signal, this effect being called *microscopic fading* [27, 31].

To characterize the random scale factor caused by multipath, we choose $s(t)$ to become an unmodulated carrier given by

$$s(t) = \Re\left\{\mathrm{e}^{\mathrm{j}\,2\pi f_c t}\right\} = \cos\left(2\pi f_c t\right)$$

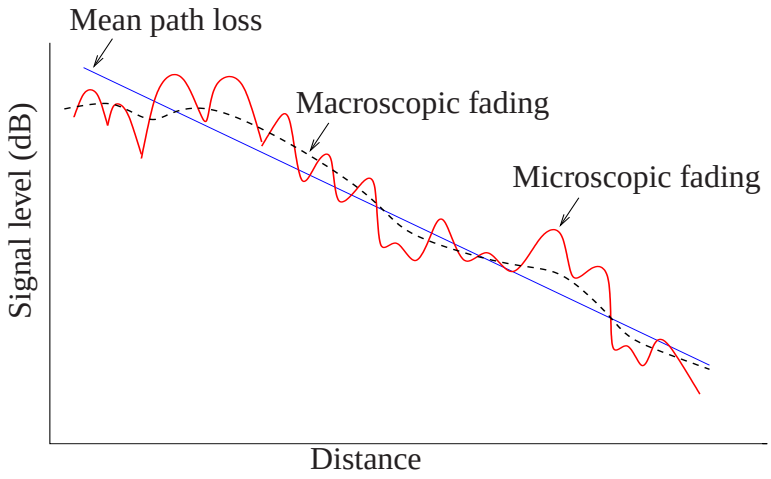


Figure 2.2: Macroscopic and Microscopic Fading.

and, therefore, $u(t)$ in Eq. (2.1) is equal to 1, for all t. Under most delay spread characterizations, the channel coherence bandwidth is much smaller than the inverse of the delay spread (see Subsection 2.1.1) which implies that the delay associated with the m–th multipath component $\tau_m(t) \leq \tau_{\text{RMS}}$ for all m and, then, we can consider a *narrowband fading model* where $u(t-\tau_m(t)) \approx u(t)$ for all m and $h(t)$ in Eq. (2.3) is now expressed as

$$h\left(t\right) = \sum_{m=0}^{M} \alpha_m(t) \mathrm{e}^{-\mathrm{j}\,\psi_m(t)} \delta(t). \tag{2.17}$$

Therefore, Eq. (2.6) can be rewritten as

$$r(t) = \Re\left\{ \left[\sum_{m=1}^{M} \alpha_m(t) \mathrm{e}^{-\mathrm{j}\,\psi_m(t)} \right] \mathrm{e}^{\mathrm{j}\,2\pi f_c t} \right\} = r_{\mathrm{I}}\left(t\right) \cos\left(2\pi f_c t\right) - r_{\mathrm{Q}}\left(t\right) \sin\left(2\pi f_c t\right)$$

where the in-phase and quadrature components are given by

$$r_{\mathrm{I}}\left(t\right) = \sum_{m=1}^{M} \alpha_m(t) \cos\left(\psi_m(t)\right)$$

$$r_{\mathrm{Q}}\left(t\right) = \sum_{m=1}^{M} \alpha_m(t) \sin\left(\psi_m(t)\right)$$

with the phase term $\psi_m(t)$ given by Eq. (2.4).

For large M we can apply the central limit theorem together with the fact that $\alpha_m(t)$ and $\psi_m(t)$ are independent for different components in order to approximate r_{I} and r_{Q} as a jointly Gaussian random process. The Gaussian property also holds if $\alpha_m(t)$ is Rayleigh

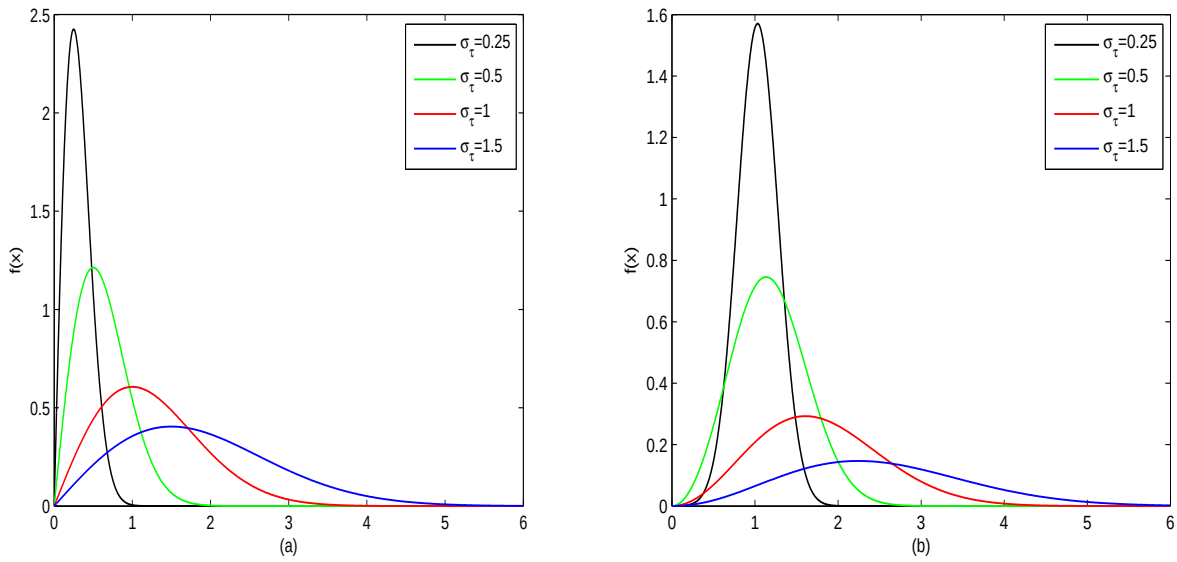


Figure 2.3: Microscopic Fading: (a) Rayleigh PDF; (b) Rice PDF with $\mu = 1$.

distributed and $\psi_m(t)$ is uniformly distributed over $[-\pi, \pi]$. Let σ^2 be the variance for both in-phase and quadrature components. Then, the signal envelope

$$x\left(t\right) = \left|r\left(t\right)\right| = \sqrt{r_{\mathrm{I}}^2(t) + r_{\mathrm{Q}}^2(t)}$$

follows a Rayleigh distribution with density function (see Fig. 2.3):

$$f\left(x\right) = \frac{2x}{\bar{P}_{\mathrm{r}}} \mathrm{e}^{-\frac{x^2}{\bar{P}_{\mathrm{r}}}} = \frac{x}{\sigma^2} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \quad x \geq 0$$

where $\bar{P}_{\mathrm{r}} = \sum_m \mathrm{E}[\alpha_m^2] = 2\sigma^2$ is the average received signal power.

If there is a direct path present between the transmitter and the receiver, the signal envelope is no longer Rayleigh and the statistics of the signal amplitude follow a Rician distribution. Rician fading is formed by the sum of a Rayleigh distributed signal and a direct or line-of-sight signal. Now, the modulus of $r(t)$ is said to follow a Rician distribution and its PDF is given by (see Fig. 2.3)

$$f\left(x\right) = \frac{x}{\sigma^2} \mathrm{e}^{-\frac{x^2+\mu^2}{2\sigma^2}} I_0\left(\frac{x\mu}{\sigma^2}\right) \quad x \geq 0$$

where $I_0(\bullet)$ is the modified Bessel function of zero-th order. $\mu^2 = \alpha_0^2$ is the power of the LOS component and $2\sigma^2 = \sum_{m,m\neq 0} \mathrm{E}[\alpha_m^2]$ is the average power of the non-LOS multipath components. The average received power for Rician fading is obtained as

$$\bar{P}_{\mathrm{r}} = \int_0^\infty x^2 f(x) dx = \mu^2 + 2\sigma^2.$$

Thus, the Rician distribution can be alternatively expressed in terms of the K factor defined as the ratio of the power in the LOS component to the power of the scattered components, i.e.

$$K = \frac{\mu^2}{2\sigma^2}$$

which leads to the alternative expression for the PDF of the Rician distribution,

$$f(x) = \frac{2\,(K+1)\,x}{\bar{P}_{\mathrm{r}}} e^{\left(-K-\frac{(K+1)x^2}{\bar{P}_{\mathrm{r}}}\right)} I_0\left(2\sqrt{\frac{K\,(K+1)}{\bar{P}_{\mathrm{r}}}}x\right) \quad x \geq 0$$

by making the substitutions $\mu^2 = K\bar{P}_{\mathrm{r}}/(K+1)$ and $2\sigma^2 = \bar{P}_{\mathrm{r}}/(K+1)$. Since $I_0(0) = 1$, the Rician distribution reduces to the Rayleigh distribution when $K = 0$. On the contrary, when $K \to \infty$ we have no fading, i.e. there is no multipath but only a LOS component.

2.4 MIMO Systems

Fig. 2.4 shows a communication system employing N_{t} transmit antennas and N_{r} receive antennas, which is called a *Multiple-Input Multiple-Output* (MIMO) system. In MIMO communication systems [27, 28, 31, 32], the multiple data streams can be sent simultaneously from a transmitter employing multiple antennas to a receiver that employs multiple receive antennas. The goal of a MIMO system is to increase the data rate through spatial multiplexing and improving the error rate performance by increasing signal diversity (this being achieved by increasing the number of transmit or receive antennas, given that the probability of a fade at the same time in all the paths is reduced) to combat fading.

A MIMO system can be seen as a single-user point-to-point communication system. The special case with $N_{\mathrm{t}} = N_{\mathrm{r}} = 1$ is called a *Single-Input Single-Output* (SISO) system. A second special case is when $N_{\mathrm{t}} = 1$ and $N_{\mathrm{r}} \geq 2$ and is called a *Single-Input Multiple-Output* (SIMO) system. Lastly, there exists another special case if $N_{\mathrm{r}} = 1$ and $N_{\mathrm{t}} \geq 2$, called a *Multiple-Input Single-Output* (MISO) system.

In MIMO systems with N_{t} transmit antennas and N_{r} receive antennas, we denote the equivalent lowpass channel impulse response between the j–th transmit antenna and the i–th receive antenna as $h_{i,j}(\tau, t)$. Thus, the randomly time-varying channel is characterized by the $N_{\mathrm{r}} \times N_{\mathrm{t}}$ matrix $\boldsymbol{H}(\tau, t)$ defined as

$$\mathbf{H}(\tau,t) = \begin{pmatrix} h_{1,1}(\tau,t) & h_{1,2}(\tau,t) & \cdots & h_{1,N_{\mathrm{t}}}(\tau,t) \\ h_{2,1}(\tau,t) & h_{2,2}(\tau,t) & \cdots & h_{2,N_{\mathrm{t}}}(\tau,t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_{\mathrm{r}},1}(\tau,t) & h_{N_{\mathrm{r}},2}(\tau,t) & \cdots & h_{N_{\mathrm{r}},N_{\mathrm{t}}}(\tau,t) \end{pmatrix}.$$

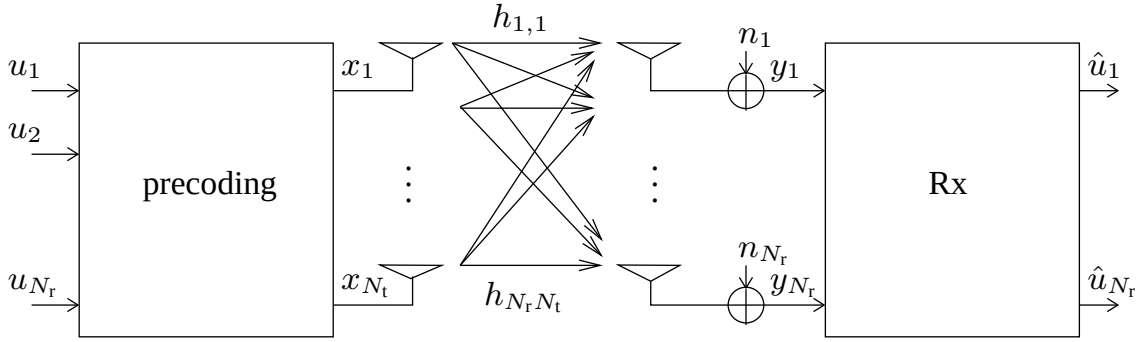


Figure 2.4: System with Precoding over Flat MIMO Channel.

Suppose that the transmitted signal from the i–th transmit antenna is $x_i(t)$. Then, the receive signal at the j–th receive antenna is given by

$$y_j(t) = \sum_{i=1}^{N_\mathrm{t}} h_{j,i}(\tau, t) * x_i(t) + \eta_j(t)$$

where $\eta_j(t)$ is the additive noise. In matrix notation, this equation can be rewritten as

$$\boldsymbol{y}(t) = \boldsymbol{H}(\tau, t) * \boldsymbol{x}(t) + \boldsymbol{\eta}(t)$$

where $\boldsymbol{x}(t) = [x_1(t), \ldots, x_{N_\mathrm{t}}(t)]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{t}}$, $\boldsymbol{y}(t) = [y_1(t), \ldots, y_{N_\mathrm{r}}(t)]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{r}}$, and $\boldsymbol{\eta}(t) = [\eta_1(t), \ldots, \eta_{N_\mathrm{r}}(t)]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{r}}$ [see Eq. (2.1)]. For flat fading channels (see Subsection 2.1.1), the channel matrix $\boldsymbol{H}(\tau, t)$ is transformed into the matrix $\mathbf{H}(t)$ given by

$$\mathbf{H}(t) = \begin{pmatrix} h_{1,1}(t) & h_{1,2}(t) & \cdots & h_{1,N_\mathrm{t}}(t) \\ h_{2,1}(t) & h_{2,2}(t) & \cdots & h_{2,N_\mathrm{t}}(t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_\mathrm{r},1}(t) & h_{N_\mathrm{r},2}(t) & \cdots & h_{N_\mathrm{r},N_\mathrm{t}}(t) \end{pmatrix}$$

and the received signal is now

$$y_j(t) = \sum_{i=1}^{N_\mathrm{t}} h_{ji}(t) x_i(t) + \eta_j(t)$$

which can be expressed in matrix form as

$$\boldsymbol{y}(t) = \boldsymbol{H}(t)\boldsymbol{x}(t) + \boldsymbol{\eta}(t). \tag{2.18}$$

In general, if we let $f[n] = f(nT_\mathrm{s} + \Delta)$ denote samples of $f(t)$ every T_s seconds with Δ being the sampling delay and T_s the symbol time, then sampling $\boldsymbol{y}(t)$ every T_s seconds yields the discrete time signal $\boldsymbol{y}[n] = \boldsymbol{y}(nT_\mathrm{s} + \Delta)$ given by

$$\boldsymbol{y}[n] = \boldsymbol{H}[q]\boldsymbol{x}[n] + \boldsymbol{\eta}(n) \tag{2.19}$$

where $n = 0, 1, 2, \ldots$ corresponds to samples spaced with T_s and q denotes the slot time. The channel remains unchanged during a block of N_B symbols, i.e, over the data frame. Note that this discrete time model is equivalent to the continuous time model in Eq. (2.18) only if ISI between samples is avoided, i.e. if the *Nyquist criterion* is satisfied. In that case, we will be able to reconstruct the original continuous signal from the samples by means of interpolation. This channel model is known as *time-varying flat block fading channels* (Subsection 2.1.1) and this assumption is made in the following.

2.5 Multiuser-MISO Systems

This work focuses on complex scenarios with multiple users and multiple communication links [27, 32]. We can distinguish between several types of multiuser communication systems. One type is the multiple access channel in which a large number of users share a common communication channel to transmit information to a single receiver. The common channel can represent the uplink in a cellular or a satellite communication system; or a cable to which a number of terminals are connected to access a central computer. For the example of a mobile cellular system, the users are the mobile terminals in a cell and the receiver side is the base station of the particular cell.

The second most common type of multiuser communication system is a broadcast channel in which a single transmitter sends information to multiple receivers (see Fig. 2.5). Examples of broadcast systems include the common radio and TV broadcast systems as well as the downlink of cellular and satellite communication systems. In this work, we focus on broadcasting methods for multiuser communications, in particular on the downlink of a cellular communication system where a base station with multiple antennas serves the corresponding cell and sends information to a number of mobile terminals in that area.

We consider a *Multi-User Multiple-Input Single-Output* (MU-MISO) system with $N_\mathrm{t} = N$ transmit antennas and K users equipped with a single antenna (i.e. $N_\mathrm{r} = 1$) as depicted in Fig. 2.5. As mentioned above, such a system is often referred to as the *broadcast* channel. Note that we work with the discrete model that is equivalent to the continuous one described in the previous section. Channel time variance is on a different scale to signal time variance since we consider a *block fading* channel, i.e. one that is considered to remain unchanged during the transmission of a frame of N_B symbols. Therefore, we will henceforth use q to indicate the time slot while n will be used to denote each one of the N_B time samples spaced with the symbol period, T_s, inside each slot.

The precoder generates the transmitted signal $\boldsymbol{x}[n]$ from all data symbols $\{u_1[n], \ldots, u_K[n]\}$ belonging to the different users $1, \ldots, K$. The signal $x_\ell[n]$ from

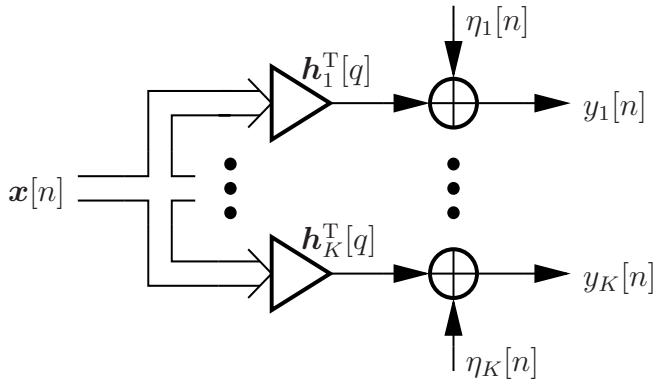


Figure 2.5: Vector BC with K Receivers.

transmit antenna ℓ propagates over the channel with the coefficient $h_{k,\ell}[n]$ to the k-th receiver, superimposes with the signals of the other transmit antennas, and is perturbed by the additive white Gaussian noise $\eta_k[n]$ with variance σ_η^2, i.e.

$$y_k[n] = \sum_{\ell=1}^{N} h_{k,\ell}[q] x_\ell[n] + \eta_k[n] = \boldsymbol{h}_k^{\mathrm{T}}[q]\boldsymbol{x}[n] + \eta_k[n] \tag{2.20}$$

where $\boldsymbol{h}_k[q] = [h_{k,1}[q], \ldots, h_{k,N}[q]]^{\mathrm{T}} \in \mathbb{C}^N$ represents the flat block fading vector channel corresponding to the k-th user and $\boldsymbol{x}[n] = [x_1[n], \ldots, x_N[n]]^{\mathrm{T}} \in \mathbb{C}^N$ is the transmit signal. The transmit signal $\boldsymbol{x}[n]$ must satisfy an average total transmit power constraint, i.e. $\mathrm{E}[\|\boldsymbol{x}[n]\|_2^2] = E_{\mathrm{tx}}$. By combining Eq. (2.20) for $k = 1, \ldots, K$, we get

$$\boldsymbol{y}[n] = \boldsymbol{H}[q]\boldsymbol{x}[n] + \boldsymbol{\eta}[n]$$

with the $K \times N$ channel matrix $\boldsymbol{H}[q]$ given by

$$\boldsymbol{H}[q] = [\boldsymbol{h}_1[q], \ldots, \boldsymbol{h}_K[q]]^{\mathrm{T}} \tag{2.21}$$

where $\boldsymbol{h}_k[q] \in \mathbb{C}^N$ is the channel vector for user k. $\boldsymbol{y}[n] = [y_1[n], \ldots, y_K[n]]^{\mathrm{T}} \in \mathbb{C}^K$ is the received vector and $\boldsymbol{\eta}[n] = [\eta_1[n], \ldots, \eta_K[n]]^{\mathrm{T}} \in \mathbb{C}^K$ is the noise vector with $f_{\boldsymbol{\eta}}(\boldsymbol{\eta}) = f_{\mathrm{G}}(\boldsymbol{\eta}, \mathbf{0}_K, \boldsymbol{C}_{\boldsymbol{\eta}})$.

2.6 Channel Model

We model the k-th user's channel vector $\boldsymbol{h}_k$ as a stationary zero-mean circularly symmetric (i.e. $\mathrm{diag}(\mathrm{e}^{\mathrm{j}\,\phi_1}, \ldots, \mathrm{e}^{\mathrm{j}\,\phi_N})$ has the same distribution as $\boldsymbol{h}_k$ for all ϕ_i) and complex Gaussian random vector with covariance matrix $\boldsymbol{C}_{\boldsymbol{h},k}$, i.e.

$$f_{\boldsymbol{h}_k}(\boldsymbol{h}_k) = f_{\mathrm{G}}\left(\boldsymbol{h}_k, \mathbf{0}_N, \boldsymbol{C}_{\boldsymbol{h},k}\right). \tag{2.22}$$

We assume that the channels corresponding to the different users are statistically independent.

In the q-th time slot, our model for the k-th user's channel vector is

$$\boldsymbol{h}_k[q] = \boldsymbol{C}_{\boldsymbol{h},k}^{1/2} \boldsymbol{h}_{\mathrm{w},k}[q] \tag{2.23}$$

with $\boldsymbol{h}_{\mathrm{w},k}[q]$ being a vector of independent stationary circularly symmetric complex white Gaussian processes (with unit variance) and where $(\bullet)^{1/2}$ represents the Cholesky decomposition. According to the modified Jakes model [33, 34] described in [35], temporal channel correlations are modeled by $\boldsymbol{h}_{\mathrm{w},k}[q]$, i.e.

$$\boldsymbol{C}_{\boldsymbol{h}_{\mathrm{w},k}}[D] = E[\boldsymbol{h}_{\mathrm{w},k}[q]\boldsymbol{h}_{\mathrm{w},k}^{\mathrm{H}}[q-D]] = J_0\left(2\pi \frac{f_{\mathrm{D,max},k}}{f_{\mathrm{slot}}} D\right) \mathbf{I}_N. \tag{2.24}$$

Here, the time scale of channel variations is in slot duration with D being the number of delay slots, J_0 denotes the zero–th order Bessel function of the first kind, $f_{\mathrm{D,max},k}$ is the *maximum* Doppler frequency [see Eq. (2.12)], and f_{slot} is the slot rate. The spatial correlations are introduced by the multiplication by $\boldsymbol{C}_{\boldsymbol{h},k}^{1/2}$.

Notice that, according to our model, the channel $\boldsymbol{h}_k$ is stationary because $\boldsymbol{h}_{\mathrm{w},k}$ is stationary. Realistic channels are usually non-stationary, i.e. either the location of the receiver or the scenario geometry can change. Thus, the channel covariance matrix has to be tracked in real situations. However, since the covariance matrix changes very slowly in comparison with the channel itself, it is realistic to assume that it is constant and perfectly known at both the receiver and the transmitter. Nevertheless, the feedback rate is limited and the feedback of the channel realizations for the precoder design must thus be optimized.

2.6.1 Spatial Channel Correlations

The development of more realistic channel models is of great interest to predict the performance of a wireless system, in particular to test the limited feedback designs proposed in this work.

It is important to stress here that for single-sensor narrowband receivers we can consider only the received signal power and/or time-varying amplitude (fading) distribution of the channel to acceptably approximate the channel variations. To this end, we use the *Third Generation Partnership Project's Spatial Channel Model* (3GPP-SCM) [36, 37], which is briefly described in Appendix A. Thus, the covariance matrix $\boldsymbol{C}_{\boldsymbol{h},k}$ in Eq. (2.23) results from considering 3GPP-SCM. This spatial channel model defines a stochastic channel model for MIMO systems. Although the description is for a downlink system where the Base Station (BS) transmits to several Mobile Stations (MS),

which reproduces exactly our assumption of a multiuser system, most of the aspects may also be applied to the uplink.

The SCM is also called *geometric* or *ray-based model* because it is based on stochastic modeling of scatterers. It defines three environments: *suburban macrocell* (approximately 3 Km. from BS to MS); *urban macrocell* (approximately 3 Km. from BS to MS); and *urban microcell* (less than 1 Km. from BS to MS). We refer to these environments as *SCM 1*, *SCM 2*, and *SCM 3*, respectively. These channels will be used in all the simulations shown throughout this work. The main spatial parameters related to each scenario (e.g. delay spread, angles of departure and arrival, average power) are defined in the 3GPP standard, and are shown in Table A.3 for SCM 1, Table A.4 for SCM 2, Table A.5 for SCM 3 and Table A.2 for all of them.

The procedure to generate the channel covariance matrices for each user according to the SCM is as follows. First, we specify the environment, i.e. we have to choose between suburban macro, urban macro, or urban microcell scenarios. After that, we obtain the corresponding parameters according to Tables A.2, A.3, A.4, and A.5. Finally, the channel coefficients $\boldsymbol{h}_{\text{SCM},k}[q]$ are generated based on the parameters and, as a result, the spatial correlations for each user given by its covariance matrix $\boldsymbol{C}_{\boldsymbol{h},k}$ are obtained as $\boldsymbol{C}_{\boldsymbol{h},k} = \text{E}[\boldsymbol{h}_{\text{SCM},k}[q]\boldsymbol{h}^{\text{H}}_{\text{SCM},k}[q]]$ [cf. Appendix A].

2.6.2 Temporal Channel Correlations

Additionally to the spatial correlations modeled by SCM, the channel also has temporal correlations modeled as described in [35]. This model is based on the sum of sinusoids of the Jakes model [33], which leads to the classical U-shape for the Doppler power spectrum (see Fig. 2.1) corresponding to spherically distributed scattering around the terminals. The detailed simulation model is described as follows [35].

Let $\boldsymbol{h}_k[q]$ be the complex channel realization for user k in the time slot q, whose i–th component is given by [see Eq. (2.20)]

$$h_{k,i}[q] = \frac{1}{\sqrt{2}} \left(h_{k,i,\text{R}}[q] + \text{j}\, h_{k,i,\text{I}}[q] \right).$$

Both real and imaginary parts are generated as

$$h_{k,i,\text{R}}[q] = \frac{2}{\sqrt{S}} \sum_{s=1}^{S} \cos\left(\psi_s\right) \cdot \cos\left(2\pi i v \cos\left(\alpha_s\right) + \phi_s\right)$$

$$h_{k,i,\text{I}}[q] = \frac{2}{\sqrt{S}} \sum_{s=1}^{S} \sin\left(\psi_s\right) \cdot \cos\left(2\pi i v \cos\left(\alpha_s\right) + \phi_s\right)$$

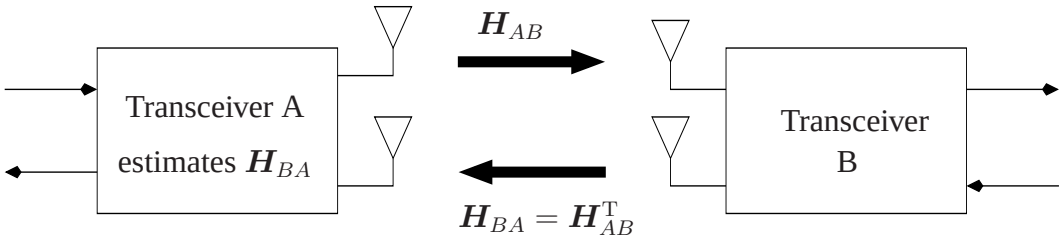


Figure 2.6: Obtaining CSIT using Reciprocity.

with

$$\alpha_s = \frac{2\pi s - \pi + \phi}{4S} \quad \text{for } s \in 1, \ldots, S$$

where ϕ, ϕ_s, and ψ_s are independent and uniformly distributed over $[-\pi, \pi]$ for all s. For the numerical simulations, the number of interfering paths is fixed at $S = 20$ [36]. Because of the central limit theorem when $S \to \infty$ and the independence of all ψ_s and ϕ_s, the real and imaginary channel parts are normally distributed, which ensures that modulus of $h_{k,i}[q]$ approximately follows a Rayleigh distribution (cf. Subsection 2.3.2) for all velocities v, even for $v = 0$.

2.7 Channel Estimation in FDD and TDD Systems

It is clear that the transmitter can only acquire the CSI indirectly, since the signal goes into the channel only after leaving the transmitter [32]. Therefore, the CSI can be obtained either by using the *reciprocity principle* or by using *feedback* from the receiver.

The reciprocity of the wireless channel implies that the channel from antenna A to antenna B can be estimated during the transmission in the opposite direction (B to A) since it is identical to the transpose of the channel from B to A (e.g. [17, 18]) as shown in Fig. 2.6. Pilot symbols are often used for channel estimation. The reciprocity holds if both forward and reverse links are located at the same frequency, the same time, and the same antenna locations. In practical systems, however, the forward and reverse links cannot use identical frequency, time, and spatial locations. In spite of that, the reciprocity principle can still hold approximately in some situations. For example, in the temporal dimension, the reciprocity principle is held if any time lag Δ_t between the forward and reverse transmission is much smaller than the channel coherence time T_c. Similarly, in the frequency dimension, any frequency offset Δ_f must be much smaller than the channel coherence bandwidth B_c, and in the spatial dimension the antenna location differences on the two links must be much smaller than the channel coherence distance D_c [28].

Since most communication systems are bi-directional, the uplink and downlink channels must be separated into orthogonal signaling dimensions. This separation is

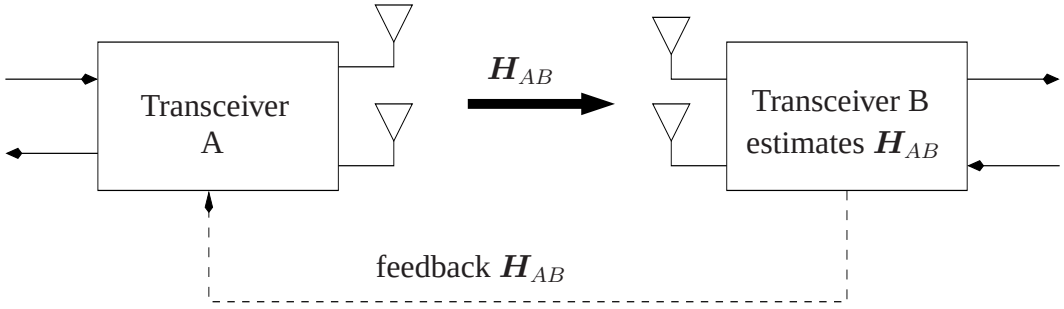


Figure 2.7: Obtaining CSIT using Feedback.

called *duplexing*.

Practical channel acquisition based on reciprocity may be applicable in TDD-TDMA (*Time-Division Duplex-Time-Division Multiple Access*) systems [27, 32, 38, 39]. TDMA consists of dividing the frame duration T_{f} into T non-overlapping subintervals, each of duration T_{f}/T. Each user who wants to transmit has to use a particular subinterval within each frame. In TDD systems, orthogonal time slots are assigned to each user to transmit to the base station and to receive from the base station. While TDD-TDMA systems have identical forward and reverse frequency bands and antennas, there is a time lag between the forward and reverse links. As mentioned above, such time lags must be negligible compared to the channel coherence time. Even in this case, reciprocity is difficult to accomplish due to the need for very good calibration (e.g. [24]).

In *Frequency-Division Multiple Access* (FDMA) systems (commonly used to accommodate multiple users for voice and data), the available channel bandwidth is split into a number of F frequency non-overlapping subchannels. Each subchannel is assigned to a user on demand. With *Frequency-Division Duplex* (FDD), separate frequency bands are assigned to each user for transmitting to or receiving from the base station. Therefore, FDD-FDMA systems often have identical temporal and spatial channel dimensions, but the frequency offset between the forward and reverse links is usually much larger than the channel coherence bandwidth. Therefore, reciprocity is usually not applicable in FDD systems. Instead, a feedback channel should be used to send the *Channel State Information* (CSI) from the transmitter to the receiver, as illustrated in Fig. 2.7. The channel response is estimated at the receiver B during the forward link (A to B) transmission, and the information is sent to the transmitter A on the reverse-link.

The same is true in a multiuser system. The transmitter is unable to obtain the CSI during reception in FDD systems because the channels are not reciprocal. This information must be sent from the users to the transmitter by means of a feedback or reverse channel, as plotted in Fig. 2.8 for a multiuser MISO system. Such reverse channels are actually implemented in most of the standards [40–42]. In this case, calibration errors

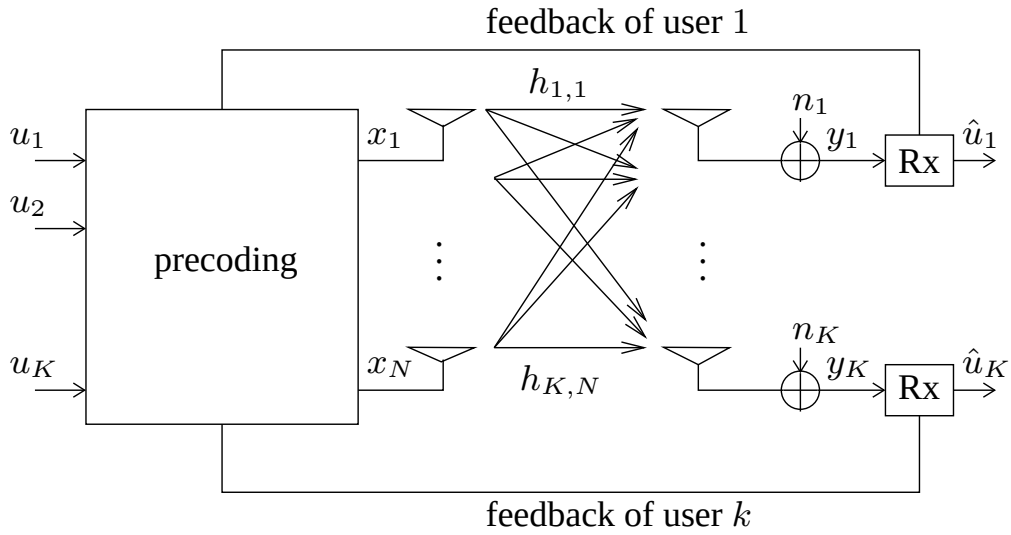


Figure 2.8: Multi-user MISO System with CSI Feedback and Precoding over Flat MISO Channels.

are estimated as part of the CSI and no special problems arise from calibration as for TDD. However, the time lag, D, between the channel measurement at the receivers and its use at the transmitter is a source of error (which will be modeled in this work by means of the feedback delay error) unless it is much smaller than the channel coherence time.

Moreover, the data rate of the feedback channel is highly limited. One drawback of feedback is the possible overhead of the reverse channel and the increasing consumption of transmit resources. Therefore, methods of reducing feedback overhead in a simple way, such as quantization or truncation of the feedback information, are crucial for practical implementations. As a consequence of the quantization, any system with limited rate CSI feedback suffers from erroneous CSI at the transmitter. Thus, the quantization operation has to be carefully designed, as done in this work.

Feedback can also be used to send channel statistics that change very slowly compared to the channel itself. In [17, 43, 44], the estimation of the statistics of the channel is discussed. As the time horizon for estimating the statistics is very large, we assume error-free knowledge of the statistics of the channel. Additionally, we assume that the channel statistics are constant and known at both the transmitter and receiver side. Nevertheless, the time lag requirement for feeding back the channel statistics is not as strong as for the feedback of the channel coefficients.

2.8 Conclusions

The goal of this chapter is the description of the radio propagation environment that exists in wireless communication systems. The main characteristics of a radio channel have been examined: mean path loss, macroscopic and microscopic fading, and signal spreading multipath effects. This analysis provides a channel model valid in general wireless environments. The 3GGP Spatial Channel Model is used to describe its spatial characteristics and time variations are modeled according to the Jakes model, so the resulting channel can be expressed as a linear and time-variant system.

Chapter 3

Multiuser MISO Transmit and SIMO Receive Processing with Perfect CSI

The main task when transmitting over channels with multiple antennas at the transmitter and/or the receiver side is the separation or equalization of the transmitted data. The *joint optimization* of transmit and receive filters was first proposed by [45] in 1952 and was widely studied in the past [46, 47]. However, this approach bears little relation to the goal of this work, since we focus on simplifying one side of the link in order to avoid filter operations at both the transmitter and the receiver side. As can be seen in [48], receive and transmit processing are outperformed by the respective jointly optimized approaches, since both receive and transmit approaches are suboptimum solutions obtained from the additional restriction that one filter is scalar. This scalar degree of freedom can be used to fulfill the transmit energy constraint and allows for closed–form solution, as has been demonstrated in [48]. Although many authors have dealt with transmit filters without this transmit energy constraint, such a constraint is necessary to avoid the dependence of the resulting transmit energy on the channel realization. So, the transmit energy constraint might be above the maximum value for bad channel realizations and thus the respective precoder solution is not valid. The transmitter may also not use the whole available transmit energy, and therefore the final quality is not as good as possible, since it could be improved by using more transmit energy. For receive processing, this constraint is also introduced to ensure the maximum transmit energy. In this case, we can make comparisons between the dual transmit and receive processing problems. Therefore, by restricting the transmit filter to being scalar we obtain the optimization for receive processing and by restricting the receive filter to being scalar the optimization for transmit processing is derived. These restrictions lead to useful schemes for the uplink or the downlink of wireless communications systems, respectively.

The goal of *receive processing* is to eliminate the distortion introduced by the channel at the receiver. The complexity of receive processing is located at the base station for the

uplink of wireless communications systems. For the downlink, however, this complexity is located at the users. It is known that the capacity in single–user MISO channels increases logarithmically with the number of antennas. In multi–user MISO systems as the considered downlink, capacity grows linearly with the number of users as long as the number of antennas is higher than the number of users. However, a single–antenna receiver is unable to separate the transmitted signals due to the lack of degrees of freedom, and also faces the problem that the requirements for the user devices become higher and infeasible. In other words, when we have non-cooperative receivers, as in the downlink of cellular communications systems, the users cannot cooperatively transform the received signals. Therefore, transmit filters are necessary to separate signals from different users before transmission through the fading channel. For all these reasons, neither joint optimization of transmit and receive filters nor receive approaches are applicable or recommendable for the downlink of multiuser MISO systems, which is the focus of this work. Thus, in many practical situations, the distortion introduced by the channel has to be compensated in advance at the transmitter instead of at the receiver as in classical single–user communications. *Transmit processing*, also termed *precoding*, is a powerful technique to reduce the tasks traditionally performed at the receiver side.

The objective of this chapter is to review most of the schemes commonly employed for transmit and receive processing. We summarize previous work as a starting point for the new contributions shown later in this thesis. We assume that the exact instantaneous channel information is known at both the transmitter and the receiver side. Therefore, channel estimation is not implemented at the receiver, whilst at the transmitter there is no need to consider the existence of a feedback channel to obtain the CSI from the different users. Although obtaining the instantaneous CSI for receive processing is relatively easy via estimation by transmitting known pilot symbols together with the unknown data, for transmit processing the major difficulty is the availability of instantaneous CSI at the transmitter, and the focus of this work is to determine optimal feedback information to be sent from the users to the transmitter [25]. Chapter 5 is exclusively dedicated to transmit processing approaches that are robust against erroneous CSI. The design of limited feedback multiuser systems is not a trivial problem in a multiuser MISO system, since the different users work in a decentralized way. This will be studied in Chapters 6 and 7.

In this chapter, we cover both linear and nonlinear systems in order to compare different schemes. We always include a constraint for the total transmit energy, since only such a formulation ensures valid solutions. We start with an analysis of different schemes for linear transmit and receive processing over MU-MISO and MU–SIMO channels, respectively. For the receive filters, we identify three filter types: *Matched Filter* (MF) [49–53], *Zero-Forcing Filter* (ZF) [52, 54], and *Wiener Filter* (WF) [55–57]. These three fundamental filter types were also found for transmit processing: *Matched*

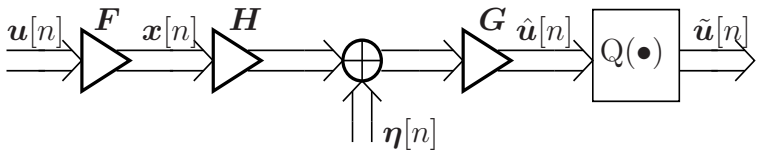


Figure 3.1: System with Linear Transmit and Receive Filters.

Filter (MF) [20, 58–60], *Zero-Forcing Filter* (ZF) [20, 48, 60, 61], and *Wiener Filter* (WF) [19, 48, 60, 62, 63]. Regarding nonlinear filters, we next focus our attention on the traditional *Decision Feedback Equalizer* (DFE) originally proposed by Austin [64], which is based on feeding back decisions in order to eliminate the interference of the previously detected symbols. DFE is a suboptimal approach to *Maximum-Likelihood Detection* (MLD), since the search over the possible data inherent to MLD is restricted so as to be successively computed. On the other hand, a search similar to that performed by MLD is done by *Vector Precoding* (VP) at the transmitter [65]. The non–linearity of VP is enabled by modulo operators introduced at the receivers. Again, when computed successively, the VP search gives us the suboptimal approach termed *Tomlinson-Harashima Precoding* (THP) [66, 67]. There also exists a close connection between DFE and THP, since the filters obtained for DFE are very similar to that of THP. The advantage of THP is that it avoids the error propagation due to the feedback of wrong decisions inherent to DFE, since for THP the fed–back signal depends exclusively on the data signal which is known to the transmitter. We focus on the standard approaches of MSE minimization with or without a *Zero-Forcing* (ZF) constraint together with a constraint of the total average transmit energy, given that these optimizations are based on the respective linear transmit processing optimizations.

3.1 MU-MISO Linear Transmit and MU–SIMO Linear Receive Processing

Fig. 3.1 shows the block diagram of a joint linear optimization scheme where the data signal $\boldsymbol{u}[n] \in \mathbb{C}^B$ is passed through the transmit filter $\boldsymbol{F} \in \mathbb{C}^{N \times B}$ to obtain the transmitted signal $\boldsymbol{x}[n] = \boldsymbol{F}\boldsymbol{u}[n] \in \mathbb{C}^N$. After propagation over the channel $\boldsymbol{H} \in \mathbb{C}^{K \times N}$ and the addition of the Gaussian noise $\boldsymbol{\eta}[n] \in \mathbb{C}^K$, the resulting signal is transformed by the receive filter $\boldsymbol{G} \in \mathbb{C}^{B \times K}$ to obtain the received signal $\hat{\boldsymbol{u}}[n]$ [20]:

$$\hat{\boldsymbol{u}}[n] = \boldsymbol{G}\boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n] + \boldsymbol{G}\boldsymbol{\eta}[n] \in \mathbb{C}^B. \tag{3.1}$$

Note that $\mathrm{Q}(\bullet)$ in Fig. 3.1 represents the quantizer operator that maps to the set of the transmitted symbols and delivers $\tilde{\boldsymbol{u}}[n]$.

Our objective is the joint optimization of the transmit and receive filter $\boldsymbol{F}$ and $\boldsymbol{G}$, respectively. The most widely used criteria for selecting $\boldsymbol{F}$ and $\boldsymbol{G}$ are the following:

- *Joint Wiener optimization*: based on the MSE minimization with only a transmit energy constraint, i.e.

$$\{\boldsymbol{F}_{\text{WF}}, \boldsymbol{G}_{\text{WF}}\} = \underset{\{\boldsymbol{F},\boldsymbol{G}\}}{\operatorname{argmin}} \operatorname{E}[\|\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2] \quad \text{s.t.: } \operatorname{E}[\|\boldsymbol{x}[n]\|_2^2] \leq E_{\text{tx}}. \tag{3.2}$$

- *Joint ZF optimization*: based on the MSE minimization together with a zero-forcing and a transmit energy constraint, i.e.

$$\begin{aligned}\{\boldsymbol{F}_{\text{ZF}}, \boldsymbol{G}_{\text{ZF}}\} = \underset{\{\boldsymbol{F},\boldsymbol{G}\}}{\operatorname{argmin}} \operatorname{E}\left[\|\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2\right] \\ \text{s.t.: } \boldsymbol{G}\boldsymbol{H}\boldsymbol{F} = \mathbf{I} \text{ and } \operatorname{E}[\|\boldsymbol{x}[n]\|_2^2] \leq E_{\text{tx}}.\end{aligned} \tag{3.3}$$

  Note that $\boldsymbol{G}\boldsymbol{H}\boldsymbol{F} = \mathbf{I}$ can only be fulfilled if $B \leq \min(K, N)$.

- *Eigenprecoder*: based on the *Signal-to-Noise Ratio* (SNR) maximization, i.e.

$$\{\boldsymbol{F}_{\text{MF}}, \boldsymbol{G}_{\text{MF}}\} = \underset{\{\boldsymbol{F},\boldsymbol{G}\}}{\operatorname{argmax}} \frac{\left|\operatorname{E}\left[\boldsymbol{u}^{\text{H}}[n]\hat{\boldsymbol{u}}[n]\right]\right|^2}{\operatorname{E}[\|\boldsymbol{u}[n]\|_2^2]\operatorname{E}[\|\boldsymbol{G}\boldsymbol{\eta}[n]\|_2^2]} \quad \text{s.t.: } \operatorname{E}[\|\boldsymbol{x}[n]\|_2^2] \leq E_{\text{tx}}. \tag{3.4}$$

As mentioned above, the restriction of either the transmit or the receive filter in Fig. 3.1 to being scalar leads to receive or transmit processing, respectively.

3.1.1 MU-SIMO Linear Receive Processing

As a result of restricting the transmit filter to being a weighted identity matrix, i.e. $\boldsymbol{F} = p\mathbf{I}$, the scheme depicted in Fig. 3.2 is obtained. As every scalar of the data signal is simply weighted with the scalar p and then applied to a transmit antenna, we conclude that $B = N$ for receive processing. The joint receive filter $\boldsymbol{G}$ implies that the receivers have to cooperate (which is called *centralized receivers*) to recover the transmitted signal. Such a setup can be found in the uplink of a cellular system, for example. However, as discussed before, for the downlink of a multiuser wireless system this assumption is not valid, and thus the channel equalization is performed at the transmitter instead of at the receiver side (see Subsection 3.1.2).

With the scalar transmit filter of Fig. 3.2, the channel is equalized only at the receiver side by means of the filter $\boldsymbol{G} \in \mathbb{C}^{N\times K}$ [49,50,52,60]. The estimated symbols are obtained as

$$\hat{\boldsymbol{u}}[n] = p\boldsymbol{G}\boldsymbol{H}\boldsymbol{u}[n] + \boldsymbol{G}\boldsymbol{\eta}[n] \in \mathbb{C}^N. \tag{3.5}$$

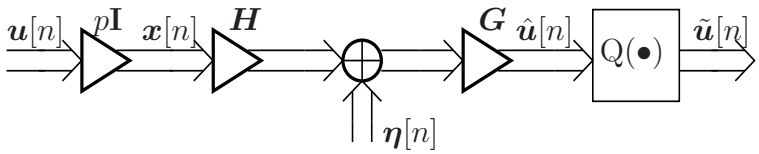


Figure 3.2: MU–SIMO System with Linear Receive Filter.

In the MU–SIMO setup, the channel matrix can be written as

$$\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N]$$

where $\boldsymbol{h}_i \in \mathbb{C}^K$ is the vector channel of the i–th user to the centralized receiver.

With the constraint $\boldsymbol{F} = p\mathbf{I}$ for the transmit filter, the optimizations for receive processing evolving from Eqs. (3.2), (3.3), and (3.4) are examined in the following sections.

MU-SIMO Receive Wiener Filter (RxWF)

The receive filter $\boldsymbol{G} \in \mathbb{C}^{N\times K}$ and the transmit weight $p \in \mathbb{C}$ are obtained by means of the following MSE minimization under a transmit energy constraint [20, 49, 55, 60], i.e.

$$\{p_{\text{WF}}, \boldsymbol{G}_{\text{WF}}\} = \underset{\{p,\boldsymbol{G}\}}{\operatorname{argmin}} \, \mathrm{E}[\|\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2] \quad \text{s.t.: } |p|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) \leq E_{\text{tx}}. \tag{3.6}$$

Taking into account that $\mathrm{E}[\|\boldsymbol{z}\|_2^2] = \mathrm{E}[\operatorname{tr}(\boldsymbol{z}\boldsymbol{z}^{\mathrm{H}})] = \operatorname{tr}(\mathrm{E}[\boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}])$, where $\boldsymbol{z}$ is a column vector, we construct the Lagrangian function in this way:

$$\begin{aligned} L(p, \boldsymbol{G}, \lambda) = {} & \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) - \operatorname{tr}\left(p^* \boldsymbol{C}_{\boldsymbol{u}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{G}^{\mathrm{H}}\right) - \operatorname{tr}\left(p \boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\right) + \operatorname{tr}\left(|p|^2 \boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) \\ & + \operatorname{tr}\left(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}}\right) + \lambda\left(|p|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) - E_{\text{tx}}\right) \end{aligned}$$

with the Lagrangian multiplier $\lambda \in \mathbb{R}^{0,+}$. The covariance matrices of the zero–mean transmit symbols and the zero–mean channel noise are given by $\boldsymbol{C}_{\boldsymbol{u}} = \mathrm{E}[\boldsymbol{u}[n]\boldsymbol{u}^{\mathrm{H}}[n]]$ and $\boldsymbol{C}_{\boldsymbol{\eta}} = \mathrm{E}[\boldsymbol{\eta}[n]\boldsymbol{\eta}^{\mathrm{H}}[n]]$, respectively.

Considering that $\operatorname{tr}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A}^{\mathrm{T}})$ (see Appendix B.2), we equate the derivatives with respect to p and $\boldsymbol{G}$ to zero (cf. Appendix C), which leads to the following KKT (Karush-Kuhn-Tucker) optimality conditions [68–71]:

$$\begin{aligned} \frac{\partial L(\bullet)}{\partial \boldsymbol{G}^*} &= -p^* \boldsymbol{C}_{\boldsymbol{u}} \boldsymbol{H}^{\mathrm{H}} + |p|^2 \boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}} = \mathbf{0} \\ \frac{\partial L(\bullet)}{\partial p} &= -\operatorname{tr}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}) + p^* \operatorname{tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right) + \lambda p^* \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) = 0 \\ |p|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) &\leq E_{\text{tx}} \\ \lambda\left(|p|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) - E_{\text{tx}}\right) &= 0 \quad \text{with } \lambda \geq 0. \end{aligned}$$

Note that these KKT conditions are only necessary conditions to find a global optimum solution since the MSE in Eq. (3.6) is not convex, as demonstrated in [48], and is therefore a nonconvex programming problem (see Appendix C).

From the first equation, we obtain the following expression for the receive filter $\boldsymbol{G}$:

$$\boldsymbol{G} = p^* \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} \left(|p|^2 \, \boldsymbol{H} \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C_\eta} \right)^{-1}.$$

By plugging this result into the second KKT condition, it is easy to demonstrate that $\lambda > 0$, and therefore the energy transmit constraint is maintained. To ensure a unique solution, we restrict $p \in \mathbb{R}^+$. Thus, p is obtained from the energy transmit constraint and we have that $p = \sqrt{\frac{E_{\mathrm{tx}}}{\mathrm{tr}(\boldsymbol{C_u})}}$. Then, the solution for the RxWF is as follows

$$\begin{aligned} \boldsymbol{G}_{\mathrm{WF}} &= p_{\mathrm{WF}} \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} \left(p_{\mathrm{WF}}^2 \boldsymbol{H} \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C_\eta} \right)^{-1} \\ p_{\mathrm{WF}} &= \sqrt{\frac{E_{\mathrm{tx}}}{\mathrm{tr}\left(\boldsymbol{C_u}\right)}}. \end{aligned} \tag{3.7}$$

Applying the matrix inversion lemma (see Appendix B.1) to the above expression for the receive filter $\boldsymbol{G}_{\mathrm{WF}}$, it can be demonstrated that

$$\begin{aligned} \boldsymbol{G}_{\mathrm{WF}} &= p_{\mathrm{WF}} \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} \left(\boldsymbol{C_\eta}^{-1} - \boldsymbol{C_\eta}^{-1} \boldsymbol{H} \left(p_{\mathrm{WF}}^{-2} \mathbf{I} + \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \boldsymbol{H} \right)^{-1} \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \right) \\ &= p_{\mathrm{WF}} \left(\boldsymbol{C_u} - \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \boldsymbol{H} \left(p_{\mathrm{WF}}^{-2} \mathbf{I} + \boldsymbol{C_u} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \boldsymbol{H} \right)^{-1} \boldsymbol{C_u} \right) \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \\ &= p_{\mathrm{WF}} \left(\boldsymbol{C_u}^{-1} + p_{\mathrm{WF}}^2 \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \boldsymbol{H} \right)^{-1} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \end{aligned}$$

and therefore Eq. (3.7) is rewritten as follows

$$\boxed{\begin{aligned} \boldsymbol{G}_{\mathrm{WF}} &= p_{\mathrm{WF}} \left(\boldsymbol{C_u}^{-1} + p_{\mathrm{WF}}^2 \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \boldsymbol{H} \right)^{-1} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C_\eta}^{-1} \\ p_{\mathrm{WF}} &= \sqrt{\frac{E_{\mathrm{tx}}}{\mathrm{tr}\left(\boldsymbol{C_u}\right)}}. \end{aligned}} \tag{3.8}$$

MU-SIMO Receive Zero-Forcing Filter (RxZF)

Receive zero-forcing processing is based on the MSE minimization with an additional zero-forcing constraint. Again, we have a transmit energy restriction, so the scalar weight $p \in \mathbb{R}$ and the receive filter $\boldsymbol{G} \in \mathbb{C}^{N \times K}$ should minimize the following expression [20, 49, 54, 60]:

$$\begin{aligned} \{p_{\mathrm{ZF}}, \boldsymbol{G}_{\mathrm{ZF}}\} = \underset{\{p, \boldsymbol{G}\}}{\mathrm{argmin}} \, & \mathrm{E}\left[\|\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2 \right] \\ & \text{s.t.: } p\boldsymbol{G}\boldsymbol{H} = \mathbf{I} \;\text{ and }\; |p|^2 \, \mathrm{tr}(\boldsymbol{C_u}) \leq E_{\mathrm{tx}}. \end{aligned} \tag{3.9}$$

Applying the zero-forcing constraint to the cost function, the MSE simplifies to the noise power at the receive filter output [54]:

$$\{p_{\text{ZF}}, \boldsymbol{G}_{\text{ZF}}\} = \underset{\{p,\boldsymbol{G}\}}{\operatorname{argmin}} \operatorname{tr}\left(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{G}^{\text{H}}\right) \quad \text{s.t.: } p\boldsymbol{G}\boldsymbol{H} = \mathbf{I} \text{ and } |p|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) \leq E_{\text{tx}}$$

and the Lagrangian function reads as

$$L\left(p, \boldsymbol{G}, \lambda\right) = \operatorname{tr}\left(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{G}^{\text{H}}\right) + 2\Re\left(\operatorname{tr}\left(\boldsymbol{\Lambda}\left(p\boldsymbol{G}\boldsymbol{H} - \mathbf{I}\right)\right)\right) + \lambda\left(|p|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) - E_{\text{tx}}\right)$$

with $\boldsymbol{\Lambda} \in \mathbb{C}^{N\times N}$ being the Lagrangian multiplier matrix and $\lambda \in \mathbb{R}^{0,+}$. The zero-forcing constraint is included in the Lagrangian function as $2\Re(\operatorname{tr}(\boldsymbol{\Lambda}(p\boldsymbol{G}\boldsymbol{H} - \mathbf{I})))$ since

$$\begin{aligned} 2\Re\left(\operatorname{tr}\left(\boldsymbol{\Lambda}\left(p\boldsymbol{G}\boldsymbol{H} - \mathbf{I}\right)\right)\right) &= \operatorname{tr}\left(\boldsymbol{\Lambda}\left(p\boldsymbol{G}\boldsymbol{H} - \mathbf{I}\right)\right) + \operatorname{tr}\left(\left(p^*\boldsymbol{H}^{\text{H}}\boldsymbol{G}^{\text{H}} - \mathbf{I}\right)\boldsymbol{\Lambda}^{\text{H}}\right) \\ &= 2\operatorname{tr}\left(\Re\left(\boldsymbol{\Lambda}\right)\Re\left(p\boldsymbol{G}\boldsymbol{H} - \mathbf{I}\right)\right) - 2\operatorname{tr}\left(\Im\left(\boldsymbol{\Lambda}\right)\Im\left(p\boldsymbol{G}\boldsymbol{H} - \mathbf{I}\right)\right) \end{aligned}$$

i.e. the complex-valued constraint is split into its real and imaginary part and each of the two real–valued constraints gets a real–valued Lagrangian multiplier.

When we set the derivatives with respect to p and $\boldsymbol{G}$ to zero, we obtain the following KKT conditions that are only necessary to find the solution since the zero-forcing constraint is bi-linear in the variables p and $\boldsymbol{G}$:

$$\begin{aligned} \frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{G}^*} &= \boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}} + p^*\boldsymbol{\Lambda}^{\text{H}}\boldsymbol{H}^{\text{H}} = \mathbf{0} \\ \frac{\partial L\left(\bullet\right)}{\partial p} &= \operatorname{tr}\left(\boldsymbol{\Lambda}\boldsymbol{G}\boldsymbol{H}\right) + \lambda p^* \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) = 0 \\ p\boldsymbol{G}\boldsymbol{H} &= \mathbf{I} \\ |p|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) &\leq E_{\text{tx}} \\ \lambda\left(|p|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) - E_{\text{tx}}\right) &= 0 \quad \text{with } \lambda \geq 0. \end{aligned}$$

With $\boldsymbol{G}$ obtained from the first KKT condition, and replacing it into the third equation, we obtain that

$$\boldsymbol{\Lambda} = -\left|p\right|^{-2}\left(\boldsymbol{H}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\right)^{-1}. \tag{3.10}$$

From this result for $\boldsymbol{\Lambda}$, we can infer that the receive zero-forcing filter only exists if $\boldsymbol{H}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}$ is invertible. Therefore, a necessary condition for the existence of the RxZF $\boldsymbol{G}_{\text{ZF}}$ is that $K \geq N$. Plugging the above expression for $\boldsymbol{\Lambda}$ into the first KKT condition leads to

$$\boldsymbol{G}_{\text{ZF}} = \frac{1}{p}\left(\boldsymbol{H}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}. \tag{3.11}$$

It is easy to see that when both expressions for $\boldsymbol{G}$ and $\boldsymbol{\Lambda}$ in Eqs. (3.10) and (3.11), respectively, are substituted into the second condition, it is obtained that $\lambda > 0$ and thus the transmit energy constraint is active.

Restricting p to being positive real, a unique solution is ensured, and the weight p is directly obtained from the transmit energy constraint. Then, we obtain that the RxZF solution to Eq. (3.9) is given by

$$\boxed{\begin{aligned}\boldsymbol{G}_{\text{ZF}} &= p_{\text{ZF}}^{-1}(\boldsymbol{H}^H\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^H\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\\ p_{\text{ZF}} &= \sqrt{\frac{E_{\text{tx}}}{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}})}}.\end{aligned}} \tag{3.12}$$

It is easy to see that when $\frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\text{tx}}} \to 0$, i.e, SNR $\to \infty$, the RxWF in Eq. (3.8) converges to the RxZF in Eq. (3.12).

MU-SIMO Receive Matched Filter (RxMF)

The receive matched filter is also known in the CDMA literature as the *rake* or conventional receiver [31]. To obtain the receive matched filter we use the eigenprecoder criterion in Eq. (3.4) to derive the optimization for the receive matched filter [20, 49, 60], i.e.

$$\begin{aligned}\{p_{\text{MF}}, \boldsymbol{G}_{\text{MF}}\} = \underset{p,\boldsymbol{G}}{\operatorname{argmax}} \frac{|\operatorname{tr}(p\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}|^2}{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}})\operatorname{tr}(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{G}^{\text{H}})}\\ \text{s.t.: } |p|^2\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) \le E_{\text{tx}}.\end{aligned} \tag{3.13}$$

We can form the Lagrangian function as follows,

$$L(p, \boldsymbol{G}, \lambda) = \frac{|\operatorname{tr}(p\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}})|^2}{\operatorname{tr}(\boldsymbol{C}_u)\operatorname{tr}(\boldsymbol{G}\boldsymbol{C}_{\eta}\boldsymbol{G}^{\text{H}})} + \lambda\left(|p|^2\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) - E_{\text{tx}}\right) \tag{3.14}$$

with $\lambda \in \mathbb{R}^{0,-}$. We set the derivatives with respect to p and $\boldsymbol{G}$ to zero, which yields the following KKT optimality conditions:

$$\begin{aligned}\frac{\partial L(\bullet)}{\partial \boldsymbol{G}^*} &= \frac{\operatorname{tr}(p\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}})\, p^*\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\text{H}}}{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}})\operatorname{tr}(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{G}^{\text{H}})} - \frac{|\operatorname{tr}(p\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}})|^2\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}}{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}})\operatorname{tr}^2(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{G}^{\text{H}})} = \boldsymbol{0}\\ \frac{\partial L(\bullet)}{\partial p} &= \frac{p^*|\operatorname{tr}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}})|^2}{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}})\operatorname{tr}(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{G}^{\text{H}})} + \lambda p^*\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) = 0\\ |p|^2\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) &\le E_{\text{tx}}\\ \lambda\left(|p|^2\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{u}}) - E_{\text{tx}}\right) &= 0 \quad \text{with } \lambda \le 0\end{aligned} \tag{3.15}$$

which are only necessary conditions since we maximize a non-concave function.

According to the second KKT condition, $\lambda < 0$, since $\boldsymbol{C}_{\boldsymbol{u}}$ and $\boldsymbol{C}_{\boldsymbol{\eta}}$ are positive definite. Then, the energy transmit constraint is satisfied with equality [see the last condition in Eq. (3.15)]. From the first KKT equation, we have

$$\boldsymbol{G} = \alpha\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\text{H}}{\boldsymbol{C}_{\boldsymbol{\eta}}}^{-1} \quad \text{with } \alpha \in \mathbb{C}$$

and, therefore, the solution is not unique. We set $\alpha = \sqrt{E_{\text{tx}}/\operatorname{tr}(\boldsymbol{C_u})}$ and $p \in \mathbb{R}^+$. Thus, the resulting MF solution is expressed as

$$\boxed{\begin{aligned} \boldsymbol{G}_{\text{MF}} &= p_{\text{MF}} \boldsymbol{C_u} \boldsymbol{H}^{\text{H}} \boldsymbol{C_\eta^{-1}} \\ p_{\text{MF}} &= \sqrt{\frac{E_{\text{tx}}}{\operatorname{tr}(\boldsymbol{C_u})}}. \end{aligned}} \tag{3.16}$$

It is easy to see that for low SNR (i.e. $\frac{\operatorname{tr}(\boldsymbol{C_\eta})}{E_{\text{tx}}} \to \infty$), the RxWF in Eq. (3.8) converges to RxMF in Eq. (3.16).

3.1.2 MU-MISO Linear Transmit Processing

As mentioned above, the equalization task can be performed at the transmitter, so the channel is pre-equalized or *precoded* before transmission with the goal of simplifying the user requirements. Such an operation prior to transmission is only possible for a centralized transmitter as in the downlink of a cellular system for example. In this subsection, we assume that the receive filter is an identity matrix (multiplied by a scalar g, with $g \in \mathbb{C}$) allowing for decentralized receivers. The goal is to find the optimum transmit filter $\boldsymbol{F}$. Therefore, the transmit and receive filter are given by the matrices $\boldsymbol{F} \in \mathbb{C}^{N \times K}$ and $\boldsymbol{G} = g\boldsymbol{I} \in \mathbb{C}^{K \times K}$, respectively. In other words, the number of scalar data streams is $B = K$. The resulting communications system is shown in Fig. 3.3. It can be seen from the figure how the data symbols $\boldsymbol{u}[n]$ are passed through the transmit filter $\boldsymbol{F}$ to form the transmit signal $\boldsymbol{x}[n] = \boldsymbol{F}\boldsymbol{u}[n] \in \mathbb{C}^N$. Note that the constraint for the transmit energy must be fulfilled, i.e.

$$\mathrm{E}\left[\|\boldsymbol{x}[n]\|_2^2\right] = \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^{\text{H}}\right) \leq E_{\text{tx}}.$$

The received signal is given by

$$\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n] + \boldsymbol{\eta}[n] \in \mathbb{C}^K$$

where $\boldsymbol{H} \in \mathbb{C}^{K \times N}$ and $\boldsymbol{\eta}[n] \in \mathbb{C}^K$ is the *Additive White Gaussian Noise* (AWGN). In the MU-MISO setup, the channel can be written as

$$\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]^{\text{T}}$$

where $\boldsymbol{h}_i^{\text{T}} \in \mathbb{C}^{1 \times N}$ is the channel from the centralized transmitter to the i–th user. Therefore, the channel $\boldsymbol{H}$ must be equalized by the transmit filter $\boldsymbol{F}$ prior to transmission. After multiplying by the receive gain g, we get the estimated symbols

$$\hat{\boldsymbol{u}}[n] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n] + g\boldsymbol{\eta}[n] \in \mathbb{C}^K. \tag{3.17}$$

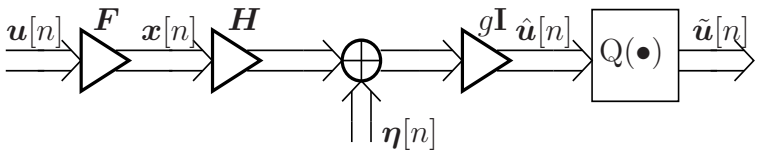


Figure 3.3: MU–MISO System with Linear Transmit Filter (Linear Precoding).

Similarly to receive processing, the optimizations for transmit processing are performed according to the three criteria described in Eqs. (3.2), (3.3), and (3.4) by restricting $\boldsymbol{G} = g\boldsymbol{I}$ as was done in [20, 60].

Clearly, the restriction that all the receivers apply the same scalar weight g is not necessary for decentralized receivers. Replacing $\boldsymbol{G}$ by a diagonal matrix suffices (e.g. [72]). However, usually no closed form can be obtained for the precoder if $\boldsymbol{G}$ is diagonal. Fortunately, $\boldsymbol{F}$ can be found in closed form for $\boldsymbol{G} = g\mathbf{I}$. Thus, we use $\boldsymbol{G} = g\mathbf{I}$ in the following.

MU-MISO Transmit Wiener Filter (TxWF)

Although Wiener filtering for precoding has been dealt with by only a few authors [63] in comparison with other criteria for precoding, it is a very powerful transmit optimization that minimizes the MSE with a transmit energy constraint [19, 48, 60, 62], i.e.

$$\{\boldsymbol{F}_{\text{WF}}, g_{\text{WF}}\} = \underset{\{\boldsymbol{F},g\}}{\operatorname{argmin}} \, \mathrm{E}\left[\|\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2\right] \quad \text{s.t.: } \operatorname{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) \leq E_{\text{tx}}. \tag{3.18}$$

We form the following Lagrangian function

$$\begin{aligned} L\left(\boldsymbol{F}, g, \lambda\right) = {} & \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) - \operatorname{tr}\left(g^*\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) - \operatorname{tr}\left(g\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right) + |g|^2 \operatorname{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) \\ & + |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) + \lambda\left(\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\text{tx}}\right). \end{aligned}$$

Setting the derivatives with respect to $\boldsymbol{F}$ and g to zero, and taking into account that the MSE in Eq. (3.18) is not convex, we obtain the necessary KKT conditions:

$$\begin{aligned} \frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{F}^*} &= -g^*\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}} + |g|^2 \boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}} + \lambda\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}} = \mathbf{0} \\ \frac{\partial L\left(\bullet\right)}{\partial g} &= -\operatorname{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right) + g^* \operatorname{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) + g^* \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) = 0 \\ \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) &\leq E_{\text{tx}} \\ \lambda\left(\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\text{tx}}\right) &= 0 \quad \text{with } \lambda \geq 0. \end{aligned} \tag{3.19}$$

The gain g^* obtained from the second equation is given by

$$g^* = \frac{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right)}{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)}. \tag{3.20}$$

Multiplying the first KKT condition by $\boldsymbol{F}^{\mathrm{H}}$ from the right and applying the trace operator, we get the following:

$$g^* \,\mathrm{tr}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) - |g|^2 \,\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) = \lambda \,\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right).$$

And now, combining this result with the expression for g^* in Eq. (3.20) yields

$$\begin{aligned}\lambda \,\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) &= \frac{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right)}{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)} \,\mathrm{tr}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) \\ &\quad - \frac{\left|\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right)\right|^2}{\mathrm{tr}^2\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)} \,\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) \\ &= |g|^2 \,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\end{aligned} \tag{3.21}$$

since $\mathrm{tr}^*(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}) = \mathrm{tr}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}})$. From the above result, $\lambda = \frac{|g|^2 \,\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{\mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}})} > 0$ if the trivial solution $\boldsymbol{F} = \boldsymbol{0}$ is not allowed. Therefore, the transmit energy constraint is an equality, i.e. $\mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) = E_{\mathrm{tx}}$ and consequently, $\lambda = |g|^2 \xi$ where, for brevity, we have introduced the notation to be used in the sequel:

$$\xi = \frac{\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)}{E_{\mathrm{tx}}}. \tag{3.22}$$

If we plug this result for λ into the first KKT condition, we get

$$\boldsymbol{F} = \frac{1}{g}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi\mathbf{I}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}. \tag{3.23}$$

By considering the transmit energy constraint $\mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) = E_{\mathrm{tx}}$ and the above expression for $\boldsymbol{F}$, it is obtained that

$$|g|^2 = \frac{\mathrm{tr}\left(\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi\mathbf{I}\right)^{-2}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}\right)}{E_{\mathrm{tx}}}$$

which leads to a unique solution if we restrict g to being positive real. Then, if we consider $g \in \mathbb{R}^+$, the solution for the Wiener filter is given by

$$\boxed{\begin{aligned}\boldsymbol{F}_{\mathrm{WF}} &= g_{\mathrm{WF}}^{-1}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi\mathbf{I}\right)^{-1}\boldsymbol{H}^{\mathrm{H}} \\ g_{\mathrm{WF}} &= \sqrt{\frac{\mathrm{tr}\left(\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi\mathbf{I}\right)^{-2}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}\right)}{E_{\mathrm{tx}}}}.\end{aligned}} \tag{3.24}$$

MU-MISO Transmit Zero-Forcing Filter (TxZF)

The transmit zero-forcing filter eliminates global interference at the output of the receive filter, and is based on the following MSE minimization under a transmit energy constraint [20, 48, 60, 61],

$$\begin{aligned}\{\boldsymbol{F}_{\text{ZF}}, g_{\text{ZF}}\} = \underset{\{\boldsymbol{F},g\}}{\operatorname{argmin}} \, & \mathrm{E}\left[\|\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2\right] \\ \text{s.t.: } & g\boldsymbol{H}\boldsymbol{F} = \mathbf{I} \quad \text{and} \quad \operatorname{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) \leq E_{\text{tx}}\end{aligned} \tag{3.25}$$

where the MSE including that zero-forcing constraint is given by

$$\mathrm{E}\left[\|\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2 \,|\, g\boldsymbol{H}\boldsymbol{F} = \mathbf{I}\right] = |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right).$$

Then, we can construct the Lagrangian function as follows,

$$L\left(\boldsymbol{F}, g, \lambda\right) = |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) + 2\Re\left(\boldsymbol{\Lambda}\left(g\boldsymbol{H}\boldsymbol{F} - \mathbf{I}\right)\right) + \lambda\left(\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\text{tx}}\right)$$

with $\boldsymbol{\Lambda} \in \mathbb{C}^{K\times K}$ and $\lambda \in \mathbb{R}^{0,+}$. This function enables us to obtain the following KKT conditions:

$$\begin{aligned}\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{F}^*} &= g^* \boldsymbol{H}^{\mathrm{H}} \boldsymbol{\Lambda}^{\mathrm{H}} + \lambda \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}} = \mathbf{0} \\ \frac{\partial L\left(\bullet\right)}{\partial g} &= g^* \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) + \operatorname{tr}\left(\boldsymbol{\Lambda}\boldsymbol{H}\boldsymbol{F}\right) = \mathbf{0} \\ g\boldsymbol{H}\boldsymbol{F} &= \mathbf{I} \\ \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) &\leq E_{\text{tx}} \\ \lambda\left(\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\text{tx}}\right) &= 0 \quad \text{with } \lambda \geq 0.\end{aligned}$$

Again, the above KKT conditions are only necessary to find the solution to Eq. (3.25) because the zero-forcing constraint is bilinear in g and $\boldsymbol{F}$.

By multiplying the first KKT condition by $\boldsymbol{F}^{\mathrm{H}}$ from the right and applying the trace operator, we get

$$\lambda \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) = |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)$$

where we have incorporated the equality $\operatorname{tr}(\boldsymbol{\Lambda}\boldsymbol{H}\boldsymbol{F}) = -g^* \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})$ obtained from the second KKT condition. Therefore, $\lambda > 0$ if $\boldsymbol{F} \neq \mathbf{0}$ and the transmit energy constraint is active with an equality.

From the first KKT condition it is obtained that the transmit filter $\boldsymbol{F}$ is

$$\boldsymbol{F} = -\frac{g^*}{\lambda} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{\Lambda}^{\mathrm{H}} {\boldsymbol{C}_{\boldsymbol{u}}}^{-1}. \tag{3.26}$$

Multiplying by $g\boldsymbol{H}$ from the left and applying the zero-forcing constraint yields

$$\boldsymbol{\Lambda}^{\mathrm{H}} = -\frac{\lambda}{|g|^2}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\boldsymbol{C}_{\boldsymbol{u}}.$$

Plugging $\boldsymbol{\Lambda}^{\mathrm{H}}$ into Eq. (3.26) leads to the following transmit filter

$$\boldsymbol{F} = \frac{1}{g}\boldsymbol{H}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}.$$

The weight g is derived from substituting the above expression for the transmit filter $\boldsymbol{F}$ into the transmit energy constraint, obtaining the following result

$$|g|^2 = \frac{\mathrm{tr}\left(\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\boldsymbol{C}_{\boldsymbol{u}}\right)}{E_{\mathrm{tx}}}$$

which leads to a unique solution if g is restricted to being positive real. Then, the solution for the TxZF is as follows

$$\boxed{\begin{aligned}\boldsymbol{F}_{\mathrm{ZF}} &= g_{\mathrm{ZF}}^{-1}\boldsymbol{H}^{H}(\boldsymbol{H}\boldsymbol{H}^{H})^{-1}\\ g_{\mathrm{ZF}} &= \sqrt{\frac{\mathrm{tr}((\boldsymbol{H}\boldsymbol{H}^{H})^{-1}\boldsymbol{C}_{\boldsymbol{u}})}{E_{\mathrm{tx}}}}.\end{aligned}} \tag{3.27}$$

By applying the matrix inversion lemma to the TxWF solution in Eq. (3.24), it is easy to demonstrate that the TxWF converges to the TxZF for $\xi = \frac{\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\mathrm{tx}}} \to 0$, i.e. for SNR $\to \infty$.

MU-MISO Transmit Matched Filter (TxMF)

The TxMF was intuitively introduced by Esmailzadeh et *al.* in [58] by moving the channel matched filter $\boldsymbol{H}^{\mathrm{H}}$ from the receiver to the transmitter. The transmit matched filter, also known as *prerake filter* [58, 59], maximizes the SNR and is obtained as follows [20, 60],

$$\begin{aligned}\{\boldsymbol{F}_{\mathrm{MF}}, g_{\mathrm{MF}}\} = \underset{\{\boldsymbol{F},g\}}{\mathrm{argmax}} \; & \frac{\left|\mathrm{E}\left[\boldsymbol{u}^{\mathrm{H}}[n]\hat{\boldsymbol{u}}[n]\right]\right|^2}{\mathrm{E}\left[\|\boldsymbol{u}[n]\|_2^2\right]\mathrm{E}\left[\|g\boldsymbol{\eta}[n]\|_2^2\right]}\\ \text{s.t.: } & \mathrm{E}\left[\|\boldsymbol{x}[n]\|\right] \le E_{\mathrm{tx}}\end{aligned} \tag{3.28}$$

it being advantageous for systems where the transmit energy or the SNR are low, since it is based on the maximization of desired signal portion in the received signal.

The above objective function can be rewritten as

$$\frac{\left|\mathrm{E}\left[\boldsymbol{u}[n]^{\mathrm{H}}\hat{\boldsymbol{u}}[n]\right]\right|^2}{\mathrm{E}\left[\|\boldsymbol{u}[n]\|_2^2\right]\mathrm{E}\left[\|g\boldsymbol{\eta}[n]\|_2^2\right]} = \frac{\left|\mathrm{tr}\left(g\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right)\right|^2}{\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right)\mathrm{tr}\left(|g|^2\boldsymbol{C}_{\boldsymbol{\eta}}\right)}.$$

Note that the above equation does not depend on g. Also the transmit energy constraint is independent from g. Therefore, the solution for the transmit matched filter is not unique.

The Lagrangian function is expressed as

$$L\left(\boldsymbol{F}, g, \lambda\right) = \frac{\left|\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C_u}\right)\right|^2}{\mathrm{tr}\left(\boldsymbol{C_u}\right)\mathrm{tr}\left(\boldsymbol{C_\eta}\right)} + \lambda\left(\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right)$$

which enables us to derive the following KKT conditions by setting the corresponding derivatives to zero:

$$\begin{aligned}
\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{F}^*} &= \frac{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C_u}\right)}{\mathrm{tr}\left(\boldsymbol{C_u}\right)\mathrm{tr}\left(\boldsymbol{C_\eta}\right)}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C_u} + \lambda\boldsymbol{F}\boldsymbol{C_u} = \boldsymbol{0} \\
\frac{\partial L\left(\bullet\right)}{\partial g} &= 0 \\
\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^{\mathrm{H}}\right) &\leq E_{\mathrm{tx}} \\
\lambda\left(\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right) &= 0 \quad \text{with } \lambda \leq 0.
\end{aligned} \tag{3.29}$$

Note that we maximize a non–concave objective function. Thus, the KKT conditions are not sufficient to find the solution (see Appendix C). After multiplying the first KKT condition by $\boldsymbol{F}^{\mathrm{H}}$ from the right and by rejecting the trivial solution $\boldsymbol{F} = \boldsymbol{0}$, the Lagrangian multiplier λ is given by

$$\lambda\,\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^{\mathrm{H}}\right) = -\frac{\left|\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C_u}\right)\right|^2}{\mathrm{tr}\left(\boldsymbol{C_u}\right)\mathrm{tr}\left(\boldsymbol{C_\eta}\right)}$$

which is smaller than zero, showing that the transmit energy constraint is active with an equality, i.e. $\mathrm{tr}(\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^{\mathrm{H}}) = E_{\mathrm{tx}}$. We also obtain from the first KKT condition that

$$\boldsymbol{F} = \alpha\boldsymbol{H}^{\mathrm{H}} \tag{3.30}$$

with $\alpha = -\frac{1}{\lambda}\frac{\mathrm{tr}(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C_u})}{\mathrm{tr}(\boldsymbol{C_u})\,\mathrm{tr}(\boldsymbol{C_\eta})} \in \mathbb{C}$. Plugging the above result into the transmit energy constraint yields:

$$|\alpha|^2 = \frac{E_{\mathrm{tx}}}{\mathrm{tr}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C_u}\boldsymbol{H}\right)}.$$

Therefore, the solution for the precoder $\boldsymbol{F}$ is not unique unless α is restricted to being positive real, for example. With this restriction, α is expressed as

$$\alpha = \sqrt{\frac{E_{\mathrm{tx}}}{\mathrm{tr}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C_u}\boldsymbol{H}\right)}}$$

and the solution for the TxMF is given by

$$\boxed{\begin{aligned}\boldsymbol{F}_{\mathrm{MF}} &= \sqrt{\frac{E_{tr}}{\mathrm{tr}(\boldsymbol{H}^{H}\boldsymbol{C_u}\boldsymbol{H})}}\boldsymbol{H}^{\mathrm{H}} \\ g_{\mathrm{MF}} &\in \mathbb{C}.\end{aligned}} \tag{3.31}$$

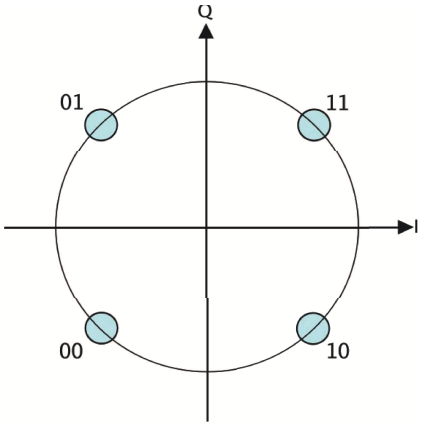


Figure 3.4: QPSK Constellation.

Note that the TxWF in Eq. (3.24) converges to the TxMF in Eq. (3.31) for low SNR scenarios, i.e, $\xi = \frac{\mathrm{tr}(\boldsymbol{C_\eta})}{E_{\mathrm{tx}}} \rightarrow \infty$.

3.1.3 Simulation Results

In this section, we show some computer simulations in order to illustrate the *Bit Error Rate* (BER) performance of the schemes discussed. The number of transmit antennas is equal to the number of users, i.e. $N = K = 4$, and the results are averaged over $5{,}000$ channel realizations. The information bits are *Quadrature Phase Shift Keying* (QPSK) modulated (Fig. 3.4). The modulation constellation is given as $\mathcal{A} = \{\pm\sqrt{2}/2 \pm \mathrm{j}\sqrt{2}/2\}$. A frame length of $N_{\mathrm{B}} = 50$ symbols is considered. We assume that $\boldsymbol{C_u} = \mathbf{I}$ and $\boldsymbol{C_\eta} = \sigma_\eta^2\mathbf{I}$, where σ_η^2 is the noise variance. We set the transmit energy to $E_{\mathrm{tx}} = N$. We use for the simulations the *SCM 2* described in Chapter 2 due to its intermediate BER performance and diversity.

Figs. 3.5 and 3.6 depict some results obtained from the comparison between the transmit and receive processing schemes described above. The results indicate that the performance of Wiener filters is always better compared to matched or zero-forcing filters. The performance of ZF schemes is worse than the corresponding MF designs for low SNR, but is better for high SNR, where the matched filters show a very poor performance. The same conclusions are obtained for receive processing.

It is apparent that the performance achieved with correlated channels, as plotted in Fig. 3.6, is worse than for uncorrelated channels, as depicted in Fig. 3.5. However, we can observe basically the same behavior as before for all the types of precoders. A slight difference can be seen between the transmit and receive processing due to the noise coloring at the receiver in receive processing. As this difference is small, we see that the application of the same filter type (e.g. RxZF in the uplink and TxZF in the downlink)

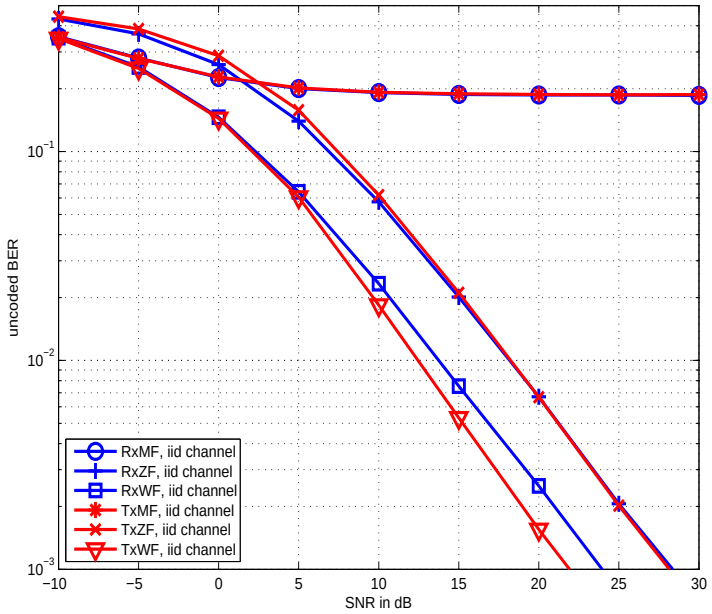


Figure 3.5: Uncoded BER vs. SNR for Linear Receive and Transmit Filters: QPSK Transmission over Uncorrelated Flat Fading MU-SIMO and MU-MISO Channels with Four Transmitting Antenna Elements and Four Users.

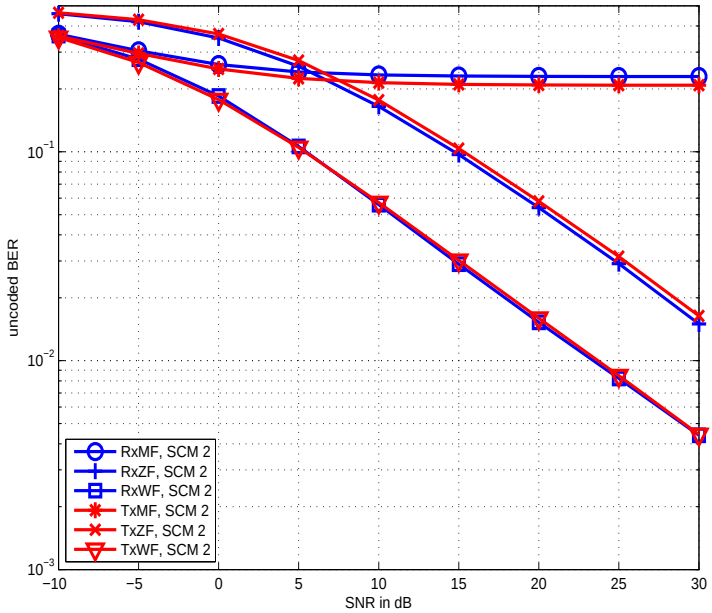


Figure 3.6: Uncoded BER vs. SNR for Linear Receive and Transmit Filters: QPSK Transmission over Correlated Flat Fading MU-SIMO and MU-MISO Channels (*SCM 2*) with Four Transmitting Antenna Elements and Four Users.

leads to similar results in the up- as in the downlink. So, there is no asymmetry of error performance.

3.2 MU-SIMO Nonlinear Receive Processing

In this section, we focus on various nonlinear systems with transmit or receive channel equalization with the goal of recovering the data at the receivers. Nonlinear receive processing requires cooperation between the receivers, this being known as *centralized* receivers. This is no limitation for the uplink of a wireless communications system, since the transmitters are located at the base station. However, we should recall that signal processing at the receiver side is quite useful for the uplink, but not for the downlink under study in this work. Moreover, transmit processing simplifies the requirements in the user devices, which implies an important reduction in terms of cost and complexity. The interest of studying these nonlinear schemes is to verify the performances obtained from both receive and transmit processing ignoring issues concerning their practical implementation.

It is known that *Maximum Likelihood Detection* (MLD) leads to full diversity and is the optimum detection scheme in the sense that it minimizes the probability of a symbol being erroneously detected. The search over all the possible data inherent to MLD can be seen as a lattice search and computed by *sphere decoding* [73–77]. However, its computational complexity is prohibitive in many cases because it grows exponentially and thus nonpolynomically. Contrary to MLD, suboptimum detection schemes such as the *Decision-Feedback Equalizer* (DFE) have been widely used in recent years. DFE, however, suffers from the major drawback of error propagation derived from feeding back erroneous decisions. This effect can be solved by performing the equalization similarly to DFE but at the transmitter side instead of the receiver side. This idea leads to *Tomlinson-Harashima Precoding* (THP), which, again, is a suboptimum approach of *Vector Precoding* (VP). Similarly to MLD, VP consists of a lattice search carried out at the transmitter instead of the receiver side. There is thus a double parallelism MLD vs. VP, and DFE vs. THP and, on the other hand, between MLD vs. DFE and VP vs. THP.

3.2.1 MU-SIMO Maximum Likelihood Detection (MLD)

MLD decides for the signal $\hat{\boldsymbol{u}}[n]$ that maximizes the likelihood of the received signal $\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{u}[n] + \boldsymbol{\eta}[n]$ [78, 79], i.e.

$$\begin{aligned}\hat{\boldsymbol{u}}_{\text{MLD}}[n] &= \underset{\boldsymbol{u}[n]\in\mathbb{A}^N}{\operatorname{argmax}} f_{\boldsymbol{y}}\left(\boldsymbol{y}[n];\boldsymbol{u}[n]\right) \\ &= \underset{\boldsymbol{u}[n]\in\mathbb{A}^N}{\operatorname{argmax}} f_{\boldsymbol{\eta}}\left(\boldsymbol{y}[n]-\boldsymbol{H}\boldsymbol{u}[n]\right) \\ &= \underset{\boldsymbol{u}[n]\in\mathbb{A}^N}{\operatorname{argmin}} \left(\boldsymbol{y}[n]-\boldsymbol{H}\boldsymbol{u}[n]\right)^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\left(\boldsymbol{y}[n]-\boldsymbol{H}\boldsymbol{u}[n]\right)\end{aligned} \tag{3.32}$$

where $\mathbb{A}$ denotes the alphabet of the data signal, that is, $\boldsymbol{u}[n] \in \mathbb{A}^N$. Assuming that the noise is spatially white [1], i.e. $\boldsymbol{C}_{\boldsymbol{\eta}} = \sigma_\eta^2\mathbf{I}$, and introducing the QR decomposition $\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{R}$, where $\boldsymbol{Q}$ is unitary and $\boldsymbol{R}$ is upper triangular, we can write Eq. (3.32) as

$$\begin{aligned}\hat{\boldsymbol{u}}_{\text{MLD}}[n] &= \underset{\boldsymbol{u}[n]}{\operatorname{argmin}} \|\boldsymbol{y}[n]-\boldsymbol{H}\boldsymbol{u}[n]\|_2^2 = \underset{\boldsymbol{u}[n]}{\operatorname{argmin}} \|\boldsymbol{y}[n]-\boldsymbol{Q}\boldsymbol{R}\boldsymbol{u}[n]\|_2^2 \\ &= \underset{\boldsymbol{u}[n]}{\operatorname{argmin}} \|\tilde{\boldsymbol{y}}[n]-\boldsymbol{R}\boldsymbol{u}[n]\|_2^2\end{aligned} \tag{3.33}$$

where $\tilde{\boldsymbol{y}}[n] = \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{y}[n]$. Therefore, the minimization for the case of a QAM constellation is a closest point search in a subset of an N-dimensional lattice. In spite of being the optimum detector for equiprobable data, MLD is often infeasible on account of its enormous complexity. *Sphere decoding* [73–77] performs this search in a more sophisticated manner than just doing a full search over the subset of a lattice, but still requires exponential complexity. In fact, sphere decoding only searches over the lattice points lying in a certain hypersphere of radius r centered on the received signal $\boldsymbol{y}[n]$. However, sorting out the points outside the sphere leads to the exponential worst case complexity.

Due to the upper triangular structure of $\boldsymbol{R}$, the i–th summand of the Euclidean norm $\|\tilde{\boldsymbol{y}}[n]-\boldsymbol{R}\boldsymbol{u}[n]\|_2^2 = \sum_{i=1}^N \lambda_i$ is

$$\lambda_i = \left|\tilde{y}_i[n] - r_{i,i}u_i[n] - \sum_{j=i+1}^{N} r_{i,j}u_j[n]\right|^2 \tag{3.34}$$

where $r_{i,j}$ corresponds to the element of the i-th row and j-th column of $\boldsymbol{R}$.

When this search is computed successively, i.e. $u_i[n]$ is found for fixed $u_{i+1}[n],\ldots,u_N[n]$, we meet the idea of the *Vertical Bell Labs Layered Space-Time* architecture (V-BLAST) which was based on this successive interference cancellation.

[1] Such a setup can be achieved by left–multiplying $\boldsymbol{y}[n]-\boldsymbol{H}\boldsymbol{y}[n]$ by $\boldsymbol{C}_{\boldsymbol{\eta}}^{-1/2}$.

The decoding algorithm presented in the first works about BLAST [80, 81] was based on interference nulling, interference cancellation, and ordering. Indeed, this decoding process is equivalent to the zero-forcing *Decision-Feedback Equalization* (ZF-DFE) [82, 83], where the interference nulling is performed by a so called *feedforward filter* and the interference cancellation by the *feedback filter*. The original BLAST ordering algorithm was based on an SNR criterion, although different ordering algorithms have been proposed since this first proposal. In fact, the DFE ordering studied throughout this work is based on an MSE criterion.

3.2.2 MU-SIMO Decision-Feedback Receiver (DFE)

The block diagram of a MU-MISO system employing DFE is depicted in Fig. 3.7. Given that we work with flat fading channels, there is no need to deal with the temporal decision-feedback equalizer [48], and therefore we henceforth only refer to spatial DFE. Again, note that the scheme depicted in Fig. 3.7 implies cooperation between the receivers. For the downlink, i.e. a multiuser MISO system, this cooperation between the users is infeasible, and therefore DFE is not a practical choice for separating the signals from the different users. This task, however, can be performed by the transmitter where a centralized base station makes this separation before transmission by means of *precoding* (see Section 3.3). Contrary to the downlink, DFE is really useful for the uplink since the base station can easily obtain the filters to be used by the DFE strategy.

The DF equalizer has been widely used in wireless communications systems to avoid the noise amplification problem in linear equalizers. Although initially proposed to equalize SISO communication links with IIR filters, DFE has been extended to multiuser MIMO channels, whilst IIR filters have been restricted to being FIR due to practical implementations. The DF equalizer uses feedback from past decisions to cancel the interference of the symbols that have already been detected. It consists of two linear filters: the feedforward filter, whose input is the received sequence, and the feedback filter, whose input is the previously detected sequence. The feedforward filter provides spatial causality and ensures that the error is white. The feedback filter, however, exploits causality for the feedback loop and ISI cancellations due to its strictly lower triangular structure [21, 84]. For achieving optimum performance, the symbols have to be detected according to a specific ordering. This issue has an enormous influence on the performance, as we will see from some computer simulations at the end of this section. However, the decision feedback receiver suffers from the major drawback of error propagation, which will be solved when the feedback and the feedforward filters are moved to the transmitter (to end up with THP, see Subsection 3.3.2).

As can be seen in Fig. 3.7, the received signals can be concisely expressed in matrix–

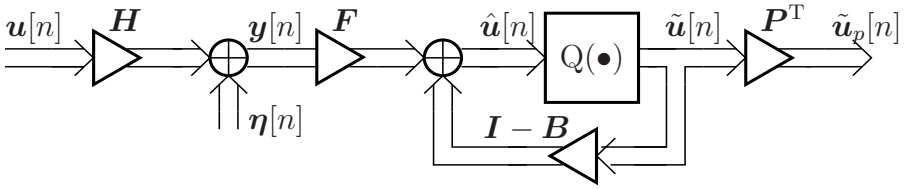


Figure 3.7: MU–SIMO System with DFE.

vector notation as

$$\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{u}[n] + \boldsymbol{\eta}[n] \in \mathbb{C}^K \tag{3.35}$$

where $\boldsymbol{u}[n] \in \mathbb{A}^N$ are the transmitted symbols and $\mathbb{A}$ denotes the modulation alphabet. $\boldsymbol{H} \in \mathbb{C}^{K\times N}$ is the flat fading channel introduced in Subsection 3.1.1 and $\boldsymbol{\eta}[n] \in \mathbb{C}^K$ is the received noise. We optimize the estimated signal $\hat{\boldsymbol{u}}[n]$ in Fig. 3.7, which can be expressed as

$$\hat{\boldsymbol{u}}[n] = \boldsymbol{F}\boldsymbol{y}[n] + (\mathbf{I} - \boldsymbol{B})\,\tilde{\boldsymbol{u}}[n] \tag{3.36}$$

with $\boldsymbol{y}[n]$ in Eq. (3.35) and $\tilde{\boldsymbol{u}}[n] \in \mathbb{A}^N$ denoting the quantized symbols.

We introduce the permutation matrix to be used in the sequel as follows

$$\boldsymbol{P} = \sum_{n=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_{k_i}^{\mathrm{T}} \in \{0,1\}^{N\times N} \tag{3.37}$$

where $\{k_1, \ldots, k_N\}$ with $k_i \in \{1, \ldots, N\} \setminus \{k_1, \ldots, k_{i-1}\}$ determine the detection ordering. The i-th column of the $N \times N$ identity matrix $\mathbf{I}_N$ is denoted by $\boldsymbol{e}_i \in \{0,1\}^N$. Thus, the recovered symbols $\tilde{\boldsymbol{u}}[n]$ are permuted by $\boldsymbol{P}^{\mathrm{T}}$ to get $\tilde{\boldsymbol{u}}_p[n]$. Given that $\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}} = \mathbf{I}$, we have that $\tilde{\boldsymbol{u}}[n] = \boldsymbol{P}\tilde{\boldsymbol{u}}_p[n]$. Then, Eq. (3.36) can be rewritten as

$$\hat{\boldsymbol{u}}[n] = \boldsymbol{F}\boldsymbol{y}[n] + (\mathbf{I} - \boldsymbol{B})\,\boldsymbol{P}\tilde{\boldsymbol{u}}_p[n].$$

MU–SIMO Wiener Decision Feedback Receiver (WF-DFE)

As the quantized symbols $\tilde{\boldsymbol{u}}[n]$ are reordered by $\boldsymbol{P}^{\mathrm{T}}$ to get the detected symbols $\tilde{\boldsymbol{u}}_p[n]$, the desired value for the estimates $\hat{\boldsymbol{u}}[n]$ is $\boldsymbol{P}\boldsymbol{u}[n]$. Assuming that decisions made prior to every detection are correct (i.e. $\tilde{\boldsymbol{u}}_p[n] = \boldsymbol{u}[n]$), we have the error vector defined as follows [48, 85],

$$\boldsymbol{\epsilon}_p[n] = \boldsymbol{P}\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n] = \boldsymbol{P}\boldsymbol{u}[n] - \boldsymbol{F}\boldsymbol{y}[n] - (\mathbf{I} - \boldsymbol{B})\,\boldsymbol{P}\boldsymbol{u}[n] = \boldsymbol{B}\boldsymbol{P}\boldsymbol{u}[n] - \boldsymbol{F}\boldsymbol{y}[n].$$

The WF-DFE feedforward and feedback filters are found by minimizing the MSE and restricting the feedback filter $\boldsymbol{B}$ to being lower triangular, i.e.

$$\left\{\boldsymbol{P}_{\mathrm{WF}}^{\mathrm{DFE}}, \boldsymbol{F}_{\mathrm{WF}}^{\mathrm{DFE}}, \boldsymbol{B}_{\mathrm{WF}}^{\mathrm{DFE}}\right\} = \mathrm{E}\left[\|\boldsymbol{P}\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2\right] \quad \text{s.t.: } \boldsymbol{B} \text{ is unit lower triangular} \tag{3.38}$$

where the MSE $\varepsilon_{\text{WF}}^{\text{DFE}}(\boldsymbol{P},\boldsymbol{B},\boldsymbol{F}) = \mathrm{E}[\|\tilde{\boldsymbol{u}}[n] - \hat{\boldsymbol{u}}[n]\|_2^2]$ is calculated as

$$\begin{aligned}\varepsilon_{\text{WF}}^{\text{DFE}}(\boldsymbol{P},\boldsymbol{B},\boldsymbol{F}) &= \mathrm{E}\left[\|\boldsymbol{P}\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]\|_2^2\right] = \mathrm{E}\left[\|\boldsymbol{B}\boldsymbol{P}\boldsymbol{u}[n] - \boldsymbol{F}\boldsymbol{H}\boldsymbol{u}[n] - \boldsymbol{F}\boldsymbol{\eta}[n]\|_2^2\right] \\ &= \mathrm{tr}\left(\boldsymbol{B}\boldsymbol{P}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{H}}\right) - \mathrm{tr}\left(\boldsymbol{B}\boldsymbol{P}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{F}^{\mathrm{H}}\right) - \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{H}}\right) \\ &\quad + \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{F}^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{F}^{\mathrm{H}}\right). \end{aligned} \tag{3.39}$$

This allows us to construct the Lagrangian function

$$\begin{aligned}L(\boldsymbol{P},\boldsymbol{F},\boldsymbol{B},\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_k) &= \mathrm{tr}\left(\boldsymbol{B}\boldsymbol{P}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{H}}\right) - \mathrm{tr}\left(\boldsymbol{B}\boldsymbol{P}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{F}^{\mathrm{H}}\right) \\ &\quad - \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{F}^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{F}^{\mathrm{H}}\right) \\ &\quad + 2\Re\left(\mathrm{tr}\left(\sum_{i=1}^{N}\left(\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{B}\boldsymbol{S}_i^{\mathrm{T}} - \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)\boldsymbol{\mu}_i\right)\right) \end{aligned} \tag{3.40}$$

where the equality $\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{B}\boldsymbol{S}_i^{\mathrm{T}} = \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}$, for $i = 1,\ldots,N$, must hold because of the unit lower triangular structure of $\boldsymbol{B}$ [2]. To mathematically formulate this restriction, we included the selection matrix $\boldsymbol{S}_i$ defined as

$$\boldsymbol{S}_i = [\mathbf{0}_{N-i+1\times i-1}, \mathbf{I}_{N-i+1}] \in \{0,1\}^{N-i+1\times N} \tag{3.41}$$

which cuts out the last $N-i+1$ rows of a matrix with N rows, when applied from the left. The Lagrangian multiplier $\boldsymbol{\mu}_i$, $i = 1,\ldots,N$, is a column vector of dimension $N-i+1$. Note that we need $2\Re(\bullet)$ in Eq. (3.40) to assure that the structural constraint is fulfilled for both the real and the imaginary part of $\boldsymbol{B}$.

By setting its derivatives with respect to $\boldsymbol{F}$ and $\boldsymbol{B}$ to zero, we obtain the following sufficient KKT conditions that lead to a unique global minimum with respect to $\boldsymbol{F}$ and $\boldsymbol{B}$ because the MSE in Eq. (3.38) is strictly convex (cf. [48] and Appendix C):

$$\begin{aligned}\frac{\partial L(\bullet)}{\partial \boldsymbol{F}^*} &= -\boldsymbol{B}\boldsymbol{P}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{F}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}} = \mathbf{0} \\ \frac{\partial L(\bullet)}{\partial \boldsymbol{B}^*} &= \boldsymbol{B}\boldsymbol{P}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{P}^{\mathrm{T}} - \boldsymbol{F}\boldsymbol{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{P}^{\mathrm{T}} + \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{\mu}_i^{\mathrm{H}}\boldsymbol{S}_i = \mathbf{0} \\ \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{B}\boldsymbol{S}_i^{\mathrm{T}} &= \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}} \quad \forall i \in \{1,\ldots,N\}. \end{aligned} \tag{3.42}$$

From the second KKT condition, we obtain

$$\boldsymbol{B} = \boldsymbol{F}\boldsymbol{H}\boldsymbol{P}^{\mathrm{T}} - \left(\sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{\mu}_i^{\mathrm{H}}\boldsymbol{S}_i\right)\boldsymbol{P}\boldsymbol{C}_{\boldsymbol{u}}^{-1}\boldsymbol{P}^{\mathrm{T}}. \tag{3.43}$$

[2] The lefthand side cuts out the last $N-i+1$ elements of the i–th row of $\boldsymbol{B}$ and the righthand side sets the first of those elements (the i–th diagonal element of $\boldsymbol{B}$) to one and the others to zero (triangularity of $\boldsymbol{B}$).

Plugging this expression for $\boldsymbol{B}$ into the first KKT condition, we get

$$\left(\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{\mathrm{H}} \boldsymbol{S}_i\right) \boldsymbol{P}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}} = \boldsymbol{0}$$

and therefore

$$\boldsymbol{F} = -\left(\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{\mathrm{H}} \boldsymbol{S}_i\right) \boldsymbol{P}\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1}. \tag{3.44}$$

Substituting into Eq. (3.43) we obtain

$$\boldsymbol{B} = -\left(\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{\mathrm{H}} \boldsymbol{S}_i\right) \boldsymbol{P} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}}. \tag{3.45}$$

Applying the restriction concerned with the unit lower triangular structure of $\boldsymbol{B}$ to the above result leads to

$$\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{B} \boldsymbol{S}_i^{\mathrm{T}} = -\boldsymbol{e}_i^{\mathrm{T}} \left(\sum_{j=1}^{N} \boldsymbol{e}_j \boldsymbol{\mu}_j^{\mathrm{H}} \boldsymbol{S}_j\right) \boldsymbol{P} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} = \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}}.$$

Then, with $\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}_j = 0$, for $j \neq i$, and 1, otherwise, $\boldsymbol{\mu}_i^{\mathrm{H}}$ reads as

$$\boldsymbol{\mu}_i^{\mathrm{H}} = -\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left[\boldsymbol{S}_i \boldsymbol{P} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}}\right]^{-1}.$$

This result for $\boldsymbol{\mu}_i^{\mathrm{H}}$ gives us the following expressions for the filters $\boldsymbol{F}$ and $\boldsymbol{B}$ of Fig. 3.7:

$$\begin{aligned}
\boldsymbol{F} &= \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left[\boldsymbol{S}_i \boldsymbol{P} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}}\right]^{-1} \boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\
\boldsymbol{B} &= \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left[\boldsymbol{S}_i \boldsymbol{P} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}}\right]^{-1} \boldsymbol{S}_i \boldsymbol{P} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right) \boldsymbol{P}^{\mathrm{T}}.
\end{aligned} \tag{3.46}$$

In order to simplify calculation, let us define $\boldsymbol{\Phi} = (\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1})^{-1}$. Since this matrix is Hermitian, there exists a permutation matrix $\boldsymbol{P}$, a unit lower triangular matrix $\boldsymbol{L}$, and a diagonal matrix $\boldsymbol{D}$, which satisfy the following relationship [85, 86]

$$\boldsymbol{P}\boldsymbol{\Phi}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{H}} \tag{3.47}$$

which will be termed *Cholesky factorization with symmetric permutation* [86]. Bearing in mind this factorization, the feedforward filter in Eq. (3.46) reduces to

$$\begin{aligned}\boldsymbol{F} &= \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left(\boldsymbol{S}_i \boldsymbol{L}^{-\mathrm{H}} \boldsymbol{D}^{-1} \boldsymbol{L}^{-1} \boldsymbol{S}_i^{\mathrm{T}}\right)^{-1} \boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ &= \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left(\boldsymbol{S}_i \boldsymbol{L}^{-\mathrm{H}} \boldsymbol{D}^{-1} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i \boldsymbol{L}^{-1} \boldsymbol{S}_i^{\mathrm{T}}\right)^{-1} \boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ &= \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i \boldsymbol{L} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ &= \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} = \sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ &= \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \end{aligned} \tag{3.48}$$

where in the derivations we have used the following properties for the selection matrix $\boldsymbol{S}_i$:

$$\boldsymbol{S}_i \boldsymbol{N} = \boldsymbol{S}_i \boldsymbol{N} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i, \quad \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i \boldsymbol{M} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i = \boldsymbol{e}_i^{\mathrm{T}}, \quad \text{and} \quad \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{N} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i = \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{N}$$

with $\boldsymbol{N}$ being an upper triangular matrix and $\boldsymbol{M}$ a unit lower triangular matrix. Comparing this result with Eq. (3.44) leads to the conclusion that $-\sum_{i=1}^{N} \boldsymbol{e}_i \boldsymbol{\mu}_i^{\mathrm{H}} \boldsymbol{S}_i = \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}}$. Hence, the feedback filter reduces to [cf. Eq. (3.45)]

$$\boldsymbol{B} = \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \boldsymbol{L}^{-\mathrm{H}} \boldsymbol{D}^{-1} \boldsymbol{L}^{-1} = \boldsymbol{L}^{-1}. \tag{3.49}$$

Therefore, the filters $\boldsymbol{B}$ and $\boldsymbol{F}$ in Fig. 3.7 corresponding to the WF-DFE solution are given by

$$\boxed{\begin{aligned} \boldsymbol{F}_{\mathrm{WF}}^{\mathrm{DFE}} &= \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ \boldsymbol{B}_{\mathrm{WF}}^{\mathrm{DFE}} &= \boldsymbol{L}^{-1}. \end{aligned}} \tag{3.50}$$

Finally, the MSE in Eq. (3.39) reads as

$$\varepsilon_{\mathrm{WF}}^{\mathrm{DFE}} = \mathrm{tr}\left(\boldsymbol{D}\right) = \sum_{k=1}^{N} d_i \tag{3.51}$$

where $\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} = \boldsymbol{P}^{\mathrm{T}} \boldsymbol{L}^{-\mathrm{H}} \boldsymbol{D}^{-1} \boldsymbol{L}^{-1} \boldsymbol{P} - \boldsymbol{C}_{\boldsymbol{u}}^{-1}$ was used. Solving Eq. (3.38) would imply that the $N!$ different factorizations in Eq. (3.47) corresponding to all possible permutations must be computed and that the permutation minimizing Eq. (3.51) must

be taken. As in [85], we avoid this very complex procedure and instead implement a successive computation to choose the optimum order, where the k_i-th entry is given by

$$k_i = \underset{k \notin \{k_1,\ldots,k_{i-1}\}}{\operatorname{argmin}} \; d_k. \tag{3.52}$$

This optimization implies that the data stream corresponding to the minimum MSE entry of the MSE matrix in Eq. (3.51) is decoded first in order to minimize the effect of error propagation inherent to DFE. This algorithm was proposed in [85] and is summarized as a pseudo code in Table 3.1. The algorithm is a Cholesky factorization (e.g. [86]) where a reordering according to Eq. (3.52) is included. Compared to previous ordering proposals, as shown in [11], its complexity order is less than the complexity related to the ordering based on BLAST described in [81] without any penalization with respect to the BER performance.

$\boldsymbol{\Phi} \leftarrow \left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1}\right)^{-1}$
$\boldsymbol{P} \leftarrow \mathbf{I}_N$, $\boldsymbol{D} \leftarrow \mathbf{0}_{N\times N}$
for $i = 1, \ldots, N$
  $q \leftarrow \underset{q'=1,\ldots,N}{\operatorname{argmin}} \boldsymbol{\Phi}(q', q')$
  $\boldsymbol{P}_i \leftarrow \mathbf{I}_N$ whose i-th and q-th rows are exchanged
  $\boldsymbol{P} \leftarrow \boldsymbol{P}_i\boldsymbol{P}$
  $\boldsymbol{\Phi} \leftarrow \boldsymbol{P}_i\boldsymbol{\Phi}\boldsymbol{P}_i^{\mathrm{T}}$
  $\boldsymbol{D}(i,i) \leftarrow \boldsymbol{\Phi}(i,i)$
  $\boldsymbol{\Phi}(i:N,i) \leftarrow \boldsymbol{\Phi}(i:N,i)/\boldsymbol{D}(i,i)$
  $\boldsymbol{\Phi}(i+1:N,i+1:N) \leftarrow \boldsymbol{\Phi}(i+1:N,i+1:N)$
    $-\boldsymbol{\Phi}(i+1:N,i)\boldsymbol{\Phi}(i+1:N,i)^{\mathrm{H}}\boldsymbol{D}(i,i)$
$\boldsymbol{L} \leftarrow$ lower triangular part of $\boldsymbol{\Phi}$
$\boldsymbol{B} \leftarrow \boldsymbol{L}^{-1}$, $\boldsymbol{F} \leftarrow \boldsymbol{D}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}}{\boldsymbol{C}_{\boldsymbol{\eta}}}^{-1}$

Table 3.1: Calculation of WF-DFE Filters with Ordering.

MU-SIMO Zero-Forcing Decision Feedback Receiver (ZF-DFE)

Under the ZF constraint $\boldsymbol{B}\boldsymbol{P} = \boldsymbol{F}\boldsymbol{H}$, i.e. the feedback filter $\mathbf{I} - \boldsymbol{B}$ removes the residual interference at the output of the feedforward filter $\boldsymbol{F}$, the MSE of Eq. (3.39) reduces to [48, 85]

$$\varepsilon_{\mathrm{ZF}}^{\mathrm{DFE}}\left(\boldsymbol{F}\right) = \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{F}^{\mathrm{H}}\right). \tag{3.53}$$

Therefore, the optimization problem can be expressed as

$$\left\{\boldsymbol{P}_{\text{ZF}}^{\text{DFE}}, \boldsymbol{F}_{\text{ZF}}^{\text{DFE}}, \boldsymbol{B}_{\text{ZF}}^{\text{DFE}}\right\} = \underset{\{\boldsymbol{P},\boldsymbol{F},\boldsymbol{B}\}}{\operatorname{argmin}} \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{F}^{\text{H}}\right)$$
$$\text{s.t.: } \boldsymbol{B}\boldsymbol{P} = \boldsymbol{F}\boldsymbol{H} \text{ with } \boldsymbol{B} \text{ unit lower triangular.} \tag{3.54}$$

This enables us to construct the following Lagrangian function,

$$L\left(\boldsymbol{P}, \boldsymbol{F}, \boldsymbol{B}, \lambda, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K\right) = \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}}\boldsymbol{F}^{\text{H}}\right) + 2\Re\left(\operatorname{tr}\left(\sum_{i=1}^{N}\left(\boldsymbol{e}_i^{\text{T}}\boldsymbol{B}\boldsymbol{S}_i^{\text{T}} - \boldsymbol{e}_i^{\text{T}}\boldsymbol{S}_i^{\text{T}}\right)\boldsymbol{\mu}_i\right)\right)$$
$$+ 2\Re\left(\operatorname{tr}\left(\boldsymbol{\Lambda}\left(\boldsymbol{B}\boldsymbol{P} - \boldsymbol{F}\boldsymbol{H}\right)\right)\right) \tag{3.55}$$

where $\boldsymbol{\Lambda} \in \mathbb{C}^{N\times N}$ is the Lagrangian matrix. The selection matrix $\boldsymbol{S}_i$ is given by Eq. (3.41).

Setting the derivatives of Eq. (3.55) to zero, we get the following sufficient KKT conditions since the constraint in Eq. (3.54) is linear:

$$\begin{aligned}
\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{F}^{*}} &= \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{\eta}} - \boldsymbol{\Lambda}^{\text{H}}\boldsymbol{H}^{\text{H}} = \boldsymbol{0} \\
\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{B}^{*}} &= \sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{\mu}_i^{\text{H}}\boldsymbol{S}_i + \boldsymbol{\Lambda}^{\text{H}}\boldsymbol{P}^{\text{T}} = \boldsymbol{0} \\
\boldsymbol{e}_i^{\text{T}}\boldsymbol{B}\boldsymbol{S}_i^{\text{T}} &= \boldsymbol{e}_i^{\text{T}}\boldsymbol{S}_i^{\text{T}} \\
\boldsymbol{P}^{\text{T}}\boldsymbol{B}^{\text{T}} &= \boldsymbol{H}^{\text{T}}\boldsymbol{F}^{\text{T}}.
\end{aligned} \tag{3.56}$$

From the second KKT condition, we obtain that

$$\boldsymbol{\Lambda}^{\text{H}} = -\left(\sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{\mu}_i^{\text{H}}\boldsymbol{S}_i\right)\boldsymbol{P}$$

and after plugging it into the first KKT condition in Eq. (3.56), the feedforward filter $\boldsymbol{F}$ is expressed as

$$\boldsymbol{F} = \boldsymbol{\Lambda}^{\text{H}}\boldsymbol{H}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1} = -\left(\sum_{i=1}^{N}\boldsymbol{e}_i\boldsymbol{\mu}_i^{\text{H}}\boldsymbol{S}_i\right)\boldsymbol{P}\boldsymbol{H}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}.$$

From the ZF constraint, we have $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{H}\boldsymbol{P}^{\text{T}}$. Multiplying from the left by $\boldsymbol{e}_i^{\text{T}}$ and from the right by $\boldsymbol{S}_i^{\text{T}}$, and then applying the constraint to ensure the unit lower triangular structure of the feedback matrix $\boldsymbol{B}$, we obtain that

$$\boldsymbol{e}_i^{\text{T}}\boldsymbol{B}\boldsymbol{S}_i^{\text{T}} = \boldsymbol{e}_i^{\text{T}}\boldsymbol{F}\boldsymbol{H}\boldsymbol{P}^{\text{T}}\boldsymbol{S}_i^{\text{T}} = -\boldsymbol{e}_i^{\text{T}}\left(\sum_{j=1}^{N}\boldsymbol{e}_j\boldsymbol{\mu}_j^{\text{H}}\boldsymbol{S}_j\right)\boldsymbol{P}\boldsymbol{H}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{H}\boldsymbol{P}^{\text{T}}\boldsymbol{S}_i^{\text{T}} = \boldsymbol{e}_i^{\text{T}}\boldsymbol{S}_i^{\text{T}}$$

and then

$$\boldsymbol{\mu}_i^{\mathrm{H}} = -\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left(\boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \right)^{-1} \quad i = 1, \ldots, N$$

since $\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}_j = 0$, for $j \neq i$, and 1 otherwise. Then, the feedforward and feedback filters are expressed as

$$\begin{aligned} \boldsymbol{F} &= \sum_{i=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left(\boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \right)^{-1} \boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ \boldsymbol{B} &= \sum_{i=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \left(\boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \right)^{-1} \boldsymbol{S}_i \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} \boldsymbol{P}^{\mathrm{T}} \end{aligned} \tag{3.57}$$

respectively.

We now define a matrix $\boldsymbol{\Phi} = (\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H})^{-1}$. Since $\boldsymbol{\Phi}$ is Hermitian and positive definite, there exists a unit lower triangular matrix $\boldsymbol{L}$ and a diagonal matrix $\boldsymbol{D}$ such that the following decomposition is satisfied [85, 86]

$$\boldsymbol{P} \boldsymbol{\Phi} \boldsymbol{P}^{\mathrm{T}} = \boldsymbol{L} \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \tag{3.58}$$

with the permutation matrix $\boldsymbol{P}$ introduced in Eq. (3.37). Taking into account the properties satisfied by the selection matrix $\boldsymbol{S}_i$ in Eq. (3.41), it is easy to get the ZF-DFE solution similarly to Eqs. (3.48) and (3.49), which results in

$$\boxed{\begin{aligned} \boldsymbol{F}_{\mathrm{ZF}}^{\mathrm{DFE}} &= \boldsymbol{D} \boldsymbol{L}^{\mathrm{H}} \boldsymbol{P} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \\ \boldsymbol{B}_{\mathrm{ZF}}^{\mathrm{DFE}} &= \boldsymbol{L}^{-1}. \end{aligned}} \tag{3.59}$$

With the above expressions for the precoding filters, the MSE in Eq. (3.53) reads as

$$\varepsilon_{\mathrm{ZF}}^{\mathrm{DFE}} = \mathrm{tr}\,(\boldsymbol{D}) = \sum_{k=1}^{N} d_i \tag{3.60}$$

i.e. the same result based on Eq. (3.58) as for WF-DFE based on Eq. (3.47) is obtained. The proposal for ordering is similar to the Wiener approach, but the starting matrix in Table 3.1 is now $\boldsymbol{\Phi} = (\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H})^{-1}$ instead of $\boldsymbol{\Phi} = (\boldsymbol{H}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{\eta}}^{-1} \boldsymbol{H} + \boldsymbol{C}_{\boldsymbol{u}}^{-1})^{-1}$.

3.2.3 Simulation Results

Again, we show some results obtained via computer simulations for $N = 4$ transmitting antenna elements and $K = 4$ receiving users. We average $5{,}000$ channel realizations and $N_{\mathrm{B}} = 50$ symbols are considered per channel realization. These symbols are QPSK modulated.

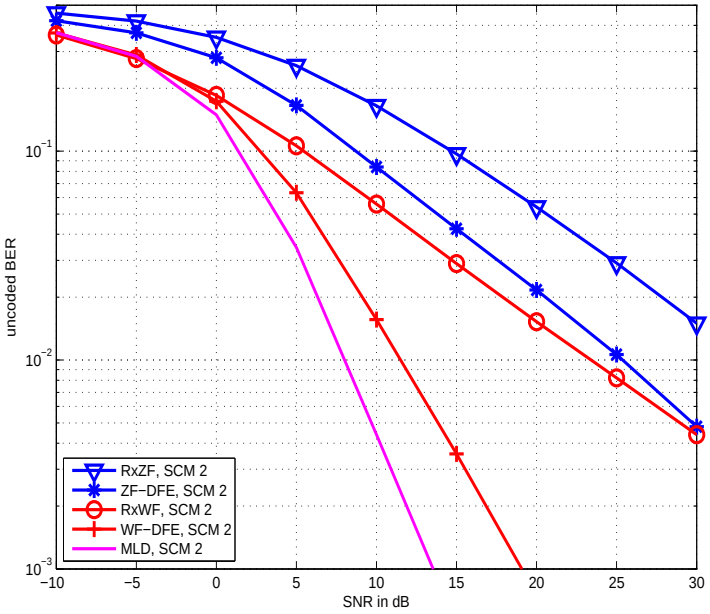


Figure 3.8: Uncoded BER vs. SNR for Receive Filters: QPSK Transmission over Correlated Flat Fading MU-SIMO Channels (*SCM 2*) with Four Transmitting Antenna Elements and Four Users.

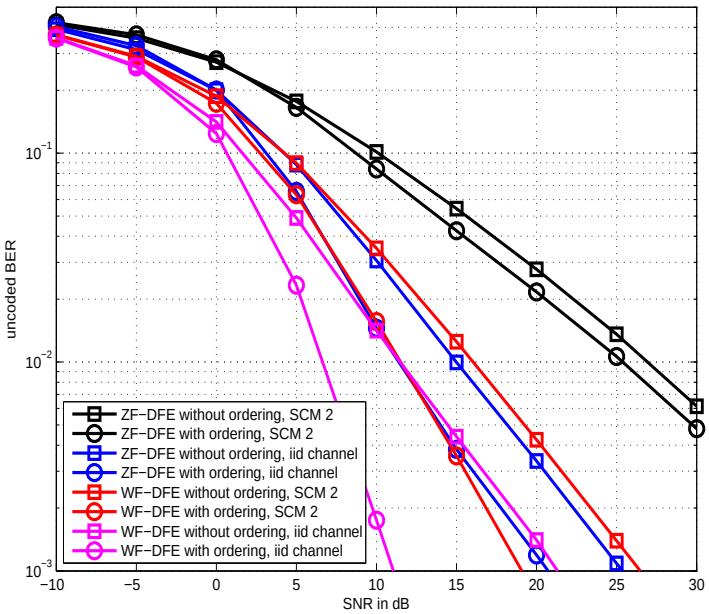


Figure 3.9: Uncoded BER vs. SNR for Receive Filters: QPSK Transmission over Correlated (*SCM 2*) and Uncorrelated Flat Fading MU-SIMO Channels with Four Transmitting Antenna Elements and Four Users.

Fig. 3.8 depicts a comparison between different types of receive filters: MLD, ZF-DFE and WF-DFE, and the RxZF and RxWF schemes described in Section 3.1.1. It can be seen that MLD is the optimum receive processing, although its computational complexity is too high for many practical systems. It is obvious that the Wiener filters are always superior to the respective zero-forcing filters. However, the WF-DFE approach shows a slight loss in performance for low SNR compared to the linear Wiener receive processing, due to the effect of error propagation.

Fig. 3.9 shows the uncoded BER performance for DFE with and without ordering. Obviously, an optimized ordering improves the final results substantially, as can be seen when comparing the curves marked with circles to the curves marked with squares, even for the case of correlated channels, where for a BER of 10^{-3} a gain of about $7.5\,$dB is obtained.

3.3 MU-MISO Nonlinear Transmit Processing

Research on transmit processing has received a great deal of interest in recent years due to the lack of degrees of freedom in the downlink and the limitations of power and complexity in the receivers of wireless communications systems. In multiuser MISO systems cooperation is not often to be found between the receivers, and transmit processing is mandatory if we wish to implement efficient filtering methods that remove interference. This filtering process prior to transmission is referred to as *precoding*. In this section we focus on nonlinear precoding due to its superior performance compared to that of the linear precoders explained in Section 3.1.2.

Assuming the CSI is available at the transmitter, a lattice search similar to that of MLD can be performed at the transmitter, resulting in the precoding scheme called *Vector Precoding* (VP). A perturbation vector is directly added to the data signal and this signal is then precoded by linear filtering [23]. In fact, *Tomlinson-Harashima Precoding* (THP) is a constrained type of vector precoding where the elements of the perturbation vector are successively computed. THP is based on a feedforward and a feedback filter, as in DFE, but with both located at the transmitter side instead of the receiver side. The error propagation of DFE is avoided by moving the filters to the transmitter, since the transmitter knows the signal to be fed back. Ordering strategies improve the achieved THP performance so we will apply successive algorithms to find the optimum ordering, in a similar manner for the algorithms shown for DFE.

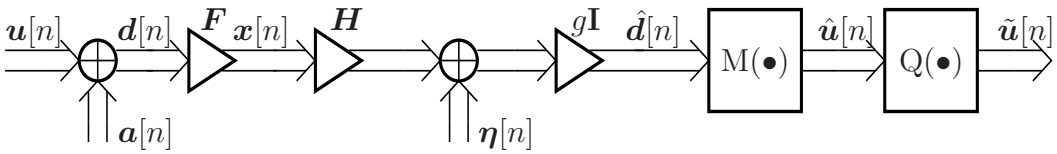


Figure 3.10: MU–MISO System with Vector Precoding.

3.3.1 MU-MISO Vector precoding (VP)

Fig. 3.10 shows the block diagram of a MU-MISO system with vector precoding. The transmitter has the freedom to add an arbitrary perturbation signal $\boldsymbol{a}[n] \in \tau\mathbb{Z}^K + \mathrm{j}\,\tau\mathbb{Z}^K$ to the data signal prior to linear transformation with the filter $\boldsymbol{F} \in \mathbb{C}^{N\times K}$, since the receivers apply the modulo operator $\mathrm{M}(\bullet)$. Here, τ denotes a constant that depends on the modulation alphabet, so we set $\tau = 2\sqrt{2}$ for QPSK modulation (see Fig. 3.11) and $\tau = 8/\sqrt{10}$ for 16QAM modulation [83]. This constant is associated with the modulo operator $\mathrm{M}(\bullet)$. This nonlinear operation is defined as

$$\mathrm{M}\left(x\right) = x - \left(\left\lfloor \frac{\Re\left(x\right)}{\tau} + \frac{1}{2} \right\rfloor \tau + \mathrm{j} \left\lfloor \frac{\Im\left(x\right)}{\tau} + \frac{1}{2} \right\rfloor \tau\right) \in \mathbb{V} \tag{3.61}$$

where $\lfloor \bullet \rfloor$ denotes the floor operator which gives the largest integer smaller than or equal to the argument. The corresponding fundamental Voronoi region is

$$\mathbb{V} = \left\{ x \in \mathbb{C} \,|\, -\frac{\tau}{2} \leq \Re\left(x\right) < \frac{\tau}{2}, -\frac{\tau}{2} \leq \Im\left(x\right) < \frac{\tau}{2} \right\}$$

which means that the modulo operator constrains the real and imaginary part of x to the interval $[-\tau/2, \tau/2]$ by adding integer multiples of τ and $\mathrm{j}\,\tau$ to the real and imaginary part, respectively. For example, for $x = 3.4 - 1.5\,\mathrm{j}$ and $\tau = 2$, when the modulo operator is applied we get $\mathrm{M}(x) = -0.6 + 0.5\,\mathrm{j}$. Note that if we apply the modulo operator to a multidimensional vector $\boldsymbol{x} = [x_1, \ldots, x_K]^{\mathrm{T}}$, it is satisfied that

$$\mathrm{M}(\boldsymbol{x}) = [\mathrm{M}(x_1), \ldots, \mathrm{M}(x_K)]^{\mathrm{T}} \in \mathbb{V}^K$$

where $\mathrm{M}(x_i)$, $i = 1, \ldots, K$ is defined as in Eq. (3.61).

As can be seen from Fig. 3.10, the data vector $\boldsymbol{u}[n] \in \mathbb{C}^K$ is first superimposed with the perturbation vector $\boldsymbol{a}[n]$, and the resulting vector is then processed by the linear filter $\boldsymbol{F}$ to form the transmit vector

$$\boldsymbol{x}[n] = \boldsymbol{F}\boldsymbol{d}[n] \in \mathbb{C}^N, \quad n = 1, \ldots, N_{\mathrm{B}}$$

where $\boldsymbol{d}[n]$ is the desired signal given by $\boldsymbol{u}[n] + \boldsymbol{a}[n]$ and n is the symbol index in a block size of N_{B} data symbols.

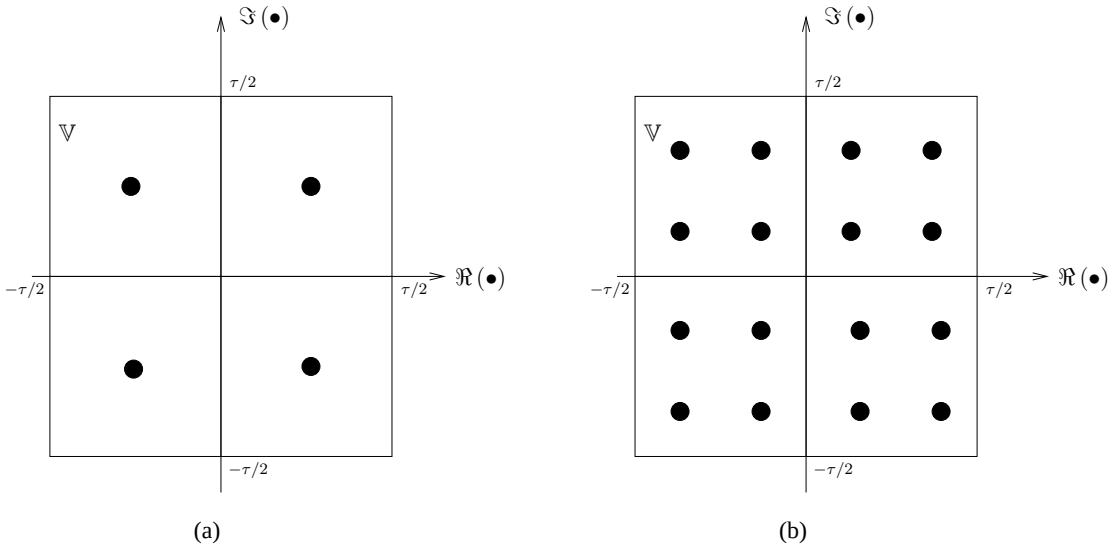


Figure 3.11: Modulo Operator. (a) QPSK, (b) 16QAM.

Similar to linear precoding filters, we impose a transmit power constraint. Since the statistics of the transmit symbols are unknown, we average over the block instead of taking the expected value, i.e.

$$\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\|\boldsymbol{x}[n]\|_2^2 \leq E_{\mathrm{tx}}.$$

After passing through the channel and by superimposing the AWGN noise, the received signal is given by

$$\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{F}\boldsymbol{d}[n] + \boldsymbol{\eta}[n].$$

The weight g in Fig. 3.10 is assumed to be constant throughout the block of N_{B} symbols. Note that we use a common weight for all the users. Thus, the weighted estimated signal is given by

$$\hat{\boldsymbol{d}}[n] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{d}[n] + g\boldsymbol{\eta}[n]. \tag{3.62}$$

The modulo operator at the receiver is used to compensate the effect of adding the perturbation $\boldsymbol{a}[n]$ at the transmitter.

The MSE can be expressed as [22, 23]

$$\begin{aligned}\varepsilon_{\mathrm{WF}}^{\mathrm{VP}}\left(\boldsymbol{a}[n],\boldsymbol{x}[n],g\right) &= \frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\mathrm{E}\left[\left\|\boldsymbol{d}[n]-\hat{\boldsymbol{d}}[n]\right\|_2^2\middle|\boldsymbol{u}[n]\right] \\ &= \frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\mathrm{E}\left[\|\boldsymbol{d}[n]-g\boldsymbol{H}\boldsymbol{x}[n]-g\boldsymbol{\eta}[n]\|_2^2\,|\boldsymbol{u}[n]\right]. \end{aligned}\tag{3.63}$$

Note that the expectation is conditioned on the full knowledge of the symbols $\boldsymbol{u}[n]$ by the transmitter. But since the statistics of $\boldsymbol{a}[n]$ are unknown, we average the symbol MSE over the whole block.

Since $\boldsymbol{a}[n]$ is discrete, we cannot derive with respect to $\boldsymbol{a}$. The optimization procedure is as follows. We start by fixing $\boldsymbol{a}$, after which $\boldsymbol{x}$ and g are optimized taking into account the transmit power constraint. For these optimum $\boldsymbol{x}$ and g we choose the best $\boldsymbol{a}$ according to the MSE criterion. Although we optimize the continuous and discrete part separately, this procedure leads to the optimum minimization of Eq. (3.63) [23].

MU-MISO Wiener Spatial Vector Precoding (WF-VP)

We have to find the joint optimum of all the perturbation vectors $\boldsymbol{a}[n]$, all the transmit vectors $\boldsymbol{x}[n]$, and the gain factors g for $n = 1, \ldots, N_{\mathrm{B}}$:

$$\left\{\boldsymbol{a}_{\mathrm{WF}}^{\mathrm{VP}}[n], \boldsymbol{x}_{\mathrm{WF}}^{\mathrm{VP}}[n], g_{\mathrm{WF}}^{\mathrm{VP}}\right\} = \underset{\{\boldsymbol{a}[n], \boldsymbol{x}[n], g\}}{\operatorname{argmin}} \varepsilon_{\mathrm{WF}}^{\mathrm{VP}}\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) \quad \text{s.t.: } \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \|\boldsymbol{x}[n]\|_2^2 \leq E_{\mathrm{tx}}. \tag{3.64}$$

The MSE $\varepsilon_{\mathrm{WF}}^{\mathrm{VP}}(\boldsymbol{a}[n], \boldsymbol{x}[n], g)$ is given by Eq. (3.63) [22, 23] and can be rewritten as

$$\begin{aligned}\varepsilon_{\mathrm{WF}}^{\mathrm{VP}}\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) = \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \big(&\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{d}[n] - g^* \boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{d}[n] - g\boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{H}\boldsymbol{x}[n] \\ &+ |g|^2 \boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{x}[n] + |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\big)\end{aligned} \tag{3.65}$$

where we use $\mathrm{E}[\|\boldsymbol{d}[n]\|_2^2 \,|\boldsymbol{u}[n]] = \|\boldsymbol{d}[n]\|_2^2$ and $\mathrm{E}[\|\boldsymbol{x}[n]\|_2^2 \,|\boldsymbol{u}[n]] = \|\boldsymbol{x}[n]\|_2^2$, since the data signal $\boldsymbol{u}[n]$, the perturbation signal $\boldsymbol{a}[n]$ and consequently, the transmitted signal $\boldsymbol{x}[n]$ are known to the transmitter.

The Lagrangian function can be expressed as

$$L\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g, \lambda\right) = \varepsilon_{\mathrm{WF}}^{\mathrm{VP}}\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) + \lambda \left(\frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{x}[n] - E_{\mathrm{tx}}\right) \tag{3.66}$$

where $\lambda \in \mathbb{R}^{0,+}$. Now, we set its derivative with respect to $\boldsymbol{x}[n], n = 1, \ldots, N_{\mathrm{B}}$ and g to

zero, which leads to the necessary KKT conditions

$$\begin{aligned}\frac{\partial L(\bullet)}{\partial \boldsymbol{x}^*[n]} &= \frac{1}{N_\mathrm{B}}\left(-g^*\boldsymbol{H}^\mathrm{H}\boldsymbol{d}[n] + |g|^2\,\boldsymbol{H}^\mathrm{H}\boldsymbol{H}\boldsymbol{x}[n]\right) + \frac{\lambda}{N_\mathrm{B}}\boldsymbol{x}[n] = \boldsymbol{0}\\ \frac{\partial L(\bullet)}{\partial g} &= \frac{1}{N_\mathrm{B}}\left(-\boldsymbol{d}^\mathrm{H}[n]\boldsymbol{H}\boldsymbol{x}[n] + g^*\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{H}^\mathrm{H}\boldsymbol{H}\boldsymbol{x}[n]\right.\\ &\quad \left. + g^*\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\right) = 0\end{aligned}$$

$$\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{x}[n] \leq E_\mathrm{tx}$$

$$\lambda\left(\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{x}[n] - E_\mathrm{tx}\right) = 0 \quad \text{with } \lambda \geq 0 \tag{3.67}$$

since the optimization problem in Eq. (3.64) is not a convex programming problem.

Then, the transmit symbols are directly obtained from the first KKT condition and are given by

$$\boldsymbol{x}[n] = \frac{1}{g}\left(\boldsymbol{H}^\mathrm{H}\boldsymbol{H} + \frac{\lambda}{|g|^2}\boldsymbol{I}\right)^{-1}\boldsymbol{H}^\mathrm{H}\boldsymbol{d}[n]. \tag{3.68}$$

First of all, we have to show that $\lambda > 0$, i.e. the power constraint as active. Multiplying the second KKT condition by g, we have

$$\frac{1}{N_\mathrm{B}}\left(-g\boldsymbol{d}^\mathrm{H}[n]\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\,\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{H}^\mathrm{H}\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\right) = 0 \tag{3.69}$$

and multiplying the Hermitian of the first KKT condition by $\boldsymbol{x}[n]$ from the right, we have

$$\frac{1}{N_\mathrm{B}}\left(-g\boldsymbol{d}^\mathrm{H}[n]\boldsymbol{H}\boldsymbol{x}[n] + |g|^2\,\boldsymbol{x}[n]^\mathrm{H}\boldsymbol{H}^\mathrm{H}\boldsymbol{H}\boldsymbol{x}[n]\right) + \frac{\lambda}{N_\mathrm{B}}\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{x}[n] = 0. \tag{3.70}$$

With Eq. (3.69) and the transmit energy constraint, the Lagrangian multiplier λ is given by

$$\lambda = |g|^2\,\frac{\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)}{\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{x}[n]}. \tag{3.71}$$

Therefore, it becomes clear that $\lambda > 0$ for the non–trivial case that $\exists n : \boldsymbol{x}[n] \neq \boldsymbol{0}$. Thus, the transmit energy constraint is active and $\lambda = |g|^2\,\xi$ with $\xi = \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})/E_\mathrm{tx}$.

Then, we reach the following solution for the WF-VP:

$$\boxed{\begin{aligned}\boldsymbol{x}_\mathrm{WF}^\mathrm{VP}[n] &= \frac{1}{g_\mathrm{WF}^\mathrm{VP}}\left(\boldsymbol{H}^\mathrm{H}\boldsymbol{H} + \xi\boldsymbol{I}\right)^{-1}\boldsymbol{H}^\mathrm{H}\boldsymbol{d}[n]\\ g_\mathrm{WF}^\mathrm{VP} &= \sqrt{\frac{\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{d}^\mathrm{H}[n]\boldsymbol{H}\left(\boldsymbol{H}^\mathrm{H}\boldsymbol{H} + \xi\boldsymbol{I}\right)^{-2}\boldsymbol{H}^\mathrm{H}\boldsymbol{d}[n]}{E_\mathrm{tx}N_\mathrm{B}}}\end{aligned}} \tag{3.72}$$

where $g_{\text{WF}}^{\text{VP}}$ is directly obtained from the transmit energy constraint and $\xi = \frac{\text{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\text{tx}}}$. Remember that g is chosen only once in each block.

We define a matrix $\boldsymbol{\Phi} = (\boldsymbol{H}\boldsymbol{H}^{\text{H}} + \xi\mathbf{I})^{-1}$. Applying the matrix inversion lemma to Eq. (3.72) shows that $\boldsymbol{x}_{\text{WF}}^{\text{VP}}[n] = \frac{1}{g_{\text{WF}}^{\text{VP}}}\boldsymbol{H}^{\text{H}}\boldsymbol{\Phi}\boldsymbol{d}[n]$ and then, $g_{\text{WF}}^{\text{VP}} = \sqrt{\sum_{n=1}^{N_{\text{B}}}(\boldsymbol{d}^{\text{H}}[n]\boldsymbol{\Phi}^{\text{H}}\boldsymbol{H}\boldsymbol{H}^{\text{H}}\boldsymbol{\Phi}\boldsymbol{d}[n])/(E_{\text{tx}}N_{\text{B}})}$. Thus, when we plug these results into the MSE expression in Eq. (3.65) we obtain that

$$\varepsilon_{\text{WF}}^{\text{VP}}\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) = \frac{\xi}{N_{\text{B}}}\sum_{n=1}^{N_{\text{B}}}\boldsymbol{d}^{\text{H}}[n]\boldsymbol{\Phi}\boldsymbol{d}[n]. \tag{3.73}$$

Since $\boldsymbol{\Phi}$ is positive definite, we can use the Cholesky factorization to obtain a lower triangular matrix $\boldsymbol{L}$ and a diagonal matrix $\boldsymbol{D}$ with the following relationship [22],

$$\boldsymbol{\Phi} = \left(\boldsymbol{H}\boldsymbol{H}^{\text{H}} + \xi\mathbf{I}\right)^{-1} = \boldsymbol{L}^{\text{H}}\boldsymbol{D}\boldsymbol{L}.$$

Thus, the perturbation signal can be found by the following search [22]

$$\boxed{\begin{aligned}\boldsymbol{a}_{\text{WF}}^{\text{VP}}[n] &= \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^K+\text{j}\,\tau\mathbb{Z}^K}{\text{argmin}}\ (\boldsymbol{u}[n]+\boldsymbol{a}[n])^{\text{H}}\boldsymbol{\Phi}(\boldsymbol{u}[n]+\boldsymbol{a}[n]) \\ &= \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^K+\text{j}\,\tau\mathbb{Z}^K}{\text{argmin}}\ ||\boldsymbol{D}^{1/2}\boldsymbol{L}(\boldsymbol{u}[n]+\boldsymbol{a}[n])||_2^2\end{aligned}} \tag{3.74}$$

This search can be solved by means of the Schnorr-Euchner sphere decoding [87, 88] where the use of real-valued notation to represent vectors and matrices has been considered to run the final computer simulations (see Appendix B.5).

Note that due to the unit lower triangular structure of $\boldsymbol{D}^{1/2}\boldsymbol{L}$, the i–th summand of the Euclidean norm $||\boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{u}[n] + \boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{a}[n]||_2^2 = \sum_{i=1}^{K}\lambda_i$ is given by

$$\lambda_i = d_{i,i}\left|u_i[n] + a_i[n] + \sum_{j=1}^{i-1} l_{i,j}\left(u_j[n] + a_j[n]\right)\right|^2 \tag{3.75}$$

where $l_{i,j}$ corresponds to the element of the i-th row and j-th column of $\boldsymbol{L}$ and $d_{i,i}$ is the i-th entry of the diagonal matrix $\boldsymbol{D}$.

When the off-diagonal elements of $\boldsymbol{L}$ are approximately zero, i.e. $l_{i,j} = 0$, for $j \neq i$, we have

$$\boldsymbol{a}[n] = \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^K+\text{j}\,\tau\mathbb{Z}^K}{\text{argmin}}\ ||\boldsymbol{u}[n]+\boldsymbol{a}[n]||_2^2 \tag{3.76}$$

which leads to $\boldsymbol{a}[n] = \mathbf{0}$, i.e. we obtain the linear precoding approach described in Subsection 3.1.2.

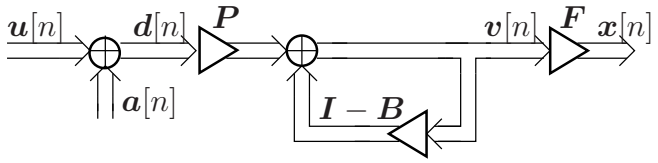


Figure 3.12: Linear Representation of Tomlinson Harashima Precoding.

When $a_1[n], \ldots, a_K[n]$ are computed successively, i.e. $a_i[n]$ is found for fixed $a_1[n], \ldots, a_{i-1}[n]$, the i–th element of $\boldsymbol{a}[n]$ is obtained as

$$a_i[n] = -\,\mathrm{Q}_{\tau\mathbb{Z}^K+\mathrm{j}\,\tau\mathbb{Z}^K}\left(u_i[n] + \sum_{j=1}^{i-1} l_{i,j}\,(u_j[n] + a_j[n])\right). \tag{3.77}$$

This successive computation of $\boldsymbol{a}[n]$ enables us to obtain the scheme depicted in Fig. 3.12, which corresponds to the linear representation of THP. According to the definition of the modulo operator as $\mathrm{M}(x) = x - (\lfloor \frac{\Re(x)}{\tau} + \frac{1}{2} \rfloor \tau + \mathrm{j} \lfloor \frac{\Im(x)}{\tau} + \frac{1}{2} \rfloor \tau)$, it is straightforward to see that the quantizer $\mathrm{Q}_{\tau\mathbb{Z}^K+\mathrm{j}\,\tau\mathbb{Z}^K}(x)$ is equivalent to the term $\lfloor \frac{\Re(x)}{\tau} + \frac{1}{2} \rfloor \tau + \mathrm{j} \lfloor \frac{\Im(x)}{\tau} + \frac{1}{2} \rfloor \tau$ and then we can write $\mathrm{Q}_{\tau\mathbb{Z}^K+\mathrm{j}\,\tau\mathbb{Z}^K}(x) = x - \mathrm{M}(x)$. Therefore, the perturbation signal $\boldsymbol{a}[n]$ can easily be included inside the feedback loop (as can be seen in Fig. 3.13) by means of the modulo operator $\mathrm{M}(x)$. This leads to the well-known suboptimum approach of Tomlinson-Harashima precoding described in the following subsection.

The above result for $\boldsymbol{a}[n]$ can be transformed to

$$\begin{aligned}
\boldsymbol{a}_{\mathrm{WF}}^{\mathrm{VP}}[n] &= \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^K+\mathrm{j}\,\tau\mathbb{Z}^K}{\operatorname{argmin}} ||\boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{u}[n] + \boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{a}[n]||_2^2 \\
&= \underset{\boldsymbol{\lambda}[n]\in\mathbb{Z}^K+\mathrm{j}\,\mathbb{Z}^K}{\operatorname{argmin}} ||\tau\boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{\lambda}[n] - (-\boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{u}[n])||_2^2 \\
&= \underset{\boldsymbol{\lambda}[n]\in\mathbb{Z}^K+\mathrm{j}\,\mathbb{Z}^K}{\operatorname{argmin}} ||\boldsymbol{G}\boldsymbol{\lambda}[n] - \boldsymbol{z}[n]||_2^2
\end{aligned} \tag{3.78}$$

where $\boldsymbol{G} = \tau\boldsymbol{D}^{1/2}\boldsymbol{L}$ and $\boldsymbol{z}[n] = -\boldsymbol{D}^{1/2}\boldsymbol{L}\boldsymbol{u}[n]$. This is the called a *closest point search* in the lattice generated by the matrix $\boldsymbol{G}$ [89].

According to the *Lenstra-Lenstra-Lovász* (LLL) algorithm [90], this matrix $\boldsymbol{G}$ can be decomposed into a matrix $\boldsymbol{\Gamma}$ and a unimodular integer matrix $\boldsymbol{T}$, i.e. the absolute value of its determinant is equal to one, as follows

$$\boldsymbol{G} = \boldsymbol{\Gamma}\boldsymbol{T}^{-1}.$$

Note that the inverse of $\boldsymbol{T}$ is also unimodular integer.

Thus, $\boldsymbol{G}\boldsymbol{\lambda}[n] = \boldsymbol{\Gamma}\boldsymbol{T}^{-1}\boldsymbol{\lambda}[n] = \boldsymbol{\Gamma}\boldsymbol{\lambda}'[n]$ with integer $\boldsymbol{\lambda}'[n] \in \mathbb{Z}^K + \mathrm{j}\,\mathbb{Z}^K$. Based on the above factorization of the generator matrix $\boldsymbol{G}$, the lattice search of Eq. (3.78) can be

rewritten as follows

$$\boldsymbol{\lambda}_{\text{opt}}[n] = \boldsymbol{T} \underset{\boldsymbol{\lambda}'[n]\in\mathbb{Z}^K+\mathrm{j}\,\mathbb{Z}^K}{\operatorname{argmin}} ||\boldsymbol{\Gamma}\boldsymbol{\lambda}'[n] - \boldsymbol{x}[n]||_2^2 = \boldsymbol{T}\boldsymbol{\lambda}'_{\text{opt}}[n].$$

Since the columns of $\boldsymbol{\Gamma}$ are closer to orthogonal than those of the original $\boldsymbol{G}$, this search can be solved more efficiently than in Eq. (3.78).

In order to find $\boldsymbol{\lambda}'_{\text{opt}}[n]$ we employ the Schnorr-Euchner algorithm [87, 88], where a sphere decoder performs this lattice search. Then, the vector $\boldsymbol{a}_{\text{WF}}^{\text{VP}}[n]$ is simply given by

$$\boldsymbol{a}_{\text{WF}}^{\text{VP}}[n] = \tau\boldsymbol{\lambda}_{\text{opt}}[n] \in \tau\mathbb{Z}^K + \mathrm{j}\,\tau\mathbb{Z}^K.$$

Note that the complexity of the sphere decoder grows exponentially with the number of users [89] which implies that the implementation of VP in real systems may be questionable.

MU-MISO Zero-Forcing Spatial Vector Precoding (ZF-VP)

By considering the ZF constraint $\mathrm{E}[\hat{\boldsymbol{d}}[n] \mid \boldsymbol{d}[n]] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{d}[n]$, for $n = 1, \ldots, N_{\text{B}}$, with [cf. Eq. (3.63)]

$$\hat{\boldsymbol{d}}[n] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{d}[n] + g\boldsymbol{\eta}[n]$$

the MSE in Eq. (3.63) reduces to

$$\varepsilon_{\text{ZF}}^{\text{VP}}\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) = |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right). \tag{3.79}$$

Thus, the optimization problem is expressed as

$$\begin{aligned}\left\{\boldsymbol{a}_{\text{ZF}}^{\text{VP}}[n], \boldsymbol{x}_{\text{ZF}}^{\text{VP}}[n], g_{\text{ZF}}^{\text{VP}}\right\} = &\underset{\{\boldsymbol{a}[n],\boldsymbol{x}[n],g\}}{\operatorname{argmin}} |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) \\ &\text{s.t.: } \frac{1}{N_{\text{B}}}\sum_{n=1}^{N_{\text{B}}} \|\boldsymbol{x}[n]\|_2^2 \leq E_{\text{tx}} \quad \text{and} \quad \mathrm{E}[\hat{\boldsymbol{d}}[n] \mid \boldsymbol{d}[n]] = \boldsymbol{d}[n].\end{aligned} \tag{3.80}$$

We can form the Lagrangian function as follows,

$$\begin{aligned}L\left(\boldsymbol{x}[n], \boldsymbol{\mu}[n], g, \lambda\right) = \varepsilon_{\text{ZF}}^{\text{VP}}\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) + 2\Re\left(\operatorname{tr}\left(\frac{1}{N_{\text{B}}}\sum_{n=1}^{N_{\text{B}}} \boldsymbol{\mu}^{\text{T}}[n]\left(g\boldsymbol{H}\boldsymbol{x}[n] - \boldsymbol{d}[n]\right)\right)\right) \\ + \lambda\left(\frac{1}{N_{\text{B}}}\sum_{n=1}^{N_{\text{B}}} \boldsymbol{x}^{\text{H}}[n]\boldsymbol{x}[n] - E_{\text{tx}}\right)\end{aligned} \tag{3.81}$$

where $\lambda \in \mathbb{R}^{0,+}$ and $\boldsymbol{\mu}[n] \in \mathbb{C}^K$, $n = 1, \ldots, N_{\text{B}}$.

From the Lagrangian function, we can obtain the following KKT conditions:

$$\begin{aligned}
\frac{\partial L(\bullet)}{\partial \boldsymbol{x}^*[n]} &= \frac{1}{N_\mathrm{B}}\left(g^*\boldsymbol{H}^\mathrm{H}\boldsymbol{\mu}^*[n]\right) + \frac{\lambda}{N_\mathrm{B}}\boldsymbol{x}[n] = \mathbf{0} \\
\frac{\partial L(\bullet)}{\partial g} &= g^*\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) + \mathrm{tr}\left(\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{\mu}[n]^\mathrm{T}\boldsymbol{H}\boldsymbol{x}[n]\right) = 0 \\
g\boldsymbol{H}\boldsymbol{x}[n] &= \boldsymbol{d}[n] \quad \forall n \in \{1,\ldots,N_\mathrm{B}\} \\
\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\mathrm{tr}\left(\boldsymbol{x}[n]\boldsymbol{x}^\mathrm{H}[n]\right) &\leq E_\mathrm{tx} \\
\lambda\left(\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{x}[n] - E_\mathrm{tx}\right) &= 0 \quad \text{with } \lambda \geq 0
\end{aligned}$$

that are only necessary to find the solution including the zero-forcing constraint.

It is easy to show that the transmit energy constraint is active. Indeed, multiplying the first KKT condition by $\boldsymbol{x}^\mathrm{H}[n]$ from the left, summing over $n = 1,\ldots,N_\mathrm{B}$, and taking into account again the transmit energy constraint, we get

$$\lambda\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{x}[n] = -\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}g^*\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{H}^\mathrm{H}\boldsymbol{\mu}^*[n] = -\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{d}^\mathrm{H}[n]\boldsymbol{\mu}^*[n]$$

where the last equality is obtained from the ZF constraint [third KKT condition in Eq. (3.81)]. With this result and the second KKT condition, we get

$$\lambda = \frac{|g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)}{\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{x}^\mathrm{H}[n]\boldsymbol{x}[n]}$$

and therefore $\lambda > 0$ as long as $|g|^2 \neq 0$. So, the transmit energy constraint is always active.

Combining the first KKT condition with the third and with the transmit energy constraint, we can obtain, respectively, the transmit symbols and the receive weights as follows

$$\boxed{\begin{aligned}
\boldsymbol{x}_\mathrm{ZF}^\mathrm{VP}[n] &= \frac{1}{g}\boldsymbol{H}^\dagger\boldsymbol{d}[n] = \frac{1}{g}\boldsymbol{H}^\mathrm{H}\left(\boldsymbol{H}\boldsymbol{H}^\mathrm{H}\right)^{-1}\boldsymbol{d}[n] \\
g_\mathrm{ZF}^\mathrm{VP} &= \sqrt{\frac{\frac{1}{N_\mathrm{B}}\sum_{n=1}^{N_\mathrm{B}}\boldsymbol{d}^\mathrm{H}[n]\left(\boldsymbol{H}\boldsymbol{H}^\mathrm{H}\right)^{-1}\boldsymbol{d}[n]}{E_\mathrm{tx}}}
\end{aligned}} \tag{3.82}$$

where $\boldsymbol{H}^\dagger = \boldsymbol{H}^\mathrm{H}(\boldsymbol{H}\boldsymbol{H}^\mathrm{H})^{-1}$ denotes the pseudo inverse of $\boldsymbol{H}$. Note that for the existence of the zero–forcing solution, i.e. $\boldsymbol{H}\boldsymbol{H}^\mathrm{H}$ is regular or the zero–forcing condition can be

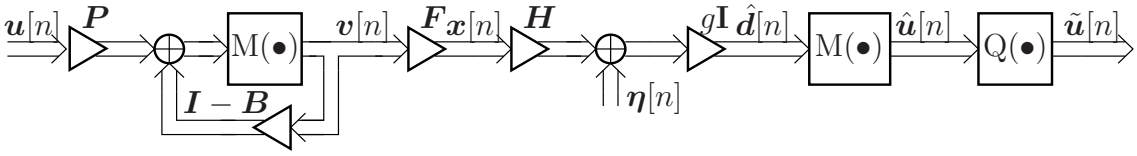


Figure 3.13: MU–MISO System with Tomlinson Harashima Precoding.

fulfilled, it is necessary that $N \geq K$. We assume that g is real and positive to ensure a unique solution. Note that by applying the matrix inversion lemma to Eq. (3.72), we get that $(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi\mathbf{I})^{-1}\boldsymbol{H}^{\mathrm{H}} = \boldsymbol{H}^{\mathrm{H}}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi\mathbf{I})^{-1}$ and it is easy to see that for $\xi = \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})/E_{\mathrm{tx}} \to 0$ the Wiener VP solution converges to the ZF approach.

Plugging g into the MSE in Eq. (3.79) yields

$$\varepsilon_{\mathrm{ZF}}^{\mathrm{VP}}\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) = \frac{\xi}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{d}^{\mathrm{H}}[n] \left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1} \boldsymbol{d}[n].$$

Due to $\boldsymbol{d}[n] = \boldsymbol{u}[n] + \boldsymbol{a}[n]$, the optimum perturbation vectors are found by the following closest point search in a lattice [23, 91]

$$\boxed{\boldsymbol{a}_{\mathrm{ZF}}^{\mathrm{VP}}[n] = \underset{\boldsymbol{a}[n] \in \tau\mathbb{Z}^K + j\tau\mathbb{Z}^K}{\operatorname{argmin}} \left\| \boldsymbol{H}^{\mathrm{H}} \left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1} \left(\boldsymbol{u}[n] + \boldsymbol{a}[n]\right) \right\|_2^2} \tag{3.83}$$

since $(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}})^{-1} = (\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}})^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}})^{-1}$ and $\boldsymbol{z}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{z} = \|\boldsymbol{A}\boldsymbol{z}\|_2^2$.

3.3.2 MU-MISO Tomlinson-Harashima Precoding (THP)

Fig. 3.13 shows the block diagram of a MU-MISO system with THP. The basic equation for the computation of the perturbation signal of THP is Eq. (3.77), i.e. it is computed successively. Based on Eq. (3.77), WF–THP can be obtained from WF–VP and ZF–THP from ZF–VP. However, the performance of THP is heavily dependent on the precoding order. Therefore, the goal of this subsection is to find the appropriate precoding order based on a THP specific optimization.

At the transmitter, the feedforward filter $\boldsymbol{F}$ linearly suppresses parts of the interference, whereas the feedback loop with the strictly lower triangular feedback filter nonlinearly $\boldsymbol{I} - \boldsymbol{B}$ subtracts the remaining interference. The feedforward matrix forces spatial causality (i.e. outputs depend on current and past entries, but not on future entries) and, additionally, the feedback filter must be strictly lower triangular to ensure causality for the feedback loop and ISI cancellation. Since the order of precoding has an important effect on the performance [21, 92], the data signal $\boldsymbol{u}[n] \in \mathbb{C}^K$ is reordered by means of the permutation filter $\boldsymbol{P} = \sum_{i=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_{k_i}^{\mathrm{T}} \in \{0,1\}^{K \times K}$, where $\boldsymbol{e}_i$ is the i-th column of the

$\boldsymbol{v} \leftarrow \boldsymbol{P}\boldsymbol{u}$
for $i = 1, \ldots, K$
    $\boldsymbol{v}(i) \leftarrow \mathrm{M}(2\boldsymbol{v}(i) - \boldsymbol{B}(i,:)\boldsymbol{v})$

Table 3.2: Computing the Feedback Loop Output from the Permuted Data.

identity matrix $\mathbf{I}_K$ and k_i is the index of the i-th data stream to be precoded. Remember that the permutation matrix $\boldsymbol{P}$ satisfies $\boldsymbol{P}^{-1} = \boldsymbol{P}^{\mathrm{T}}$.

The signal $\boldsymbol{P}\boldsymbol{u}[n]$ is passed through the feedback loop to get the output $\boldsymbol{v}[n]$, which can be computed from the pseudo code in Table 3.2. The entries of $\boldsymbol{v}[n]$ have statistical properties which approximately only depend on the modulo constant τ (see e.g. [83]). Remember that τ depends on the modulation alphabet (see Fig. 3.11). Interestingly, the covariance matrix of $\boldsymbol{v}[n]$ is approximately given by a diagonal matrix

$$\boldsymbol{C}_{\boldsymbol{v}} = \mathrm{E}\left[\boldsymbol{v}[n]\boldsymbol{v}^{\mathrm{H}}[n]\right] = \mathrm{diag}\left(\sigma_{v,1}^2, \ldots, \sigma_{v,K}^2\right) \tag{3.84}$$

whose entries are given by $\sigma_{v,i}^2 = \tau^2/6$, $i = 2, \ldots, K$ and $\sigma_{v,1}^2 = \sigma_u^2 = 1$. The modulo operator $\mathrm{M}(\bullet)$ of the feedback loop limits the amplitude of $\boldsymbol{v}[n]$ [83, 93, 94] and, thus, the power of the transmit signal $\boldsymbol{x}[n]$.

The signal $\boldsymbol{v}[n]$ obtained from Table 3.2 is then transformed by the feedforward filter $\boldsymbol{F} \in \mathbb{C}^{N \times K}$ to get the transmit signal $\boldsymbol{x}[n] \in \mathbb{C}^N$, which must satisfy an average total transmit power constraint, i.e. $E[\|\boldsymbol{x}[n]\|_2^2] = E_{\mathrm{tr}}$. The estimated signal is expressed in matrix–vector notation as

$$\hat{\boldsymbol{d}}[n] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{v}[n] + g\boldsymbol{\eta}[n] \in \mathbb{C}^K. \tag{3.85}$$

Note that we restrict g to being common to all the users so that it acts as an automatic gain control. Then, the modulo operator is applied again at the receiver to invert the effect of the modulo operator at the transmitter [83, 93, 94]. The most appropriate interpretation of the modulo operator is that it gives the transmitter degrees of freedom, since the same output can be generated with different inputs. The THP feedback loop with the modulo operator at the transmitter is simply a suboptimal means of exploiting these degrees of freedom.

The linear representation of THP [95] is depicted in Fig. 3.12. The desired signal is denoted by $\boldsymbol{d}[n]$ and from this figure it is easy to see that

$$\boldsymbol{d}[n] = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v}[n]. \tag{3.86}$$

MU-MISO Wiener Spatial Tomlinson-Harashima Precoding (WF-THP)

The Wiener THP for flat fading channels results from the minimization of the mean square error and the restriction of a spatially causal feedback filtering. The MSE can be expressed as [11, 60, 92, 96]

$$\begin{aligned}\varepsilon_{\text{WF}}^{\text{THP}}\left(\boldsymbol{P},\boldsymbol{B},\boldsymbol{F},g\right) &= \mathrm{E}\left[\left\|\boldsymbol{d}[n]-\hat{\boldsymbol{d}}[n]\right\|_2^2\right] = \mathrm{E}\left[\left\|\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v}[n]-g\boldsymbol{H}\boldsymbol{F}\boldsymbol{v}[n]-g\boldsymbol{\eta}[n]\right\|_2^2\right] \\ &= \operatorname{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) - g^{*}\operatorname{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) \\ &\quad - g\operatorname{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) + |g|^2\operatorname{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) + |g|^2\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right).\end{aligned} \tag{3.87}$$

With this MSE, the WF-THP optimization can be expressed as

$$\begin{aligned}\left\{\boldsymbol{F}_{\text{WF}}^{\text{THP}},\boldsymbol{B}_{\text{WF}}^{\text{THP}},g_{\text{WF}}^{\text{THP}},\boldsymbol{P}_{\text{WF}}^{\text{THP}}\right\} = &\underset{\{\boldsymbol{F},\boldsymbol{B},g,\boldsymbol{P}\}}{\operatorname{argmin}}\ \varepsilon_{\text{WF}}^{\text{THP}}\left(\boldsymbol{P},\boldsymbol{B},\boldsymbol{F},g\right) \\ &\text{s.t.: } \mathrm{E}\left[\|\boldsymbol{x}\|_2^2\right] \leq E_{\text{tx}} \text{ and } \boldsymbol{B} \text{ is unit lower triangular.}\end{aligned} \tag{3.88}$$

The restriction for the unit lower triangular structure of $\boldsymbol{B}$ can be written as

$$\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i = \boldsymbol{S}_i\boldsymbol{e}_i, \quad i = 1,\ldots,K$$

where $\boldsymbol{S}_i$ is a selection matrix defined as [cf. Eq. (3.41)]

$$\boldsymbol{S}_i = \left[\boldsymbol{I}_i, \boldsymbol{0}_{i\times(K-i)}\right] \in \{0,1\}^{i\times K} \tag{3.89}$$

which cuts out the first i rows of a matrix with K rows when applied from the left. Therefore, we have K linear constraints that are defined using K Lagrangian vectors $\boldsymbol{\mu}_i \in \mathbb{C}$, $i = 1,\ldots,K$.

Then, the MSE in Eq. (3.87) enables us to construct the Lagrangian function as follows

$$\begin{aligned}L\left(\boldsymbol{P},\boldsymbol{B},\boldsymbol{F},g,\lambda,\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_K\right) = &\varepsilon_{\text{WF}}^{\text{THP}}\left(\boldsymbol{P},\boldsymbol{B},\boldsymbol{F},g\right) + \lambda\left(\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\text{tx}}\right) \\ &+ 2\Re\left(\sum_{i=1}^{K}\operatorname{tr}\left(\boldsymbol{\mu}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i - \boldsymbol{S}_i\boldsymbol{e}_i\right)\right)\right)\end{aligned} \tag{3.90}$$

with $\lambda \in \mathbb{R}^{0,+}$, $\boldsymbol{\mu}_i \in \mathbb{C}^i, i = 1,\ldots,K$, and $2\Re(\sum_{i=1}^{K}\operatorname{tr}(\boldsymbol{\mu}_i^{\mathrm{T}}(\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i - \boldsymbol{S}_i\boldsymbol{e}_i)))$ comes from the restriction for the unit lower triangular structure of the feedback matrix $\boldsymbol{B}$.

The solution of Eq. (3.87) can be obtained by setting the derivatives of the Lagrangian function with respect to $\boldsymbol{B}$, $\boldsymbol{F}$, and g to zero. Then, we have a nonconvex programming

problem and the following KKT conditions are necessary for the global solution (see Appendix C):

$$\begin{aligned}
\frac{\partial L(\bullet)}{\partial \boldsymbol{F}^*} &= -g^* \boldsymbol{H}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B} \boldsymbol{C}_{\boldsymbol{v}} + |g|^2 \boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} + \lambda \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} = \boldsymbol{0} \\
\frac{\partial L(\bullet)}{\partial \boldsymbol{B}^*} &= \boldsymbol{B} \boldsymbol{C}_{\boldsymbol{v}} - g \boldsymbol{P} \boldsymbol{H} \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} + \sum_{i=1}^{K} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \boldsymbol{e}_i^{\mathrm{T}} = \boldsymbol{0} \\
\frac{\partial L(\bullet)}{\partial g} &= -\operatorname{tr}\left(\boldsymbol{H} \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{B}^{\mathrm{H}} \boldsymbol{P}\right) + g^* \operatorname{tr}\left(\boldsymbol{H} \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}}\right) \\
&\quad + g^* \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) = 0 \\
\boldsymbol{S}_i \boldsymbol{B} \boldsymbol{e}_i &= \boldsymbol{S}_i \boldsymbol{e}_i \\
\operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}}\right) &\leq E_{\mathrm{tx}} \\
\lambda\left(\operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right) &= 0 \quad \text{with } \lambda \geq 0.
\end{aligned} \tag{3.91}$$

We first demonstrate that the transmit energy constraint in Eq. (3.88) is always active, i.e. $\lambda > 0$. To this end, the weight g^* resulting from the third KKT condition is expressed as

$$g^* = \frac{\operatorname{tr}\left(\boldsymbol{H} \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{B}^{\mathrm{H}} \boldsymbol{P}\right)}{\operatorname{tr}\left(\boldsymbol{H} \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)}. \tag{3.92}$$

If we multiply the first KKT condition from the right by $\boldsymbol{F}^{\mathrm{H}}$ and afterwards apply the trace operator we get

$$\lambda \operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}}\right) = g^* \operatorname{tr}\left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}}\right) - |g|^2 \operatorname{tr}\left(\boldsymbol{H} \boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}}\right).$$

Plugging Eq. (3.92) into the above equation, we can easily derive that

$$\lambda \operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}}\right) = |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)$$

and then, $\lambda = |g|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}}) / \operatorname{tr}(\boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}}) > 0$ if we omit the trivial solution $\boldsymbol{F} = \boldsymbol{0}$. Therefore, the transmit energy constraint is active, i.e. $\operatorname{tr}(\boldsymbol{F} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}^{\mathrm{H}}) = E_{\mathrm{tx}}$ and $\lambda = |g|^2 \xi$ with $\xi = \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}}) / E_{\mathrm{tx}}$.

Thus, the resulting feedforward filter $\boldsymbol{F}$ obtained from the first equality in Eq. (3.91) is given by

$$\boldsymbol{F} = \frac{1}{g} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} + \xi \mathbf{I}\right)^{-1} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B} = \frac{1}{g} \boldsymbol{H}^{\mathrm{H}} \left(\boldsymbol{H} \boldsymbol{H}^{\mathrm{H}} + \xi \mathbf{I}\right)^{-1} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B} \tag{3.93}$$

where we applied the matrix inversion lemma to get the last equality (see Appendix B.1). Remember that $\xi = \frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\mathrm{tx}}}$.

By plugging the above result into the second KKT condition, we obtain that

$$\begin{aligned}\frac{\partial L(\bullet)}{\partial \boldsymbol{B}^*} &= \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} - \boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} + \sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}} \\ &= \xi\boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} + \sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}} = \mathbf{0}.\end{aligned}$$

Therefore, the feedback filter $\boldsymbol{B}$ is expressed as

$$\boldsymbol{B} = -\xi^{-1}\boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}}\sigma_{v,i}^{-2} \tag{3.94}$$

where we included the assumption that the entries of $\boldsymbol{v}[n]$ are uncorrelated [see Eq. (3.84)].

Multiplying this result by $\boldsymbol{S}_i$ from the left and by $\boldsymbol{e}_i$ from the right, we have

$$\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i = -\xi^{-1}\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\sigma_{v,i}^{-2} = \boldsymbol{S}_i\boldsymbol{e}_i$$

where for the last equality the constraint for the unit lower triangular structure of $\boldsymbol{B}$ is applied and we used that $\boldsymbol{e}_j^{\mathrm{T}}\sigma_{v,j}^{-2}\boldsymbol{e}_i = 0, j \neq i$, and $\boldsymbol{e}_i^{\mathrm{T}}\sigma_{v,i}^{-2}\boldsymbol{e}_i = \sigma_{v,i}^{-2}$, otherwise. Then, the Lagrangian multipliers $\boldsymbol{\mu}_i^*$, $i = 1, \ldots, K$ are given by

$$\boldsymbol{\mu}_i^* = -\sigma_{v,i}^2\xi\left(\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i. \tag{3.95}$$

We can now substitute $\boldsymbol{\mu}_i^*$ of Eq. (3.95) into Eqs. (3.93) and (3.94) so we have the following expressions for the feedforward and feedback filters:

$$\begin{aligned}\boldsymbol{F} &= \frac{1}{g}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} \\ \boldsymbol{B} &= \boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\end{aligned} \tag{3.96}$$

respectively.

We introduce the matrix $\boldsymbol{\Phi} = (\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{I})^{-1}$ similar to [92], which is factorized by using the following symmetrically permuted Cholesky decomposition

$$\boldsymbol{P}\boldsymbol{\Phi}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{L} \tag{3.97}$$

which exists since $\boldsymbol{\Phi}$ is positive definite by definition. The matrices $\boldsymbol{L}$ and $\boldsymbol{D}$ are, respectively, unit lower triangular and diagonal.

Taking into account this factorization, the precoder filters in Eq. (3.96) can be rewritten as

$$\begin{aligned}
\boldsymbol{F} &= \frac{1}{g_{\text{WF}-\text{THP}}} \boldsymbol{H}^{\text{H}} \boldsymbol{P}^{\text{T}} \sum_{i=1}^{K} \boldsymbol{S}_i^{\text{T}} \left(\boldsymbol{S}_i \boldsymbol{L}^{-1} \boldsymbol{D}^{-1} \boldsymbol{L}^{-\text{H}} \boldsymbol{S}_i^{\text{T}} \right)^{-1} \boldsymbol{S}_i \boldsymbol{e}_i \boldsymbol{e}_i^{\text{T}} \\
&= \frac{1}{g_{\text{WF}-\text{THP}}} \boldsymbol{H}^{\text{H}} \boldsymbol{P}^{\text{T}} \sum_{i=1}^{K} \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{L}^{\text{H}} \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{D} \boldsymbol{L} \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{e}_i \boldsymbol{e}_i^{\text{T}} \\
&= \frac{1}{g_{\text{WF}-\text{THP}}} \boldsymbol{H}^{\text{H}} \boldsymbol{P}^{\text{T}} \sum_{i=1}^{K} \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{L}^{\text{H}} \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{D} \boldsymbol{L} \boldsymbol{e}_i \boldsymbol{e}_i^{\text{T}} \\
&= \frac{1}{g_{\text{WF}-\text{THP}}} \boldsymbol{H}^{\text{H}} \boldsymbol{P}^{\text{T}} \sum_{i=1}^{K} \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{L}^{\text{H}} \boldsymbol{D} \boldsymbol{e}_i \boldsymbol{e}_i^{\text{T}} = \frac{1}{g_{\text{THP}}} \boldsymbol{H}^{\text{H}} \boldsymbol{P}^{\text{T}} \boldsymbol{L}^{\text{H}} \boldsymbol{D} \\
\boldsymbol{B} &= \boldsymbol{L}^{-1} \boldsymbol{D}^{-1} \boldsymbol{L}^{-\text{H}} \boldsymbol{L}^{\text{H}} \boldsymbol{D} = \boldsymbol{L}^{-1}
\end{aligned}$$

by considering the following properties of the selection matrix $\boldsymbol{S}_i$ in Eq. (3.89)

$$\boldsymbol{S}_i \boldsymbol{M} = \boldsymbol{S}_i \boldsymbol{M} \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i, \quad \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{M} \boldsymbol{e}_i = \boldsymbol{e}_i \quad \text{and} \quad \boldsymbol{S}_i^{\text{T}} \boldsymbol{S}_i \boldsymbol{N} \boldsymbol{e}_i = \boldsymbol{N} \boldsymbol{e}_i \tag{3.98}$$

with $\boldsymbol{M}$ being a unit lower triangular matrix and with $\boldsymbol{N}$ having an upper triangular structure.

To summarize the previous derivation, the WF-THP solution to Eq. (3.88) is given by

$$\boxed{\begin{aligned}
\boldsymbol{F}_{\text{WF}}^{\text{THP}} &= \frac{1}{g_{\text{WF}}^{\text{THP}}} \boldsymbol{H}^{\text{H}} \boldsymbol{P}^{\text{T}} \boldsymbol{L}^{\text{H}} \boldsymbol{D} \\
\boldsymbol{B}_{\text{WF}}^{\text{THP}} &= \boldsymbol{L}^{-1}.
\end{aligned}} \tag{3.99}$$

The receive weight $g_{\text{WF}}^{\text{THP}}$ directly follows from the transmit energy constraint. Assuming that it is real and positive, it is given by

$$\boxed{g_{\text{WF}}^{\text{THP}} = \sqrt{\frac{\text{tr}\left(\boldsymbol{H}^{\text{H}} \boldsymbol{P}^{\text{T}} \boldsymbol{L}^{\text{H}} \boldsymbol{D}^2 \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{L} \boldsymbol{P} \boldsymbol{H}\right)}{E_{\text{tx}}}}.} \tag{3.100}$$

Plugging the above results into the cost function in Eq. (3.87), it can be demonstrated that

$$\varepsilon_{\text{WF}}^{\text{THP}}\left(\boldsymbol{P}, \boldsymbol{B}, \boldsymbol{F}, g\right) = \xi \,\text{tr}\left(\boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{D}\right) = \xi \sum_{i=1}^{K} \sigma_{v,i}^2 d_{i,i}. \tag{3.101}$$

As can be seen from Eq. (3.96), the filters are determined column by column, and each column requires one matrix inverse which results in a total complexity order of

$O(K^4)$ as proposed in [96]. For a large number of users, the filter computation becomes quite complex. Note that this high complexity is for some permutation $\boldsymbol{P}$. With the decomposition in Eq. (3.97), the complexity was reduced to $O(K^3)$. Nevertheless, all $K!$ possible user permutations must be tested to find the optimum of Eq. (3.88) with respect to $\boldsymbol{P}$. For this reason, some heuristic ordering strategies based on Eq. (3.96) were developed in [92] that include the ordering optimization in the computation of the decomposition in Eq. (3.97). Contrary to DFE, where the latter detection stages are less constrained, the precoding filter optimization for the latter precoding stages are more constrained, because after each stage, one additional transmit signal is subject to precoding. Therefore, in each iteration the algorithm finds the minimum weighted diagonal entry of $\boldsymbol{\Phi}$ (i.e. the entry corresponding to the minimum MSE) to be precoded first. So, the precoding filter optimization corresponding to that entry is computed last. Therefore, the precoding filter optimization is performed in the opposite direction to the precoding ordering (in Table 3.3, for $i = K, \ldots, 1$, compare with Table 3.1) [92]. This greedy MSE minimization leads to an ordering algorithm that achieves an excellent trade-off between performance and complexity for computing the symmetrically permuted factorization in Eq. (3.97) [11,92]. The pseudo code for the filter calculation according to Eq. (3.101) is shown in Table 3.3.

$\boldsymbol{\Phi} \leftarrow (\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \xi\mathbf{I})^{-1}$
$\boldsymbol{P} \leftarrow \mathbf{I}_K$, $\boldsymbol{D} \leftarrow \mathbf{0}_{K\times K}$
for $i = K, \ldots, 1$
  $q \leftarrow \underset{q'=1,\ldots,i}{\operatorname{argmin}} \boldsymbol{\Phi}(q', q')$
  $\boldsymbol{P}_i \leftarrow \mathbf{I}_K$ whose i-th and q-th rows are exchanged
  $\boldsymbol{P} \leftarrow \boldsymbol{P}_i\boldsymbol{P}$
  $\boldsymbol{\Phi} \leftarrow \boldsymbol{P}_i\boldsymbol{\Phi}\boldsymbol{P}_i^{\mathrm{T}}$
  $\boldsymbol{D}(i,i) \leftarrow \boldsymbol{\Phi}(i,i)$
  $\boldsymbol{\Phi}(1:i,i) \leftarrow \boldsymbol{\Phi}(1:i,i)/\boldsymbol{D}(i,i)$
  $\boldsymbol{\Phi}(1:i-1,1:i-1) \leftarrow \boldsymbol{\Phi}(1:i-1,1:i-1)$
    $-\boldsymbol{\Phi}(1:i-1,i)\boldsymbol{\Phi}(1:i-1,i)^{\mathrm{H}}\boldsymbol{D}(i,i)$
$\boldsymbol{L}^{\mathrm{H}} \leftarrow$ upper triangular part of $\boldsymbol{\Phi}$
$\boldsymbol{B} \leftarrow \boldsymbol{L}^{-1}$, $\boldsymbol{F} \leftarrow \boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}$

Table 3.3: Calculation of THP Filter with Optimum Ordering.

MU-MISO Zero-Forcing Spatial Tomlinson-Harashima Precoding (ZF-THP)

The MSE optimization for ZF-THP can be expressed as (cf. [11, 60, 92, 96])

$$\begin{aligned}\left\{\boldsymbol{F}_{\text{ZF}}^{\text{THP}}, \boldsymbol{B}_{\text{ZF}}^{\text{THP}}, g_{\text{ZF}}^{\text{THP}}, \boldsymbol{P}_{\text{ZF}}^{\text{THP}}\right\} = \underset{\{\boldsymbol{F},\boldsymbol{B},g,\boldsymbol{P}\}}{\operatorname{argmin}} \; & \mathrm{E}\left[\left\|\boldsymbol{d}[n] - \hat{\boldsymbol{d}}[n]\right\|_2^2\right] \\ \text{s.t.:} \quad & \mathrm{E}\left[\|\boldsymbol{x}\|_2^2\right] \leq E_{\text{tx}} \quad \text{and} \quad \mathrm{E}[\hat{\boldsymbol{d}}[n] \mid \boldsymbol{d}[n]] = \boldsymbol{d}[n] \\ & \text{with } \boldsymbol{B} \text{ unit lower triangular}\end{aligned} \tag{3.102}$$

where $\hat{\boldsymbol{d}}[n]$ and $\boldsymbol{d}[n]$ are given by Eqs. (3.85) and (3.86), respectively. Due to the zero–forcing constraint

$$\mathrm{E}[\hat{\boldsymbol{d}}[n] \mid \boldsymbol{d}[n]] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{v}[n] = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v}[n] = \boldsymbol{d}[n]$$

which implies that $g\boldsymbol{H}\boldsymbol{F} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}$, the MSE reduces to

$$\mathrm{E}\left[\left\|\boldsymbol{d}[n] - \hat{\boldsymbol{d}}[n]\right\|_2^2\right] = |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right).$$

We construct the Lagrangian function corresponding to Eq. (3.102):

$$\begin{aligned}L\left(\boldsymbol{F}, \boldsymbol{B}, g, \boldsymbol{P}, \lambda, \boldsymbol{\Lambda}, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K\right) &= |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) + 2\Re\left(\operatorname{tr}\left(\boldsymbol{\Lambda}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B} - g\boldsymbol{H}\boldsymbol{F}\right)\right)\right) \\ &+ \lambda\left(\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}} - E_{\text{tx}}\right)\right) + 2\Re\left(\sum_{i=1}^{K} \operatorname{tr}\left(\boldsymbol{\mu}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i - \boldsymbol{S}_i\boldsymbol{e}_i\right)\right)\right)\end{aligned} \tag{3.103}$$

where $\lambda \in \mathbb{R}^{0,+}$, $\boldsymbol{\Lambda} \in \mathbb{C}^{K\times K}$, and $\boldsymbol{\mu}_i \in \mathbb{C}^i$, $i = 1, \ldots, K$. The selection matrix $\boldsymbol{S}_i$ has already been defined in Eq. (3.89).

By setting the derivatives with respect to $\boldsymbol{F}$, $\boldsymbol{B}$, and g to zero, we get

$$\begin{aligned}\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{F}^*} &= \lambda\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}} - g^*\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Lambda}^{\mathrm{H}} = \boldsymbol{0} \\ \frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{B}^*} &= \sum_{i=1}^{K} \boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}} + \boldsymbol{P}\boldsymbol{\Lambda}^{\mathrm{H}} = \boldsymbol{0} \\ \frac{\partial L\left(\bullet\right)}{\partial g} &= g^* \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) - \operatorname{tr}\left(\boldsymbol{\Lambda}\boldsymbol{H}\boldsymbol{F}\right) = 0 \\ \boldsymbol{P}^{\mathrm{T}}\boldsymbol{B} &= g\boldsymbol{H}\boldsymbol{F} \\ \boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i &= \boldsymbol{S}_i\boldsymbol{e}_i \\ \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\right) &\leq E_{\text{tx}} \\ \lambda\left(\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}} - E_{\text{tx}}\right)\right) &= 0 \quad \text{with } \lambda \geq 0.\end{aligned} \tag{3.104}$$

The above equations are the necessary KKT conditions for the global solution to the nonconvex programming problem of Eq. (3.102).

The third KKT condition directly leads to

$$g^* = \frac{\mathrm{tr}\left(\boldsymbol{\Lambda H F}\right)}{\mathrm{tr}\left(\boldsymbol{C_\eta}\right)}$$

and plugging this result into the first KKT condition yields

$$\lambda \boldsymbol{F C_v} = \frac{\mathrm{tr}^*\left(\boldsymbol{\Lambda H F}\right)}{\mathrm{tr}\left(\boldsymbol{C_\eta}\right)} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{\Lambda}^{\mathrm{H}}. \tag{3.105}$$

Multiplying now the above equation from the right by $\boldsymbol{F}^{\mathrm{H}}$ and applying the trace operator we have

$$\lambda\, \mathrm{tr}\left(\boldsymbol{F C_v F}^{\mathrm{H}}\right) = \frac{\left|\mathrm{tr}\left(\boldsymbol{\Lambda H F}\right)\right|^2}{\mathrm{tr}\left(\boldsymbol{C_\eta}\right)}$$

where we are taking into account that $\mathrm{tr}(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{\Lambda}^{\mathrm{H}} \boldsymbol{F}^{\mathrm{H}}) = \mathrm{tr}^*(\boldsymbol{F \Lambda H})$ (see Appendix B.2). Therefore, $\lambda > 0$ and the transmit energy constraint is always active if we omit the trivial solution $\boldsymbol{F} = \boldsymbol{0}$.

From the second KKT condition in Eq. (3.104) we obtain the following,

$$\boldsymbol{\Lambda}^{\mathrm{H}} = -\boldsymbol{P}^{\mathrm{T}} \sum_{i=1}^{K} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \boldsymbol{e}_i^{\mathrm{T}}. \tag{3.106}$$

When we plug this result into the first KKT condition we obtain that

$$\boldsymbol{F} = -\frac{g^*}{\lambda} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{T}} \left(\sum_{i=1}^{K} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \boldsymbol{e}_i^{\mathrm{T}} \right) \boldsymbol{C_v}^{-1}.$$

With the zero–forcing constraint, this result shows that

$$\boldsymbol{B} = -\frac{|g|^2}{\lambda} \boldsymbol{P H H}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{T}} \sum_{i=1}^{K} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \boldsymbol{e}_i^{\mathrm{T}} \sigma_{v,i}^{-2}$$

due to Eq. (3.84). Substituting this expression for the feedback filter into the constraint for unit lower triangularity, we get

$$-\frac{|g|^2}{\lambda} \boldsymbol{S}_i \boldsymbol{P H H}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \sigma_{v,i}^{-2} = \boldsymbol{S}_i \boldsymbol{e}_i$$

and consequently,

$$\boldsymbol{\mu}_i^* = -\frac{\lambda}{|g|^2} \sigma_{v,i}^2 \left(\boldsymbol{S}_i \boldsymbol{P H H}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{S}_i^{\mathrm{T}} \right)^{-1} \boldsymbol{S}_i \boldsymbol{e}_i.$$

This result for $\boldsymbol{\mu}_i^*$ leads to the optimal feedforward and feedback filter

$$\begin{aligned}\boldsymbol{F} &= \frac{1}{g}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} \\ \boldsymbol{B} &= \boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\end{aligned} \tag{3.107}$$

respectively.

The receive gain factor g is directly obtained from the transmit energy constraint, i.e.

$$g = \sqrt{\frac{\sum_{i=1}^{K}\mathrm{tr}\left(\sigma_{v,i}^2\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\right)}{E_{\mathrm{tx}}}}$$

where we consider that g is positive and real to guarantee a unique solution.

The Cholesky factorization with symmetric permutation of $\boldsymbol{\Phi} = \boldsymbol{P}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}})^{-1}\boldsymbol{P}^{\mathrm{T}}$ can be computed as [11, 92]

$$\boldsymbol{P}\boldsymbol{\Phi}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{P}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{L} \tag{3.108}$$

where $\boldsymbol{L}$ and $\boldsymbol{D}$ are a unit lower triangular and a diagonal matrix, respectively. This factorization leads to an algorithm achieving a trade-off between performance and complexity, as shown in [92]. With the factorization in Eq. (3.108), the feedback and feedforward filters are obtained similarly to the WF-THP formulation as

$$\boxed{\begin{aligned}\boldsymbol{F}_{\mathrm{ZF}}^{\mathrm{THP}} &= \frac{1}{g_{\mathrm{THP}}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D} \\ \boldsymbol{B}_{\mathrm{ZF}}^{\mathrm{THP}} &= \boldsymbol{L}^{-1}\end{aligned}} \tag{3.109}$$

using the properties for the selection matrix $\boldsymbol{S}_i$ pointed out in Eq. (3.89). The common gain $g_{\mathrm{ZF}}^{\mathrm{THP}}$ is directly obtained from the transmit energy constraint assuming that it is real and positive as follows

$$\boxed{g_{\mathrm{ZF}}^{\mathrm{THP}} = \sqrt{\frac{\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{D}\right)}{E_{\mathrm{tx}}}}.} \tag{3.110}$$

This leads to the MSE in Eq. (3.102)

$$\varepsilon_{\mathrm{ZF}}^{\mathrm{THP}} = \xi\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{D}\right) = \xi\sum_{i=1}^{K}\sigma_{v,i}^2 d_i. \tag{3.111}$$

Note that this MSE is not equal to the MSE for WF-THP since the matrix $\boldsymbol{D}$ is obtained via Eq. (3.108) [for WF–THP, we use Eq. (3.97)]. WF-THP converges to ZF-THP for

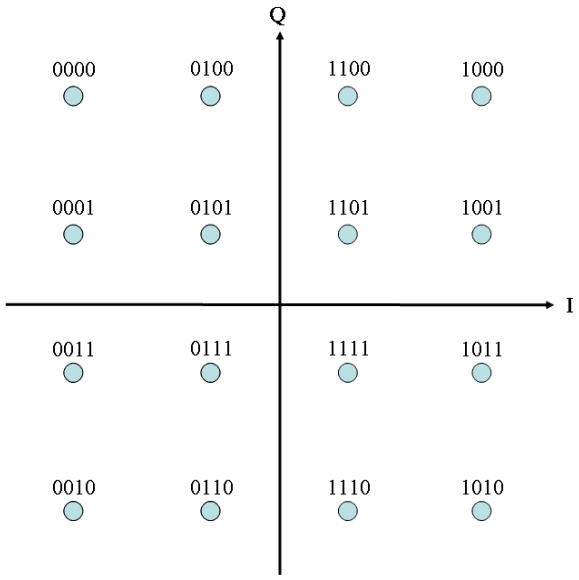


Figure 3.14: 16QAM Constellation.

high SNR scenarios, i.e. for $\xi \to 0$, since the respective matrices $\boldsymbol{\Phi}$ converge. As a result of this MSE, the ordering algorithm to compute the filters is similar to the algorithm explained for WF-THP, but changing the initial matrix in Table 3.3 to $\boldsymbol{\Phi} = (\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}})^{-1}$.

3.3.3 Simulation Results

In this subsection we present the results of some computer simulations carried out to illustrate the performance of the previously described nonlinear transmit processing techniques. We consider a MU-MISO system with $N = 4$ transmit antennas and $K = 4$ receiving users. A frame length of $N_{\mathrm{B}} = 50$ symbols is considered and $5{,}000$ channel realizations are averaged. We assume that the transmitted symbols are either QPSK or 16QAM modulated. The QPSK modulation constellation is given as $\mathbb{A} = \{\pm\sqrt{2}/2 \pm \mathrm{j}\sqrt{2}/2\}$ (see Fig. 3.4) and the alphabet in the case of 16QAM is $\mathbb{A} = \{\pm 3 \pm \mathrm{j}\,3, \pm 3 \pm \mathrm{j}, \pm 1 \pm \mathrm{j}\,3, \pm 1 \pm \mathrm{j}\}$ normalized by the factor $1/\sqrt{10}$ (see Fig. 3.14).

Fig. 3.15 depicts the BER performance vs. SNR for the zero-forcing nonlinear approaches: MLD, DF equalizer, TH precoding, and vector precoding. As can be seen from the figure, there is some advantage for the receive DF filter at low SNR and for the transmit filters THP and VP at high SNR. Transmit processing suffers from some loss at low SNR due to the modulo operator at the receiver since the modulo operation introduces more allowed constellation points. Since the number of constellation points in 16QAM is larger than in QPSK, the impact of the modulo operator is not so pronounced for 16QAM. Moreover, given that this effect depends on the noise, it will be more important for low

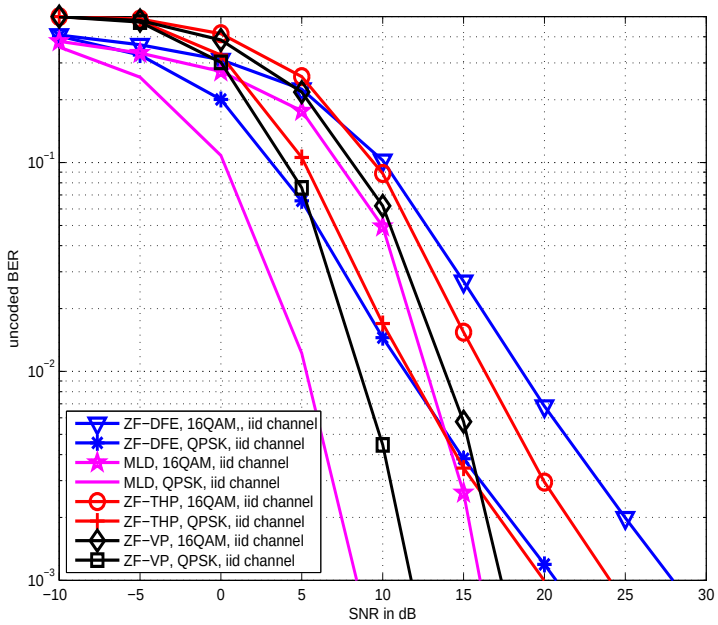


Figure 3.15: Uncoded BER vs. SNR for Nonlinear ZF Transmit and Receive Filters: QPSK and 16QAM Transmission over Uncorrelated Flat Fading MU-MISO Channels with Four Transmitting Antenna Elements and Four Users.

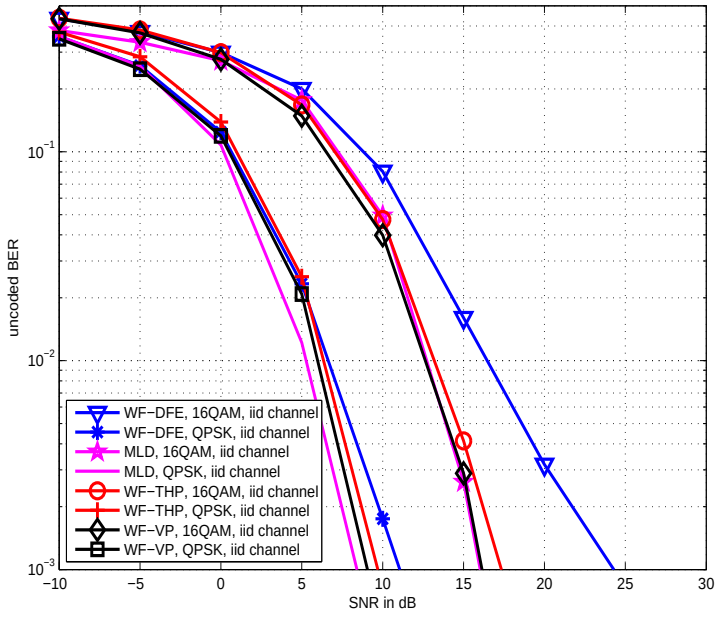


Figure 3.16: Uncoded BER vs. SNR for Nonlinear WF Transmit and Receive Filters: QPSK and 16QAM Transmission over Uncorrelated Flat Fading MU-MISO Channels with Four Transmitting Antenna Elements and Four Users.

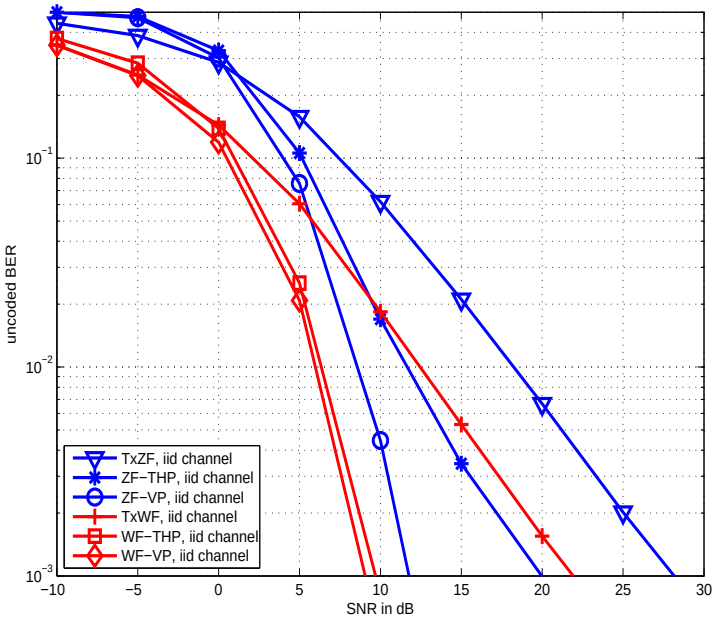


Figure 3.17: Uncoded BER vs. SNR for Transmit Filters: QPSK Transmission over Uncorrelated Flat Fading MU-MISO Channels with Four Transmitting Antenna Elements and Four Users.

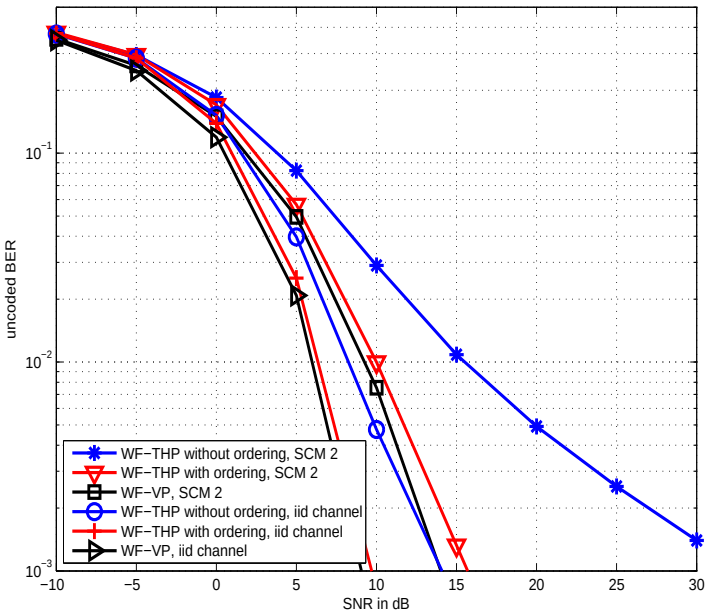


Figure 3.18: Uncoded BER vs. SNR for Nonlinear Transmit Filters: QPSK Transmission over Uncorrelated and Correlated (*SCM 2*) Flat Fading MU-MISO Channels with Four Transmitting Antenna Elements and Four Users.

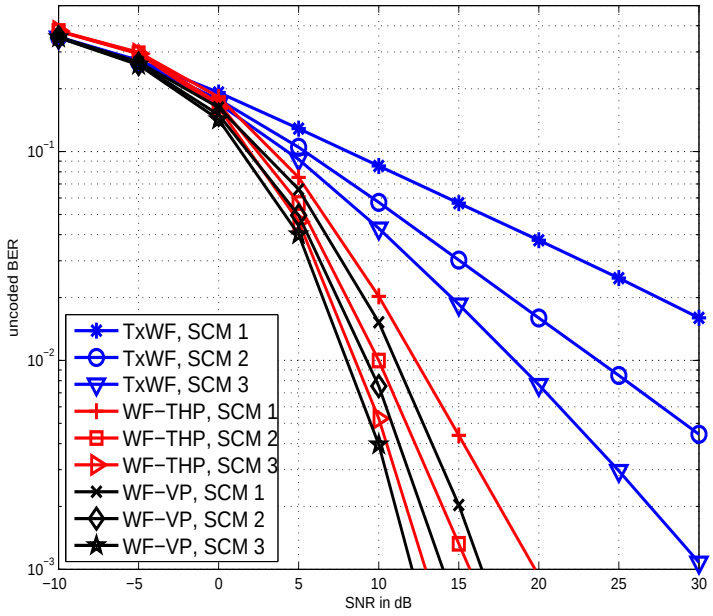


Figure 3.19: Uncoded BER vs. SNR for Transmit Filters: QPSK Transmission over Correlated Flat Fading MU-MISO Channels with Four Transmitting Antenna Elements and Four Users.

SNR. For high SNR scenarios, however, transmit processing shows better performance than DFE since error propagation is smaller. As expected, the performance achieved with Wiener filters instead of ZF filters is much better because zero-forcing amplifies the channel noise, as also shown in Fig. 3.16. From this figure the same conclusions as before can be extracted, but at lower SNR values for the same BERs. Figs. 3.15 and 3.16 also show the BER performance obtained when we directly perform ML detection at the receivers. As expected, MLD is the best technique for separating the signals from the different users, in spite of its lower performance for low scenarios with respect to the vector precoding approach caused by the effect of VP modulo operator at the receivers.

Fig. 3.17 shows a comparison between ZF and Wiener linear and nonlinear transmit processing for QPSK transmission. ZF approaches are clearly outperformed by the respective Wiener approaches. A small loss in performance is observed for the nonlinear approaches THP and VP for very low SNR with respect to linear transmit processing due to the modulo operators at the receiver. However, both THP and VP show an important gain for medium and high SNR scenarios.

Fig. 3.18 represents the uncoded BER vs. SNR for different nonlinear transmit processing approaches. For uncorrelated channels, there is a gain of about $1.5\,\mathrm{dB}$ when comparing ordered and non-ordered THP schemes at the practical operation point in

coded transmission corresponding to BER=10^{-2}. This gain is even greater for correlated channels such as *SCM 2*, where the difference is larger than 5 dB for this point. However, the ordering optimization shows a slight performance degradation in the low uncoded BER region, although this loss is negligible. The enormous advantage of this suboptimum ordering solution proposed in [92] is in fact its complexity, which is comparable to linear zero-forcing or WF filters.

Finally, Fig. 3.19 plots the BER performance achieved for the three spatial channel models considered throughout this work. SCM 1, corresponding to the suburban macrocell environment, shows the worst performance due to its small diversity since spatial correlations are larger than in the other two types of channels. On the other hand, SCM 3, corresponding to an urban microcell environment, obviously shows the best performance since all the eigenvalues of its covariance matrix are not negligible, i.e. SCM 3 is the most spatially uncorrelated channel compared to the others. The channel correlations are affected by the angle spread at the base station, which is larger in microcell than in macrocell environments [37]. The same is true when comparing urban and suburban areas, since the scattering process in the vicinity of the base station is increased due to its location usually being at the same height as the surrounding scatterers in urban areas [36,37]. For all the types of channels, nonlinear transmit techniques exhibit better performance than linear precoding with VP being the best solution at the cost of increasing the computational complexity at the transmitter caused by the lattice search. However, at low SNR, linear precoding outperforms the other two nonlinear schemes due to the presence of the modulo operators at the receiver sides for both THP and VP.

3.4 Conclusions

In this chapter we have reviewed most of the commonly implemented approaches for transmit and receive processing in multiuser MISO and SIMO systems. In general, we have shown that nonlinear schemes lead to better performance in terms of BER. Moreover, transmit processing shows better performance compared to receive processing for high SNR scenarios, with the enormous advantage of simplifying the receiver side of the downlink. Although MLD and VP achieve the full diversity of the channel they typically suffer from much higher complexity. For this reason, given the superiority of nonlinear schemes over linear ones and the need to perform the signal separation at the transmitter in the downlink due to the absence of cooperation between the different users, THP is the best choice as a trade-off between performance and complexity.

Chapter 4

Imperfect CSI: Error Sources

Most recent standards in wireless communication systems include feedback channels for sending information relating to different link parameters from the users to be used by the transmitter. However, the data rate of these feedback channels is naturally limited (e.g. [25]). When the transmitter has no full knowledge about the channel necessary to construct the precoding filters, this information has to be supplied by the users. Therefore, this CSI must be compressed to ensure that the tight scheduling constraints on the limited date rate of the feedback channel are satisfied. It is still a subject of research to determine what kind of information has to be sent from the receivers to the transmitter and the way of recovering it at the transmitter side. Therefore, the main contribution of this work is the design of this limited CSI feedback to minimize the mean squared error with the minimum number of bits of fed–back information. This topic is henceforth called *limited feedback*.

Erroneous CSI at the transmitter involves a performance degradation due to the mismatch between the true channel and the erroneous channel used for the design of the precoder filters. Therefore, we have to incorporate a robust design of these filters to compensate this mismatch effect, which is termed *robust precoding* [17, 43, 44]. In the next chapter we derive the design of the transmit processing schemes studied in this work to be robust against errors occurring on the channel information available at the transmitter. Our ultimate objective will be the design of the limited feedback channel by taking into account the CSI errors that are introduced below in this chapter.

Fig. 4.1 plots the block diagram of the limited feedback channel that we will be assuming throughout this work. The chapter starts by estimating the channel at the receivers using pilot symbols sent from all the transmit antennas. This enables the receivers to estimate their respective vector channels. Then, we reduce the estimates to a low-dimensional representation, which is also called *rank reduction* or *truncation*, by projecting them onto a basis. This basis depends only on the channel statistics and the projection leads to d coefficients per user, with $d \leq N$. It is assumed that the channel

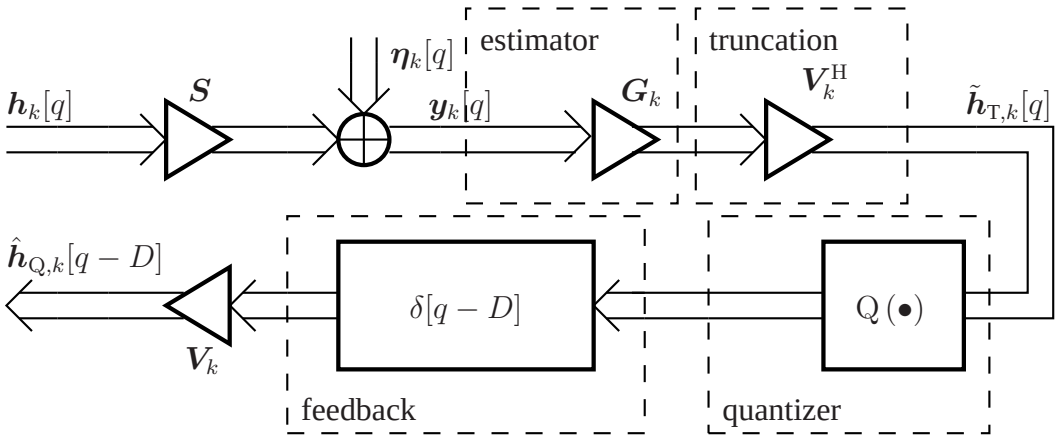


Figure 4.1: Model of Limited Feedback: Channel Estimation, Truncation, Quantization, and Feedback Delay.

statistics are known at both the transmitter and the receiver side. The d coefficients are quantized prior to transmission over the feedback channel which also introduces a delay. However, we assume that the feedback channel does not suffer from transmission errors.

For brevity, the errors due to the estimation process will be termed *Type A errors*; the errors including estimation and rank reduction will be called *Type B errors*; and finally, all the errors without feedback delay will be termed *Type C errors*. We describe all the error sources in the following sections.

4.1 Channel Estimation

Fig. 4.2 depicts the estimation process performed at the receiver side. We use linear estimators at the receivers based on $N_{\mathrm{tr}} \times N$ pilot symbols per user to enable the channel vector estimation for the k-th user. The vector comprising the N_{tr} received symbols for the k–th user at the time slot q reads as

$$\boldsymbol{y}_k[q] = \boldsymbol{S}\boldsymbol{h}_k[q] + \boldsymbol{\eta}_k[q] \in \mathbb{C}^{N_{\mathrm{tr}}} \tag{4.1}$$

where the matrix $\boldsymbol{S}$ contains the training symbols and is given by

$$\boldsymbol{S} = \begin{pmatrix} S_{1,1} & \dots & S_{1,N} \\ \vdots & \ddots & \vdots \\ S_{N_{\mathrm{tr}},1} & \dots & S_{N_{\mathrm{tr}},N} \end{pmatrix} \in \mathbb{C}^{N_{\mathrm{tr}} \times N}$$

with $\boldsymbol{h}_k[q] = [h_{1,k}[q], \dots, h_{N,k}[q]]^{\mathrm{T}} \in \mathbb{C}^N$ corresponding to the channel vector for user k and $\boldsymbol{y}_k[q] = [y_{1,k}[q], \dots, y_{N_{\mathrm{tr}},k}[q]]^{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{tr}}}$ and $\boldsymbol{\eta}_k[q] = [\eta_{1,k}[q], \dots, \eta_{N_{\mathrm{tr}},k}[q]]^{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{tr}}}$

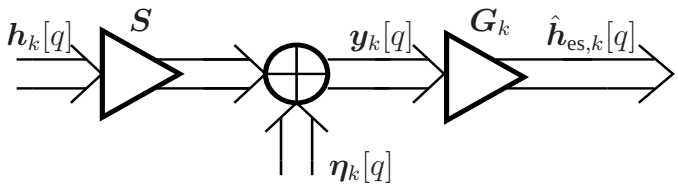


Figure 4.2: Type A Error: Estimation.

being, respectively, the received pilot signal and the additive white Gaussian noise with covariance matrix $\boldsymbol{C}_{\boldsymbol{\eta}_k} = \mathrm{E}[\boldsymbol{\eta}_k[q]\boldsymbol{\eta}_k^{\mathrm{H}}[q]]$. Remember that we have a fading block channel with q denoting the time slot of size N_{B} symbols spaced with T_{s}. As shown in Fig. 4.2, the above received signal $\boldsymbol{y}_k[q]$ is passed through a linear filter $\boldsymbol{G}_k \in \mathbb{C}^{N \times N_{\mathrm{tr}}}$ to produce the linear channel estimate

$$\hat{\boldsymbol{h}}_{\text{es},k}[q] = \boldsymbol{G}_k \boldsymbol{y}_k[q] \in \mathbb{C}^N. \tag{4.2}$$

As mentioned before, it is assumed that the channel statistics are known at the transmitter and the receiver side. If not, the information from the training channel can be exploited to estimate the channel and noise covariance matrices. These covariance matrices can be communicated to the transmitter without significantly increasing the overhead of the feedback channel since these statistics change very slowly. For example, when a MIMO testbed is employed (see Chapter 8), it is apparent that we need to know the channel statistics and, therefore, it is mandatory to use some method to estimate these second order moments [97].

Note that the error due to estimation cannot be neglected even though the number of pilot symbols is increased, as we illustrate in Chapter 6. Therefore, perfect CSI is not even available at the receivers and erroneous CSI is sent to the transmitter. Supervised techniques for channel tracking, such as Kalman filtering [98–102], were employed as initial trials to illustrate the effect of that erroneous CSI caused by estimation on the final performance. Also some blind techniques, as explained in [100–103] or in [104, 105] where particle filtering or adaptive blind source separation algorithms such as *Equivariant Adaptive Separation via Independence* (EASI) were implemented, have been tested to follow the channel variations by means of estimating and predicting the channel coefficients at the cost of increasing the loss in performance compared to the previous non-blind methods.

We describe in the following subsections the channel estimation methods used throughout this work: the *Least-Squares* (LS) and the *Linear Minimum Mean Square Error* (LMMSE) channel estimators. The LS estimation is based on the minimization of the sum of the squared errors between the observations, $\boldsymbol{y}_k[q]$, and the desired signal $\boldsymbol{S}\boldsymbol{h}_k[q]$. However, for MMSE estimation, the channel estimate is found such that the

estimate $\boldsymbol{G}_k\boldsymbol{y}_k[q]$ is the most similar signal to the channel $\boldsymbol{h}_k[q]$ in the MMSE sense. There exists a strong connection between the LS and MMSE criteria. The MMSE filter, which is also called *Wiener* filter, is obtained by minimizing the mean square error, and is therefore a stochastic criterion. The LS solution, however, is obtained by minimizing the squared error on a given block of data, which constitutes a deterministic problem. But when the noise is white, i.e. $\boldsymbol{C}_{\boldsymbol{\eta},k} = \sigma^2_{\boldsymbol{\eta},k}\mathbf{I}$, and the noise variance converges to zero, the MMSE solution converges to the LS solution. Therefore, LS estimation can be considered as a practical alternative to the MMSE method, since no statistics of any signal must be available to design the estimator.

4.1.1 Least-Squares (LS) Channel Estimation

As mentioned above, the LS channel estimator minimizes the following squared error quantity

$$\hat{\boldsymbol{h}}^{(\mathrm{LS})}_{\mathrm{es},k}[q] = \underset{\boldsymbol{h}_k[q]}{\operatorname{argmin}} \|\boldsymbol{y}_k[q] - \boldsymbol{S}\boldsymbol{h}_k[q]\|_2^2 \tag{4.3}$$

where $\boldsymbol{y}_k[q]$ is the received signal in Eq. (4.1).

The expression for the gradient of the squared error $\varepsilon = \|\boldsymbol{y}_k[q] - \boldsymbol{S}\boldsymbol{h}_k[q]\|_2^2$ with respect to $\boldsymbol{h}_k^*[q]$ is

$$\frac{\partial \varepsilon}{\partial \boldsymbol{h}_k^*[q]} = -\boldsymbol{S}^{\mathrm{H}}\boldsymbol{y}_k[q] + \boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}\boldsymbol{h}_k[q].$$

As the second gradient $\frac{\partial^2 \varepsilon}{\partial \boldsymbol{h}_k^*[q]\partial \boldsymbol{h}_k^{\mathrm{T}}[q]}$ [1] with respect to $\boldsymbol{h}_k^{\mathrm{T}}[q]$ is given by $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$, which is positive definite, the cost function of Eq. (4.3) is strictly convex. Therefore, there exists a unique global minimum solution to Eq. (4.3) that can be obtained via the above gradient.

The least-squares estimate sets this gradient to zero to produce the solution [97, 106, 107]

$$\hat{\boldsymbol{h}}^{(\mathrm{LS})}_{\mathrm{es},k}[q] = \boldsymbol{G}_{\text{LS-estim},k}\boldsymbol{y}_k[q] = \left(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}\right)^{-1}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{y}_k[q] = \boldsymbol{S}^{\dagger}\boldsymbol{y}_k[q] \tag{4.4}$$

where the last equality is obtained from the pseudo inverse definition of $\boldsymbol{S}$, denoted as $\boldsymbol{S}^{\dagger} = (\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S})^{-1}\boldsymbol{S}^{\mathrm{H}}$. $\boldsymbol{G}_{\text{LS-estim},k}$ is implicitly defined in Eq. (4.4) and is given by

$$\boldsymbol{G}_{\text{LS-estim},k} = \left(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}\right)^{-1}\boldsymbol{S}^{\mathrm{H}}. \tag{4.5}$$

Note that the LS estimation matrix $\boldsymbol{S}^{\dagger}$ is common to all the users. Also note that the expression $(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S})^{-1}\boldsymbol{S}^{\mathrm{H}}$ for the Moore–Penrose pseudo inverse $\boldsymbol{S}^{\dagger}$ is only valid if $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$

[1] More precisely, the definiteness of the Hessian matrix $\frac{\partial^2 \varepsilon}{\partial \begin{bmatrix}\boldsymbol{h}_k^*[q]\\ \boldsymbol{h}_k[q]\end{bmatrix} \partial[\boldsymbol{h}_k^{\mathrm{T}}[q], \boldsymbol{h}_k^{\mathrm{H}}[q]]}$ has to be checked. Since the Hessian matrix is $\mathbf{I} \otimes \boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$, and thus positive definite, the error ε is convex.

is regular. Under the assumption of linear independent columns of $\boldsymbol{S}$, this leads to the condition $N_{\text{tr}} \geq N$. The matrix $\boldsymbol{S}^{\text{H}}\boldsymbol{S}$ is called the *Grammian* or *Gram* matrix of $\boldsymbol{S}$ [107].

The error covariance matrix derived from LS channel estimation is easily obtained as follows

$$\begin{aligned}\boldsymbol{C}_{\boldsymbol{e},k}^{(\text{LS})} &= \text{E}\left[\left(\boldsymbol{h}_k[q] - \boldsymbol{G}_{\text{LS-estim},k}\left(\boldsymbol{S}\boldsymbol{h}_k[q] + \boldsymbol{\eta}_k[q]\right)\right)\left(\boldsymbol{h}_k[q] - \boldsymbol{G}_{\text{LS-estim},k}\left(\boldsymbol{S}\boldsymbol{h}_k[q] + \boldsymbol{\eta}_k[q]\right)\right)^{\text{H}}\right] \\ &= \left(\boldsymbol{S}^{\text{H}}\boldsymbol{S}\right)^{-1}\boldsymbol{S}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta},k}\boldsymbol{S}\left(\boldsymbol{S}^{\text{H}}\boldsymbol{S}\right)^{-1}. \end{aligned} \tag{4.6}$$

4.1.2 Linear Minimum Mean Squared Error (LMMSE) Channel Estimation

Let us assume that the estimator of $\boldsymbol{h}_k[q]$ is constrained to be a linear function of $\boldsymbol{y}_k[q]$. The problem is to find the matrix $\boldsymbol{G}_k$ that minimizes the mean squared error between $\boldsymbol{h}_k[q]$ and the linear estimate $\hat{\boldsymbol{h}}_{\text{es},k} = \boldsymbol{G}_k\boldsymbol{y}_k[q]$. The mean-squared error between $\boldsymbol{h}_k[q]$ and $\hat{\boldsymbol{h}}_{\text{es},k}$ is given by [83, 107],

$$\begin{aligned}\varepsilon\left(\boldsymbol{G}_k\right) &= \text{E}\left[\|\boldsymbol{h}_k[q] - \boldsymbol{G}_k\boldsymbol{y}_k[q]\|_2^2\right] = \text{E}\left[\text{tr}\left(\left(\boldsymbol{h}_k[q] - \boldsymbol{G}_k\boldsymbol{y}_k[q]\right)\left(\boldsymbol{h}_k[q] - \boldsymbol{G}_k\boldsymbol{y}_k[q]\right)^{\text{H}}\right)\right] \\ &= \text{E}\left[\text{tr}\left(\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\text{H}}[q] - \boldsymbol{h}_k[q]\boldsymbol{y}_k^{\text{H}}[q]\boldsymbol{G}_k^{\text{H}} - \boldsymbol{G}_k\boldsymbol{y}_k[q]\boldsymbol{h}_k^{\text{H}}[q] + \boldsymbol{G}_k\boldsymbol{y}_k[q]\boldsymbol{y}_k^{\text{H}}[q]\boldsymbol{G}_k^{\text{H}}\right)\right] \\ &= \text{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{C}_{\boldsymbol{h}\boldsymbol{y},k}\boldsymbol{G}_k^{\text{H}} - \boldsymbol{G}_k\boldsymbol{C}_{\boldsymbol{y}\boldsymbol{h},k} + \boldsymbol{G}_k\boldsymbol{C}_{\boldsymbol{y},k}\boldsymbol{G}_k^{\text{H}}\right) \end{aligned} \tag{4.7}$$

where the covariance matrices $\boldsymbol{C}_{\boldsymbol{h},k}$, $\boldsymbol{C}_{\boldsymbol{h}\boldsymbol{y},k}$, $\boldsymbol{C}_{\boldsymbol{y}\boldsymbol{h},k}$, and $\boldsymbol{C}_{\boldsymbol{y},k}$ are implicitly defined. Notice that the MSE is strictly convex in $\boldsymbol{G}_k$. Therefore, its solution is unique.

By setting the derivative of $\varepsilon(\boldsymbol{G}_k)$ with respect to $\boldsymbol{G}_k^*$ to zero, we obtain

$$\frac{\partial\varepsilon\left(\boldsymbol{G}_k\right)}{\partial\boldsymbol{G}_k^*} = -\boldsymbol{C}_{\boldsymbol{h}\boldsymbol{y},k} + \boldsymbol{G}_k\boldsymbol{C}_{\boldsymbol{y},k} = \boldsymbol{0}$$

which leads to the final expression for the MMSE linear filter

$$\boldsymbol{G}_{\text{MMSE-estim},k} = \boldsymbol{C}_{\boldsymbol{h}\boldsymbol{y},k}\boldsymbol{C}_{\boldsymbol{y},k}^{-1}. \tag{4.8}$$

Bearing in mind Eq. (4.1), the resulting LMMSE filter is expressed as

$$\boldsymbol{G}_{\text{MMSE-estim},k} = \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}}(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1} \in \mathbb{C}^{N\times N_{\text{tr}}} \tag{4.9}$$

by incorporating $\boldsymbol{C}_{\boldsymbol{h}\boldsymbol{y},k} = \text{E}[\boldsymbol{h}_k[q]\boldsymbol{y}_k^{\text{H}}[q]] = \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}}$ and $\boldsymbol{C}_{\boldsymbol{y},k} = \text{E}[\boldsymbol{y}_k[q]\boldsymbol{y}_k^{\text{H}}[q]] = \boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}$ into Eq. (4.8).

Thus, the MMSE channel estimate is obtained as

$$\hat{\boldsymbol{h}}_{\text{es},k}^{(\text{MMSE})}[q] = \boldsymbol{G}_{\text{MMSE-estim},k}\boldsymbol{y}_k[q]. \tag{4.10}$$

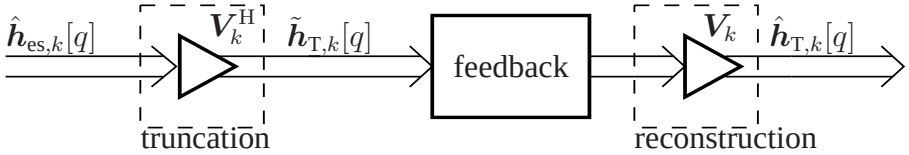


Figure 4.3: Rank reduction: Truncation at the Receiver and Reconstruction at the Transmitter.

Note that the MMSE solution (after applying the inversion lemma) reduces to the LS approach for white noise ($\boldsymbol{C}_{\boldsymbol{\eta},k} = \sigma^2_{\boldsymbol{\eta},k}\mathbf{I}$) and low noise scenarios ($\sigma^2_{\boldsymbol{\eta},k} \to 0$). Therefore, the LS approach will be clearly suboptimum for the case of correlated noise.

Finally, it is easy to demonstrate using the inversion lemma (Appendix B) that the error associated with the linear estimator $\boldsymbol{G}_{\text{MMSE-estim},k}\boldsymbol{y}_k[q]$ has the following covariance matrix

$$\begin{aligned}\boldsymbol{C}_{\boldsymbol{e},k}^{(\text{MMSE})} &= \text{E}\left[\|\boldsymbol{h}_k[q] - \boldsymbol{G}_{\text{MMSE-estim},k}\boldsymbol{S}\boldsymbol{h}_k[q] - \boldsymbol{G}_{\text{MMSE-estim},k}\boldsymbol{\eta}_k[q]\|_2^2\right] \\ &= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}}\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1}\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k} = \left(\boldsymbol{C}_{\boldsymbol{h},k}^{-1} + \boldsymbol{S}^{\text{H}}\boldsymbol{C}_{\boldsymbol{\eta},k}^{-1}\boldsymbol{S}\right)^{-1}.\end{aligned} \tag{4.11}$$

4.2 Rank Reduction

In wireless communications, the feedback channel is often limited in terms of data rate. It is interesting, then, to compress the CSI in order to reduce the amount of information sent from the users to the transmitter. To this end, in this section we explain how the CSI estimation can be compressed by truncating the *Karhunen-Loève* (KL) transformation. The basic premise of the truncated KL transform is that it is optimum in the sense that it provides dimensionality reduction with the smallest possible MSE. The robust designs proposed in this work are based on this decomposition, although the starting point matrix may be different from the matrix used in the Karhunen-Loève truncation.

Fig. 4.4 plots the block diagram that includes the rank reduction and quantization process. The procedure to obtain the truncation filters depicted in the figure is based on the eigenvalue decomposition of the channel covariance matrix of user k, $\boldsymbol{C}_{\boldsymbol{h},k}$, which reads as

$$\boldsymbol{C}_{\boldsymbol{h},k} = E[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\text{H}}[q]] = \sum_{i=1}^{r_k}\lambda_{k,i}\boldsymbol{v}_{k,i}^{'}\boldsymbol{v}_{k,i}^{'\text{H}} = \boldsymbol{V}_k^{'}\boldsymbol{\Lambda}_k\boldsymbol{V}_k^{'\text{H}} \tag{4.12}$$

where r_k is the rank of $\boldsymbol{C}_{\boldsymbol{h},k}$ and $\boldsymbol{v}_{k,i}^{'}$ and $\lambda_{k,i}$ are, respectively, the i-th eigenvector (or the i–th column of the matrix $\boldsymbol{V}_k^{'}$) and the i-th eigenvalue of $\boldsymbol{C}_{\boldsymbol{h},k}$ (or the i–th entry of the diagonal matrix $\boldsymbol{\Lambda}_k$). Note that $\boldsymbol{V}_k^{'}$ is a unitary matrix that satisfies that $\boldsymbol{V}_k^{'}\boldsymbol{V}_k^{'\text{H}} = \mathbf{I}_N$.

The KL transform defines a vector whose coefficients are called *coefficients of the KL transform* given by

$$\tilde{\boldsymbol{h}}_k[q] = \boldsymbol{V}_k^{'\mathrm{H}} \hat{\boldsymbol{h}}_{\mathrm{es},k}[q] \in \mathbb{C}^N.$$

So, the transmitter can reconstruct the channel vector by multiplying these channel coefficients from the left with $\boldsymbol{V}_k^{'}$, i.e.

$$\hat{\boldsymbol{h}}_k[q] = \boldsymbol{V}_k^{'} \tilde{\boldsymbol{h}}_k[q] \in \mathbb{C}^N.$$

Note that no errors are added to our channel estimation if all the coefficients of the KL transform are employed since $\boldsymbol{V}_k^{'}$ is an N-dimensional orthonormal basis.

Taking into account the good energy compaction properties of the KL decomposition, we can reduce the number of KL coefficients sent from each user by means of a basis $\boldsymbol{V}_k$ (see Fig. 4.4) that gives the following KL coefficients

$$\tilde{\boldsymbol{h}}_{\mathrm{T},k}[q] = \boldsymbol{V}_k^{\mathrm{H}} \hat{\boldsymbol{h}}_{\mathrm{es},k}[q] \in \mathbb{C}^d \tag{4.13}$$

where the new basis $\boldsymbol{V}_k$ is defined as $\boldsymbol{V}_k = [\boldsymbol{v}'_{k,1}, \ldots, \boldsymbol{v}'_{k,d}] \in \mathbb{C}^{N \times d}$ and $d \leq N$ denotes the number of KL coefficients sent from the receiver after truncation, i.e. the dimensionality of the rank reduction. The subindex $_{\mathrm{T}}$ highlights that the CSI errors are due to the truncation of the KL coefficients at the receiver, together with the errors due to the estimation process, which are always considered, as shown in Fig. 4.2. Note that now $\boldsymbol{V}_k$, the so called *rank reduction basis*, satisfies $\boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{V}_k = \mathbf{I}_d$, but $\boldsymbol{V}_k \boldsymbol{V}_k^{\mathrm{H}} \neq \mathbf{I}_N$. This leads to errors resulting from the compression of the information due to the coefficient truncation. Finally, the vector channel recovered at the transmitter is given by

$$\hat{\boldsymbol{h}}_{\mathrm{T},k}[q] = \boldsymbol{V}_k \tilde{\boldsymbol{h}}_{\mathrm{T},k}[q] \in \mathbb{C}^N. \tag{4.14}$$

Fig. 4.3 depicts this overall rank reduction process as will be used throughout this work.

Under the assumption that the channel statistics do not depend on time, the modal matrix obtained from the eigenvalue decomposition in Eq. (4.12) is also constant over time. With this assumption, only the coefficients of the reduced rank approximation have to be sent from the receiver to the transmitter to capture the fast variations of the channel (referred to as *short–term* variations).

4.3 Quantization Error

Quantization is the process of constraining some quantity from a continuous set of values to a discrete one. It is widely used in image and speech processing, for example, and also some compression schemes related to music use quantization, leading to lossy data compression. In the context of this work, quantization is motivated by the need to reduce

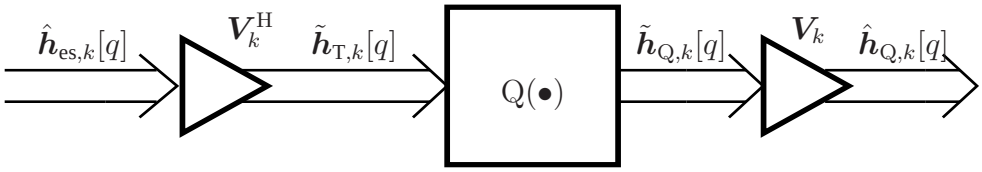


Figure 4.4: Type B and C Errors: Truncation and Quantization.

the amount of data necessary to represent the CSI sent from the users to the transmitter. We distinguish between scalar and vector quantization (both depicted as $\text{Q}(\bullet)$ in Fig. 4.2). Both methods will be applied to compress the fed–back CSI throughout this work.

The quantization process is described by the following errors model,

$$\hat{\boldsymbol{h}}_{\text{Q},k}[q] = \hat{\boldsymbol{h}}_{\text{es},k}[q] + \boldsymbol{\eta}_{\text{Q},k}[q] \tag{4.15}$$

where $\hat{\boldsymbol{h}}_{\text{Q},k}[q] \in \mathbb{C}^N$ denotes the quantized version of $\hat{\boldsymbol{h}}_{\text{es},k}[q]$ and is equal to one of the codebook entries. Additionally, the errors $\boldsymbol{\eta}_{\text{Q},k}[q] \in \mathbb{C}^N$ are assumed to be mutually independent and independent with the channel estimates $\hat{\boldsymbol{h}}_{\text{es},k}[q]$. Note that we directly quantize the channel estimates without performing rank reduction after estimation. This model is appropriate for *Random Vector Quantization* (RVQ), which is explained below.

Alternatively, if the quantizer is applied after the rank reduction of the channel estimates, Eq. (4.15) can be rewritten as

$$\tilde{\boldsymbol{h}}_{\text{Q},k}[q] = \tilde{\boldsymbol{h}}_{\text{T},k}[q] + \tilde{\boldsymbol{\eta}}_{\text{Q},k}[q] \tag{4.16}$$

where $\tilde{\boldsymbol{h}}_{\text{Q},k}[q] \in \mathbb{C}^d$ comprises the representants or codebook entries, as plotted in Figs. 4.1 and 4.4. Again, the errors $\tilde{\boldsymbol{\eta}}_{\text{Q},k}[q] \in \mathbb{C}^d$ are assumed to be mutually independent and independent of the truncated channel estimates $\tilde{\boldsymbol{h}}_{\text{T},k}[q] \in \mathbb{C}^d$. For this setup, each of the scalar coefficients of $\tilde{\boldsymbol{h}}_{\text{T},k}[q]$ is scalarly quantized as is shown in the following subsection.

4.3.1 Scalar Quantization

The scalar quantizer (as any quantizer) can be explicitly separated into two parts, an encoder and a decoder. The encoder maps $\mathcal{E} : \mathbb{R} \rightarrow \mathbb{I}$, where $\mathbb{R}$ are the real numbers and $\mathbb{I}$ is the index set for partition cells. The decoder is the mapping $\mathcal{D} : \mathbb{I} \rightarrow \mathbb{R}$, so the quantizer can be written as $\text{Q}(x) = \mathcal{D}(\mathcal{E}(x))$. Note that we restrict ourselves to real scalar quantizers, since the scalar quantization of a complex number is in fact the vector quantization of two real–valued quantities.

The output set or *codebook* $\mathcal{C} = \{y_1, y_2, y_3, \ldots, y_M\}$ with $y_i \in \mathbb{R}$ for the decoder process satisfies that $y_1 \leq y_2 \leq \ldots y_M$, with the codebook size $|\mathcal{C}| = M$. Therefore,

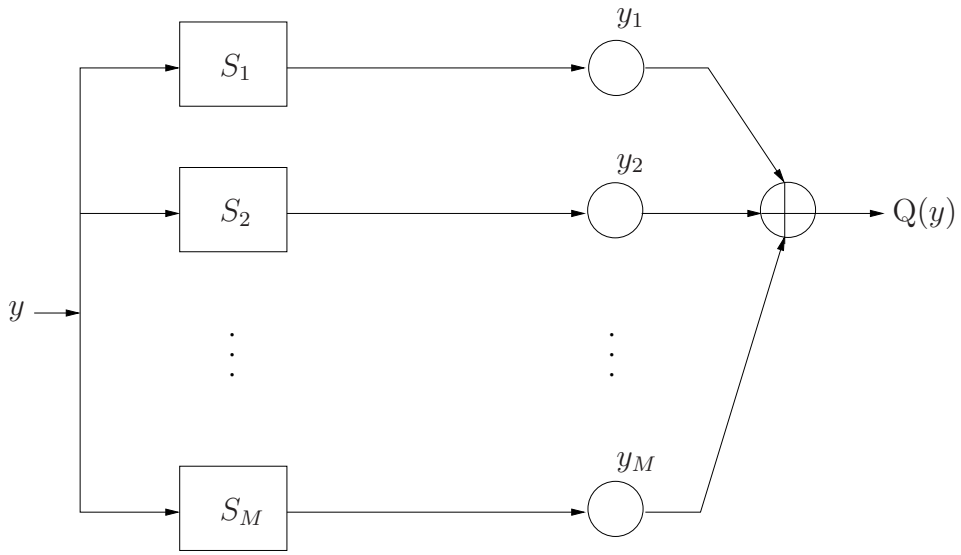


Figure 4.5: Structure of a Scalar Quantizer.

the number of bits needed to specify the quantized value is given by $\log_2 M$, the so–called *resolution*. The resolution indicates the accuracy with which the original analog amplitude is described.

The encoder operation can be modeled by means of a selector function $\mathrm{S}_i(y)$. The selector function $\mathrm{S}_i(y)$ is 1 if $y \in \mathcal{R}_i$, i.e. the i–th partition region, and 0 elsewhere, and the quantization task can be expressed as

$$\mathrm{Q}(y) = \sum_{i=1}^{M} y_i \, \mathrm{S}_i(y). \tag{4.17}$$

Note that for any given input value x, only one term of the sum is nonzero. In the sequel, we will use this notation for scalar quantization. Fig. 4.5 depicts the structure of a scalar quantizer according to this notation.

The main goal of the quantizer design is to select the representants and the partition regions or cells to provide the minimum possible average distortion $\mathrm{E}[\mathrm{d}(\bullet, \bullet)]$ for a fixed number of levels M. In general, this problem has no explicit, closed-form solution but some efficient algorithms can be used.

By assuming that one part (the encoder or the decoder) is fixed, it becomes easy to specify a condition for optimality of the other part. Specifically, the encoder part of an optimal quantizer must be optimal for the given decoder while the decoder must be optimal for the given encoder [108]. The two conditions are necessary for a quantizer to be optimal.

The best encoder for a given codebook satisfies the *nearest neighbor condition*. This

requires that the i–th region of the partition consists of all input values closer to y_i than to any other output level, i.e. the partition cells satisfy

$$\mathcal{R}_i \subseteq \{x \in \mathbb{R} : \mathrm{d}(x, y_i) \leq \mathrm{d}(x, y_j) \quad \forall j\}$$

that is,

$$\mathrm{Q}(x) = y_i \quad \text{only if } \mathrm{d}(x, y_i) \leq \mathrm{d}(x, y_j) \quad \forall j.$$

Thus, given the decoder, the encoder minimizes the distortion:

$$\mathrm{d}(x, \mathrm{Q}(x)) = \min_{y_i \in \mathcal{C}} \mathrm{d}(x, y_i). \tag{4.18}$$

This result holds in general if the goal of the quantizer is the minimization of the average distortion. The most convenient and widely used measure of distortion between an input x and its quantized value $\mathrm{Q}(x)$ is the *squared error* or *squared Euclidean distance* between two scalar values, defined as

$$\mathrm{d}(x, \mathrm{Q}(x)) = |x - \mathrm{Q}(x)|^2 . \tag{4.19}$$

We now examine the second necessary condition for optimality which is obtained by fixing the encoder (partition) and optimizing the decoder (codebook). The *centroid condition* is found based on this optimization provided that the squared error distortion measure is used. The centroid condition is simply the condition that the optimal output level, y_i, for the i–th cell of the partition is the *centroid*, or *center of mass*, of that part of the input PDF that lies in the region $\mathcal{R}_i$, i.e.

$$y_i = \mathrm{E}[x | x \in \mathcal{R}_i] = \int_{\mathbb{R}} x \frac{f_{x, x \in \mathcal{R}_i}(x, x \in \mathcal{R}_i)}{\Pr[x \in \mathcal{R}_i]} \, \mathrm{d}x = \frac{\mathrm{E}\left[x S_i(x)\right]}{\mathrm{E}\left[S_i(x)\right]}. \tag{4.20}$$

When the quantizer satisfies the centroid condition, the following properties are fulfilled

$$\begin{aligned} \mathrm{E}\left[\mathrm{Q}\left(x\right)\right] &= \mathrm{E}\left[x\right] \\ \mathrm{E}\left[x\,\mathrm{Q}\left(x\right)\right] &= \mathrm{E}\left[\mathrm{Q}\left(x\right)^2\right] \\ \mathrm{E}\left[\mathrm{Q}\left(x\right)^2\right] &= \mathrm{E}\left[x^2\right] - \mathrm{E}\left[\left(x - \mathrm{Q}\left(x\right)\right)^2\right]. \end{aligned} \tag{4.21}$$

Uniform Quantizer

The most common of all scalar quantizers is the uniform quantizer whose principle is rather simple (see [108]). A uniform quantizer is a quantizer where the boundary points $x_1, \ldots, x_{M-1}$ defining the partition cells $\mathcal{R}_i$ are equally spaced and the representants are the midpoints of the quantization interval. The first condition implies that with step size Δ, $x_i - x_{i-1} = \Delta$, for $i = 2, 3, \ldots, M-1$, whilst the second implies that

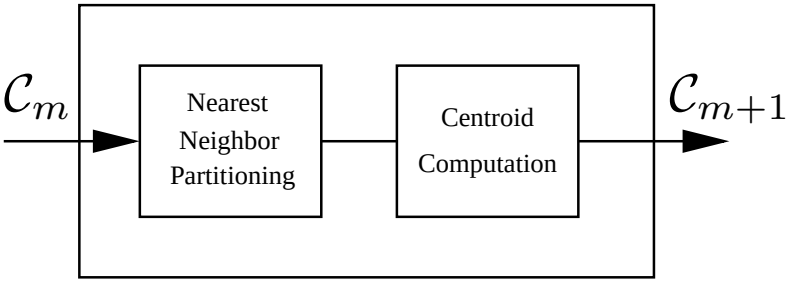


Figure 4.6: Lloyd Iteration for Codebook Improvement.

$y_i = (x_{i-1} + x_i)/2$, for $i = 2, 3, \ldots, M-1$. Consider the case of a uniform quantizer where the input is bounded with values lying in the range (a, b). When the input PDF is uniformly distributed over the region, the quantizer error $\epsilon = \mathrm{Q}(x) - x$ has a uniform PDF on $[-\Delta/2, \Delta/2]$, where $\Delta = (b-a)/M$ is the step size. Given the cell, the average distortion is simply the variance of a random variable that is uniformly distributed on an interval of width Δ, that is $\frac{\Delta^2}{12}$. Obviously, the mean of the quantizer error is zero. Although the input distribution might not be uniform, the uniform quantizer gives a reasonably good performance for a wide variety of input signals. In fact, due to this reason and also for simplicity, the uniform quantizer is widely used in A/D conversion.

In the context of this work we have the coefficients of the rank-reduced representation $\tilde{\boldsymbol{h}}_{\mathrm{T},k}[q]$ in Eq. (4.13). As we will see, although the input PDF is not uniform but Gaussian, we make the simplifying assumption that the input is bounded, i.e. we assume that the real and imaginary parts of every entry of $\tilde{\boldsymbol{h}}_{\mathrm{T},k}[q]$ lie in some interval (a, b) (see Chapter 6). The overload region has a very low probability (less than $5\,\%$) of containing an input sample. Thus, we choose representants between a and b to construct an initial codebook that is stored at both transmitter and receiver. After transmission, every receiver performs a search to find for each component of the coefficients (real and imaginary part) the element in the corresponding codebook that is closest. Then, the codebook index is fed back to the transmitter. Finally, the transmitter simply looks into its codebook and reconstructs the estimated channel from the selected codeword [103]. This estimate will be used to obtain the precoder filters.

Non-uniform Quantizer based on Gaussian Inputs and Lloyd Algorithm

When considering non-uniform quantizers, we have to select the quantizer step sizes to fit the input distribution (Gaussian in our case) employing non–uniformly spaced levels. We employ the Lloyd algorithm based not on a training sequence, but rather on the known exact input distribution. Note that it is crucial to avoid a new training sequence from the point of view of efficiency and performance. Given a codebook $\mathcal{C}_m$, the Lloyd algorithm

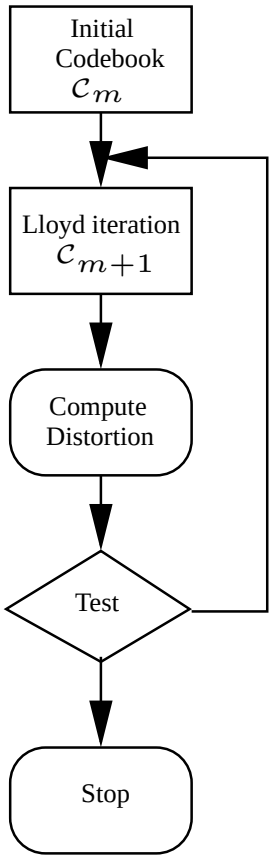


Figure 4.7: Lloyd Algorithm for Quantizer Design.

finds the optimal partition by means of using the nearest neighbor condition. The partition cells are thus defined as

$$\mathcal{R}_i \subset \{x \in \mathbb{R} : \mathrm{d}(x, y_i) \leq \mathrm{d}(x, y_j) \quad \forall j \neq i\}.$$

Basically, the procedure is as follows. We begin with an initial codebook $\mathcal{C}_1(m = 1)$, for example, the uniform codebook. Then, in the next step, given the codebook $\mathcal{C}_m$, we perform the Lloyd algorithm to generate the improved codebook $\mathcal{C}_{m+1}$. After that, we compute the average distortion for $\mathcal{C}_{m+1}$. If it is less than a fixed threshold, we stop. Otherwise, we set $m + 1 \rightarrow m$ and repeat the previous step. In this way, we get a locally optimal codebook for the Gaussian input with the minimum average distortion (see Figs. 4.6 and 4.7). [2]

This quantization scheme is modified when considering the final approach shown in

[2]The nearest neighbor condition and the centroid condition are only necessary but not sufficient for global optimality. Therefore, only local optimality is ensured by the Lloyd algorithm.

Chapter 7, where the codebook entries are not the channel coefficients but the employed precoders. Therefore, the index fed back by each user corresponds to a set of precoders and the intersection of the sets gives us the best precoder (see Chapter 7).

4.3.2 Vector Quantization

Vector quantization (VQ) is a lossy data compression method based on the principle of block coding. Vector quantization is the generalization of scalar quantization to higher dimensions. Although many of the ideas of scalar quantization can be applied to the more general scheme of vector quantization, VQ offers much greater compressing potential than scalar quantization. Again, the main goal when designing vector quantizers is to find a codebook, i.e. the decoder, and a partition or encoding rule. As for scalar quantization, the coefficients of the rank-reduced representation are the input to the quantizer

$$\mathrm{Q}(\boldsymbol{y}) = \sum_{i=1}^{M} \boldsymbol{y}_i S_i(\boldsymbol{y})$$

where $\boldsymbol{y}_i \in \mathbb{C}^d, i = 1, \ldots, M$ are the codebook entries. Note that this structure is equivalent to the scalar one in Fig. 4.5 and is depicted in Fig. 4.8. The disjoint partition cells $\mathcal{R}_i$ fulfill $\bigcup_{i=1}^{M} \mathcal{R}_i = \mathbb{C}^d$ and $\mathcal{R}_i \bigcap \mathcal{R}_j = \emptyset$, i.e. $\mathrm{Q}(\boldsymbol{y})$ is regular (e.g. [108]).

We assume that the codebook size M is given; the d–dimensional input random vector $\boldsymbol{x}$ (i.e. the channel coefficient) is statistically specified; and a particular distortion measure $d(\bullet, \bullet)$ has been selected. As for the scalar quantizer, we choose the *squared Euclidean distance* between two complex vectors defined as

$$d\left(x, \mathrm{Q}\left(x\right)\right) = \left\|\boldsymbol{x} - \mathrm{Q}(\boldsymbol{x})\right\|_2^2 = \left(\boldsymbol{x} - \mathrm{Q}(\boldsymbol{x})\right)^{\mathrm{H}} \left(\boldsymbol{x} - \mathrm{Q}(\boldsymbol{x})\right) = \sum_{i=1}^{d} \left|x_i - \mathrm{Q}(x_i)\right|^2 .$$

We wish to determine the necessary conditions for a quantizer to be optimal in the sense that it minimizes the average distortion for the given conditions. Recall that the encoder is completely specified by the partition cells and the decoder is completely specified by the codebook $\mathcal{C}$. For a given codebook, an optimal partition is one specifying the *nearest neighbor condition*: for each i, all input points closer to code vector $\boldsymbol{y}_i$ than to any other code vector should be assigned to region $\mathcal{R}_i$. Thus, by considering the *nearest neighbor condition* we have that the optimal partition cells satisfy

$$\mathcal{R}_i \subset \{\boldsymbol{x} \in \mathbb{C}^d : \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}_i) \leq \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}_j) \quad \forall j\}$$

for a given codebook $\mathcal{C} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_M\}$, that is,

$$\mathrm{Q}(\boldsymbol{x}) = \boldsymbol{y}_i \quad \text{only if} \quad \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}_i) \leq \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}_j) \quad \forall j.$$

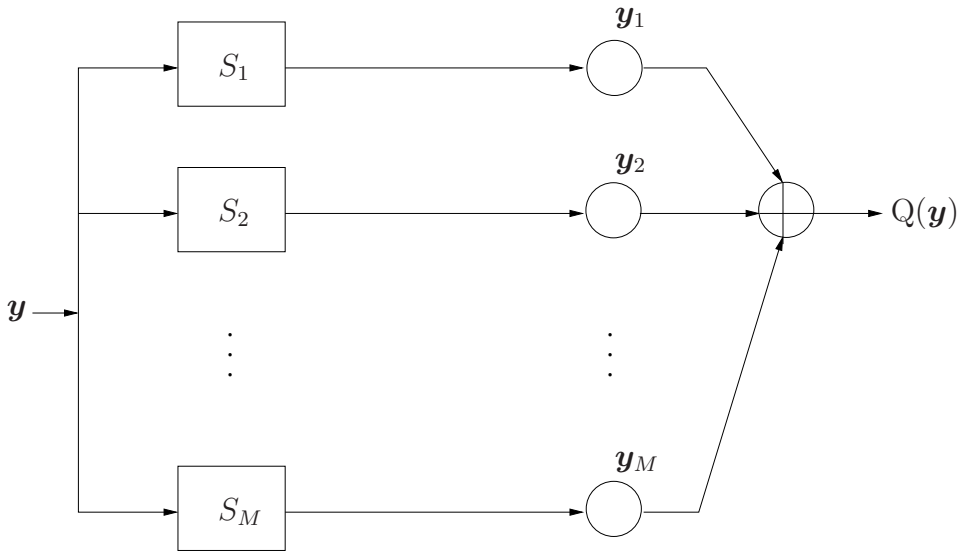


Figure 4.8: Structure of a Vector Quantizer.

Thus, given the decoder, the encoder is a minimum distortion or nearest neighbor mapping, and hence

$$\mathrm{d}(\boldsymbol{x}, \mathrm{Q}(\boldsymbol{x})) = \min_{\boldsymbol{y}_i \in \mathcal{C}} \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}_i).$$

We next consider the optimality of the codebook for a given partition. This leads to the *centroid condition* for specifying the code vector associated with each partition region. For a given partition $\mathcal{R}_i, i = 1, \ldots, M$, the optimal code vectors satisfy

$$\boldsymbol{y}_i = \mathrm{cent}(\mathcal{R}_i) = \mathrm{E}[\boldsymbol{x} | \boldsymbol{x} \in \mathcal{R}_i] = \frac{\mathrm{E}[\boldsymbol{x}\, \mathrm{S}_i(\boldsymbol{x})]}{\mathrm{E}[\mathrm{S}_i(\boldsymbol{x})]}.$$

A vector quantizer which satisfies the centroid condition for the squared error distortion measure has the following properties

$$\begin{aligned} \mathrm{E}[\mathrm{Q}(\boldsymbol{x})] &= \mathrm{E}[\boldsymbol{x}] \\ \mathrm{E}[\boldsymbol{x}^{\mathrm{H}}\, \mathrm{Q}(\boldsymbol{x})] &= \mathrm{E}[\| \mathrm{Q}(\boldsymbol{x})\|^2] \\ \mathrm{E}[\| \mathrm{Q}(\boldsymbol{x})\|^2] &= \mathrm{E}[\|\boldsymbol{x}\|^2] - \mathrm{E}[\|\boldsymbol{x} - \mathrm{Q}(\boldsymbol{x})\|^2]. \end{aligned}$$

The result is an exact generalization of the conditions for scalar quantizers explained above [cf. Eq. (4.21)].

Random Vector Quantization (RVQ)

In the context of this work, the codebook entries of each user are generated such that

$$\boldsymbol{y}_{k,i} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{h},k}) \quad i = 1, \ldots, M \tag{4.22}$$

where $\boldsymbol{C}_{\boldsymbol{h},k}$ is the channel covariance matrix of user k. Note that the codebook entries now contain channel vectors of dimension N as opposed to the channel coefficients of dimension $d \leq N$ stored in the scalar codebook explained in Subsection 4.3.1. These codebooks are simply randomly selected according to the distribution of $\boldsymbol{y}_{k,i}$ in Eq. (4.22). We henceforth refer to this method for quantizing the vector inputs as *Random Vector Quantization* (RVQ).

The procedure for RVQ is as follows. The estimated channel $\hat{\boldsymbol{h}}_{\text{es},k}[q]$ in Eq. (4.2) constitutes the input to the random vector quantizer. The RVQ tries to approximate $\hat{\boldsymbol{h}}_{\text{es},k}[q]$ by one of the M entries $\boldsymbol{y}_{k,i}$ of the codebook, so the squared error between the estimated channel and the codebook entries is minimized, i.e.

$$i_{\min,k} = \min_{i} \left\| \hat{\boldsymbol{h}}_{\text{es},k}[q] - \boldsymbol{y}_{k,i} \right\|_2^2 \tag{4.23}$$

where $i_{\min,k}$ is the index for the k-th user corresponding to the codebook entry that minimizes the above squared error. Finally, this index is transmitted over the error-free feedback channel. The transmitter collects the indices of all the users and recovers the corresponding codebook entries to construct the erroneous channel matrix that will be used for the design of the precoder filters. Note that rank reduction is not performed before vector quantization given that the channel vectors, rather than the channel coefficients, are quantized.

4.4 Feedback Delay Error

The transmission over the feedback channel introduces a certain delay of D slots, where the precoder is designed during the time slot q and the most recent channel estimate was obtained during the time slot $q - D$. The feedback delay error is modeled by the Dirac's delta shown in Fig. 4.1. This delay can equivalently be modeled as follows. The channel estimate is obtained from outdated training data, i.e. the observation of the estimator is delayed by D slots and, then, the respective feedback channel has no delay.

Bearing in mind the temporal correlation properties of $\boldsymbol{h}_k[q]$ and $\boldsymbol{h}_{\text{w},k}[q]$ described in Chapter 2, we have that

$$\mathrm{E}\left[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q-D]\right] = \mathrm{J}_0\left(\alpha_k D\right)\boldsymbol{C}_{\boldsymbol{h},k} = r_k \boldsymbol{C}_{\boldsymbol{h},k} \tag{4.24}$$

where r_k is implicitly defined. $\mathrm{J}_0(\bullet)$ denotes the zero–th order Bessel function of the first kind and $\alpha_k = 2\pi \frac{f_{\text{D,max},k}}{f_{\text{slot}}}$, where $f_{\text{D,max},k}$ is the *maximum* Doppler frequency of user k and f_{slot} is the slot rate [34].

As we showed in [109], a performance improvement is obtained when channel prediction takes into account a greater number of delayed channel vectors from each user

instead of only one outdated estimate. Let L be the number of delayed vectors processed at the transmitter, with $L \geq 1$. When the transmitter processes multiple feedback vectors, the channel information vector has to be stacked as

$$\bar{\boldsymbol{h}}_k[q] = \left[\boldsymbol{h}_k^{\mathrm{T}}[q - D_1], \cdots, \boldsymbol{h}_k^{\mathrm{T}}[q - D_L]\right]^{\mathrm{T}} \in \mathbb{C}^{dL} \tag{4.25}$$

where $D_i, i = 1, \ldots, L$, is the delay expressed as the number of slots for the i-th vector. The covariance matrix of $\bar{\boldsymbol{h}}_k$ is given by $\boldsymbol{C}_{\bar{\boldsymbol{h}},k} = \boldsymbol{C}_{\text{temp},k} \otimes \boldsymbol{C}_{\boldsymbol{h},k}$, where $\boldsymbol{C}_{\text{temp},k}$ comprises the temporal correlations and its i-th element in the j-th column is

$$\left[\boldsymbol{C}_{\text{temp},k}\right]_{i,j} = \begin{cases} \mathrm{J}_0\left(\alpha_k \left(D_i - D_j\right)\right) & j \neq i \\ 1 & j = i. \end{cases} \tag{4.26}$$

Remember that $\alpha_k = 2\pi f_{\text{D,max},k} / f_{\text{slot}}$ as given in Eq. (2.24). Moreover, we have that

$$\mathrm{E}\left[\bar{\boldsymbol{h}}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q]\right] = \boldsymbol{\beta}_k \otimes \boldsymbol{C}_{\boldsymbol{h},k} \tag{4.27}$$

where $\boldsymbol{\beta}_k$ is defined as $\boldsymbol{\beta}_k = [\mathrm{J}_0(\alpha_k D_1), \ldots, \mathrm{J}_0(\alpha_k D_L)]^{\mathrm{T}} \in \mathbb{R}^L$.

The effect of delay is quite important, especially in real systems or MIMO testbeds (cf. Chapter 8). Performance can be strongly degraded when the channel varies while CSI is sent through the feedback channel, since we would be designing our precoder with an outdated channel estimate. Therefore, it is crucial to ensure that the delay introduced by the feedback channel is less than the channel coherence time (cf. Chapter 2). It could be interesting to determine the maximum delay that the system can support before losing channel tracking.

4.5 Simulation Results

Some computer simulations were carried out to illustrate the performance degradation caused by the mismatch between the true channel and the erroneous channel available at the transmitter. The results are the mean of $5{,}000$ channel realizations and 50 symbols were transmitted per channel realization. The input bits are QPSK modulated. A feedback delay of $D = 2$ slots is considered for all the users. The slot duration was $T_{\text{slot}} = 6.67$ ms at a carrier frequency of 2 GHz. We use the second channel described in Chapter 2 for the simulations due to its intermediate BER performance and diversity. The velocity of each user is set to $v = 10$ km/h. The results are obtained for a system with $N = 4$ antennas at the transmitter and $K = 4$ single antenna users. Figs. 4.9, 4.10, and 4.11 show the BER performance achieved when respectively linear precoding, Tomlinson-Harashima precoding, or vector precoding are implemented according to the Wiener criterion.

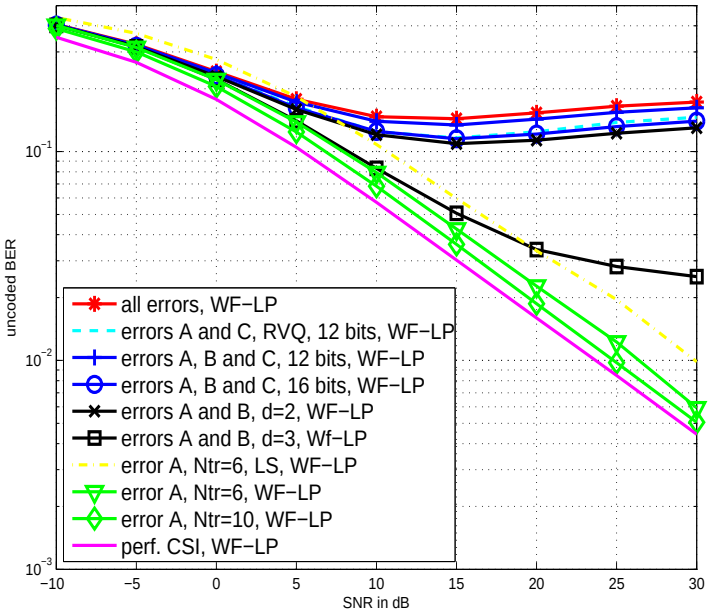


Figure 4.9: Effect of Different Types of Errors on the TxWF Scheme in an Urban Macrocell Environment. Error A: Estimation; Error B: Rank Reduction; Error C: Quantization; All Errors: Estimation, Rank Reduction, Quantization, and Delay.

First, one can see the effect of different lengths of the training sequence. The figures illustrate the performance degradation caused by channel estimation errors when this is the only source of errors in the system. As a compromise between training sequence length and performance degradation, we pick for our subsequent simulations the value $N_{\text{tr}} = 6$, which implies a loss of about 1 dB at BER=10^{-2} for the TxWF and of 2 dB for THP and VP at this same operation point with respect to the case of perfect CSI knowledge. A comparison between both LS and MMSE estimation methods for $N_{\text{tr}} = 6$ can also be observed and, as expected, the MMSE solution clearly outperforms the LS approach in an urban macrocell environment, i.e. for correlated channels.

Then, rank reduction is applied so that the number of coefficients sent from each user to the transmitter is reduced from $d = N = 4$ to $d = 3$ or $d = 2$. In spite of an important deterioration in performance when the number of coefficients is reduced, we have the enormous advantage of reducing the overhead of the feedback channel, especially for a high number of antennas at the transmitter. After the quantizer, only the codebook indices corresponding to the real and imaginary part of each user's coefficients are transmitted through the feedback channel. These coefficients are scalarly quantized in these simulations using 8 bits (4 bits per complex dimension), which yields 16 bits per user (for $d = 2$); or 6 bits (3 bits per complex dimension), which yields 12 bits per

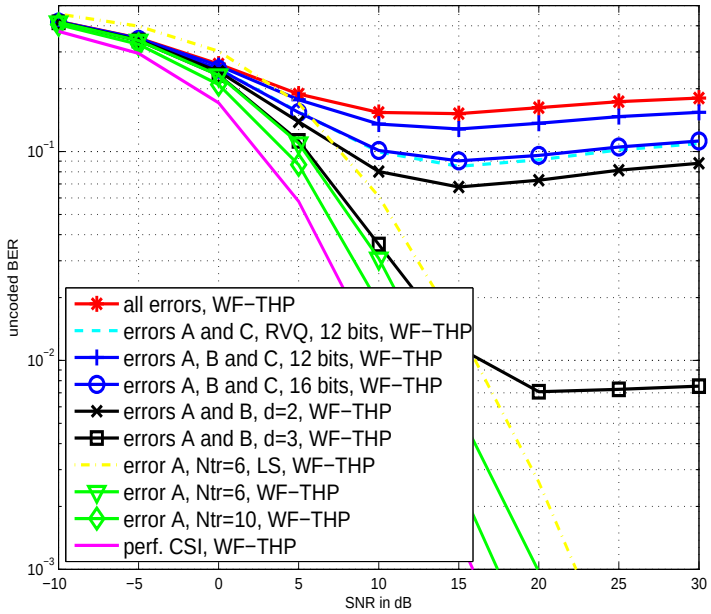


Figure 4.10: Effect of Different Types of Errors on the WF-THP Scheme in an Urban Macrocell Environment. Error A: Estimation; Error B: Rank Reduction; Error C: Quantization; All Errors: Estimation, Rank Reduction, Quantization, and Delay.

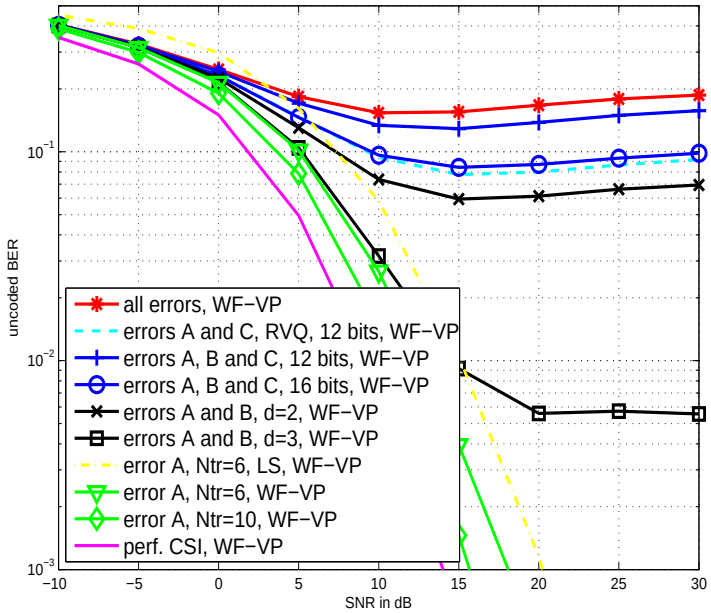


Figure 4.11: Effect of Different Types of Errors on the WF-VP Scheme in an Urban Macrocell Environment. Error A: Estimation; Error B: Rank Reduction; Error C: Quantization; All Errors: Estimation, Rank Reduction, Quantization, and Delay.

user. When the codebook is larger it is obvious that the results are much better at the cost of decreasing the compression for the CSI sent through the feedback channel and greatly extending the storage capability necessary at the users. The figures also plot the uncoded BER when random vector quantization (RVQ) feeding back 12 bits is applied instead of scalar quantization. As expected, the system performance is better when VQ is used. This is because VQ carries out a joint quantization that uses a much larger codebook ($2^{12} = 4{,}096$ entries per user) and compares an N-dimensional vector with $4{,}096$ complex vectors to choose the closest one for each channel realization. Therefore, its computational complexity is much greater than that of scalar quantization, where the search is reduced to a comparison with $2^3 = 8$ scalar values for the real and imaginary parts of each fed–back coefficient. For the considered number of 12 fed–back bits per user, it is clear that the vector quantization performance for medium and high SNR is better than that of scalar quantization.

Finally, one can see the influence on the uncoded BER of the feedback delay in the figures. Even though it is obvious that each new error source degrades the system more and more, note the strong performance degradation when moving from $d = 3$ truncated coefficients to $d = 2$.

It can also be seen how the curves go up for high SNR due to imperfect CSI. This effect can be explained by the convergence of the WF designs to the ZF precoders, which are highly sensitive to CSI errors. Moreover, the effect of imperfect CSI is more pronounced for THP or VP than for LP due to the modulo operators. We will see in the following chapters how an improvement in performance can be achieved when robust designs are considered.

4.6 Conclusions

In this chapter, we introduced the errors of the CSI available at the transmitter side instead of full CSI knowledge. In that case, it is crucial to determine what kind of information is sent from the different users and how this information can be obtained by the transmitter. First, we considered the effect of estimating the channel using supervised methods. To this end, we briefly discussed the least-squares and the linear MMSE channel estimation approaches. Although the least-squares approach shows less quality than the LMMSE channel estimator when full knowledge about the statistical second order moments is assumed, this LS estimator is quite useful due to its simplicity and relatively good performance. We also explained how we can compress the channel information sent through the feedback channel by means of the Karhunen-Loève decomposition. Then, we quantized the KL resulting coefficients in order to ensure that the feedback channel overhead is strongly reduced. We introduced scalar and vector quantization as different

methods for compressing the CSI. The issue about the superiority of vector quantization over scalar quantization for random input vectors is evident from the increased freedom in choosing the partition geometry for VQ compared to the highly restrictive geometry in the case where each vector element is scalarly quantized and the resulting quantization cells are rectangles. Thus, scalar quantization is simply a restricted special case of VQ. Indeed, scalar quantization will be very useful for us due to its simpler partition regions that produce robust designs for the optimized limited feedback. Finally, we introduced the feedback delay error that adds a new mismatch between the true channel and the erroneous channel resulting from the estimation, rank reduction and quantization processes.

We showed by means of some computer simulations how each type of error degrades the system more and more at the cost of reducing the overhead of the feedback channel. However, there exists an important effect at high SNR consisting of increasing BER. This effect may be corrected by an adequate limited feedback design combined with the robust precoders shown in the next chapter.

Chapter 5

Robust Wiener Precoders for Imperfect CSI

In FDD systems, the transmitter is unable to obtain the CSI by estimating the uplink channels in a wireless communication system since the channels at different frequencies are not reciprocal. Therefore, this CSI has to be obtained at the receiver side by the users to be sent to the transmitter by means of a feedback channel that often suffers from limited bandwidth. The CSI is affected by different sources of errors, so that all the precoding parameters are designed with an erroneous version of the channel instead of the true channel. Errors in the available CSI have a significant impact on the performance of precoding. In this chapter, we present a new *robust* precoder design based on the error model described in the previous chapter. The objective is to compensate this mismatch between the true channel and the erroneous channel sent from receivers by means of an adequate design for precoding that involves the error model to mitigate its effect, i.e. to be more *robust* against errors in CSI.

We first focus our attention on the general statistical error model to be used for the robust proposals throughout this work. The chapter continues with the MMSE derivation of the different types of precoders shown in Chapter 3 to be robust against CSI errors. As opposed to the receive coefficients arising from the precoder optimization as used in Chapter 3, we next include a brief discussion about the MMSE coefficients that are employed instead in order to compensate the mismatch between perfect and erroneous CSI available at the transmitter side. The chapter concludes with some comments about the training symbols that must be sent to enable the proposed systems to work properly.

We focus on the MMSE criterion since precoders based on this criterion clearly outperform precoders based on the ZF criterion, as seen in Chapter 3 for the perfect CSI case. Since an erroneous channel version can be seen as the sum of the true channel and some noise, there is no sense in using ZF precoders in such a situation because the effect of noise amplifying inherent to ZF worsens its performance in the presence of errors in

CSI when compared to an MMSE precoder design.

5.1 Statistical Errors Model

The basic premise of our channel error model is the interpretation of the true channel $\boldsymbol{H}$ as a matrix of random variables, whereas the channel estimate $\hat{\boldsymbol{H}}$ available at the transmitter is deterministic. The relationship between $\boldsymbol{H}$ and $\hat{\boldsymbol{H}}$ is given by [17, 43, 44]

$$\boldsymbol{H}[q] = \hat{\boldsymbol{H}}[q] + \boldsymbol{\Theta}[q] \tag{5.1}$$

where $\boldsymbol{\Theta}[q] = \boldsymbol{H}[q] - \hat{\boldsymbol{H}}[q]$ is the CSI error at the transmitter. The matrix $\hat{\boldsymbol{H}}[q] = [\hat{\boldsymbol{h}}_{\mathrm{Q},1}[q], \ldots, \hat{\boldsymbol{h}}_{\mathrm{Q},K}[q]]^{\mathrm{T}} \in \mathbb{C}^{K\times N}$ comprises the estimates obtained from the feedback of the quantized coefficients of the rank reduced channel and $\boldsymbol{\Theta}[q] = [\boldsymbol{\theta}_1[q], \ldots, \boldsymbol{\theta}_K[q]]^{\mathrm{T}} \in \mathbb{C}^{K\times N}$ collects the CSI error vectors of the different users. Each row of this matrix can be written as the column vector

$$\boldsymbol{\theta}_k[q] = \boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_{\mathrm{Q},k}[q] \in \mathbb{C}^N, \quad k = 1, \ldots, K.$$

The error covariance matrix of the zero–mean matrix $\boldsymbol{\Theta}[q]$ is given by

$$\begin{aligned}\boldsymbol{C}_{\boldsymbol{\Theta}} &= \mathrm{E}\left[\boldsymbol{\Theta}^{\mathrm{H}}[q]\boldsymbol{\Theta}[q]\right] = \mathrm{E}\left[\left(\boldsymbol{\theta}_1^*[q], \ldots, \boldsymbol{\theta}_k^*[q]\right)\left(\boldsymbol{\theta}_1[q], \ldots, \boldsymbol{\theta}_k[q]\right)^{\mathrm{T}}\right] \\ &= \sum_{k=1}^{K} \mathrm{E}\left[\left(\boldsymbol{h}_k^*[q] - \hat{\boldsymbol{h}}_{\mathrm{Q},k}^*[q]\right)\left(\boldsymbol{h}_k^{\mathrm{T}}[q] - \hat{\boldsymbol{h}}_{\mathrm{Q},k}^{\mathrm{T}}[q]\right)\right] = \sum_{k=1}^{K} \boldsymbol{C}_{\boldsymbol{\Theta},k}^*\end{aligned}$$

where $\boldsymbol{C}_{\boldsymbol{\Theta},k} = \mathrm{E}[(\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_{\mathrm{Q},k}[q])(\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_{\mathrm{Q},k}[q])^{\mathrm{H}}]$ is the error covariance matrix of user k.

Taking into account the above error model and assuming $\mathrm{E}[\boldsymbol{\Theta}[q]] = \boldsymbol{0}$, the channel mean and the channel Gram mean are given by

$$\begin{aligned}\mathrm{E}_{\boldsymbol{\Theta}}\left[\boldsymbol{H}[q]\right] &= \hat{\boldsymbol{H}}[q] \\ \mathrm{E}_{\boldsymbol{\Theta}}\left[\boldsymbol{H}^{\mathrm{H}}[q]\boldsymbol{H}[q]\right] &= \hat{\boldsymbol{H}}^{\mathrm{H}}[q]\hat{\boldsymbol{H}}[q] + \boldsymbol{C}_{\boldsymbol{\Theta}}\end{aligned} \tag{5.2}$$

respectively. The subindex $_{\boldsymbol{\Theta}}$ for the expectation denotes that the expectation is taken only over this matrix of random variables since $\hat{\boldsymbol{H}}$ is deterministic.

In this chapter, we optimize each type of precoder for a MU-MISO system according to the MMSE criterion, but incorporating the model for errors of Eq. (5.1) to compensate the effect due to erroneous CSI at the transmitter.

Note that for the sake of brevity we will henceforth omit the time slot index q used in Chapter 2.

5.2 MU–MISO Robust Linear Wiener Precoding

Fig. 5.1 depicts the block diagram of a MU-MISO system with linear precoding. The transmitted signal $\boldsymbol{x}[n] \in \mathbb{C}^N$ results from a linear transformation of the symbols $\boldsymbol{u}[n] \in \mathbb{C}^K$, i.e. $\boldsymbol{x}[n] = \boldsymbol{F}\boldsymbol{u}[n]$ as explained in Subsection 3.1.2. For robust linear MMSE precoding, $\boldsymbol{F} \in \mathbb{C}^{N\times K}$ together with the common receive weight $g \in \mathbb{R}$ minimizes the mean of the MSE under a transmit power constraint, i.e.

$$\begin{aligned}\{\boldsymbol{F}_{\text{Rlin}}, g_{\text{Rlin}}\} = \underset{\{\boldsymbol{F},g\}}{\operatorname{argmin}}\, \mathrm{E}_{\boldsymbol{\Theta}}\left[\mathrm{E}\left[\|\boldsymbol{u}[n] - g\boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n] - g\boldsymbol{\eta}[n]\|_2^2\right]\right] \\ \text{s.t.:} \quad \mathrm{E}\left[\|\boldsymbol{x}[n]\|_2^2\right] \leq E_{\text{tx}}.\end{aligned} \tag{5.3}$$

This optimization can be solved following similar steps as for the standard MMSE precoder in Subsection 3.1.2. First of all, we develop the MSE cost function in Eq. (5.3) as

$$\begin{aligned}\varepsilon(\boldsymbol{F}, g) &= \mathrm{E}_{\boldsymbol{\Theta}}\left[\mathrm{E}\left[\|\boldsymbol{u}[n] - g\boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n]\|_2^2\right]\right] + \mathrm{E}_{\boldsymbol{\Theta}}\left[\mathrm{E}\left[\|g\boldsymbol{\eta}[n]\|_2^2\right]\right] \\ &= \mathrm{E}_{\boldsymbol{\Theta}}\left[\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) - \mathrm{tr}\left(g^*\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right) - \mathrm{tr}\left(g\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right) + \mathrm{tr}\left(|g|^2\,\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right)\right] \\ &\quad + |g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\end{aligned} \tag{5.4}$$

where $\boldsymbol{C}_{\boldsymbol{u}} = \mathrm{E}[\boldsymbol{u}[n]\boldsymbol{u}^{\mathrm{H}}[n]] \in \mathbb{C}^{K\times K}$ and $\boldsymbol{C}_{\boldsymbol{\eta}} = \mathrm{E}[\boldsymbol{\eta}[n][\eta]^{\mathrm{H}}[n]] \in \mathbb{C}^{K\times K}$ are the spatial covariance matrices of symbols and noise. The transmit power constraint can be written as $\mathrm{E}[\|\boldsymbol{x}[n]\|_2^2] = \mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) \leq E_{\text{tx}}$. Substituting Eq. (5.2) into Eq. (5.4), we get

$$\begin{aligned}\varepsilon(\boldsymbol{F}, g) &= \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) - \mathrm{tr}\left(g^*\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\right) - \mathrm{tr}\left(g\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\hat{\boldsymbol{H}}\right) \\ &\quad + \mathrm{tr}\left(|g|^2\,\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\right) + \mathrm{tr}\left(|g|^2\,\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\right) + |g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\end{aligned} \tag{5.5}$$

where the property of the trace operator $\mathrm{tr}(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}^{\mathrm{T}})$ (see Appendix B.2) has been applied.

The above results enable us to construct the Lagrangian function for the MSE optimization in Eq. (5.3) as follows

$$L(\boldsymbol{F}, g, \lambda) = \varepsilon(\boldsymbol{F}, g) + \lambda\left(\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\text{tx}}\right) \tag{5.6}$$

where the Lagrangian multiplier is $\lambda \in \mathbb{R}^{0,+}$.

By setting the derivatives of Eq. (5.6) with respect to $\boldsymbol{F}^*$ and g to zero, we obtain the

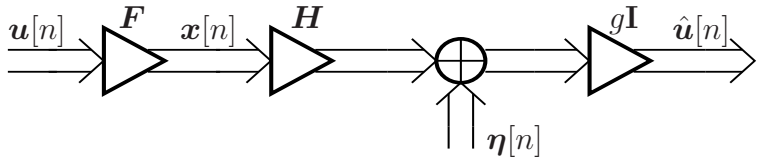


Figure 5.1: MU–MISO System with Linear Precoding.

necessary *Karush-Kuhn-Tucker* (KKT) optimality conditions [cf. Eq. (3.19)]

$$\begin{aligned}
\frac{\partial L(\bullet)}{\partial \boldsymbol{F}^*} &= -g^* \hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{C_u} + |g|^2 \hat{\boldsymbol{H}}^{\mathrm{H}} \hat{\boldsymbol{H}} \boldsymbol{F} \boldsymbol{C_u} + |g|^2 \boldsymbol{C}_{\Theta} \boldsymbol{F} \boldsymbol{C_u} + \lambda \boldsymbol{F} \boldsymbol{C_u} = \mathbf{0} \\
\frac{\partial L(\bullet)}{\partial g} &= -\operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C_u} \hat{\boldsymbol{H}}\right) + g^* \operatorname{tr}\left(\hat{\boldsymbol{H}} \boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}} \hat{\boldsymbol{H}}^{\mathrm{H}}\right) \\
&\quad + g^* \operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{C_\Theta}\right) + g^* \operatorname{tr}\left(\boldsymbol{C_\eta}\right) = 0 \\
\operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}}\right) &\leq E_{\mathrm{tx}} \\
\lambda \left(\operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right) &= 0 \quad \text{with } \lambda \geq 0
\end{aligned} \tag{5.7}$$

where the property of the trace operator $\frac{\partial \operatorname{tr}(\boldsymbol{AB})}{\partial \boldsymbol{A}} = \boldsymbol{B}^{\mathrm{T}}$ (see Appendix B.3) has been applied.

First of all, note that the constraint in Eq. (5.3) is always active, i.e. $\lambda > 0$. Indeed, from the second KKT condition we have

$$g^* = \frac{\operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C_u} \hat{\boldsymbol{H}}\right)}{\operatorname{tr}\left(\hat{\boldsymbol{H}} \boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}} \hat{\boldsymbol{H}}^{\mathrm{H}} + \boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{C}_{\Theta} + \boldsymbol{C_\eta}\right)}. \tag{5.8}$$

Additionally, the first KKT condition can be expressed as

$$\lambda \boldsymbol{F} \boldsymbol{C_u} = g^* \hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{C_u} - |g|^2 \hat{\boldsymbol{H}}^{\mathrm{H}} \hat{\boldsymbol{H}} \boldsymbol{F} \boldsymbol{C_u} - |g|^2 \boldsymbol{C}_{\Theta} \boldsymbol{F} \boldsymbol{C_u}.$$

Multiplying by $\boldsymbol{F}^{\mathrm{H}}$ and taking the trace yields

$$\lambda \operatorname{tr}\left(\boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}}\right) = g^* \operatorname{tr}\left(\hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}}\right) - |g|^2 \operatorname{tr}\left(\hat{\boldsymbol{H}}^{\mathrm{H}} \hat{\boldsymbol{H}} \boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}} + \boldsymbol{C}_{\Theta} \boldsymbol{F} \boldsymbol{C_u} \boldsymbol{F}^{\mathrm{H}}\right).$$

By plugging Eq. (5.8) into the last expression, we obtain that

$$\begin{aligned}\lambda \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) &= \frac{\left|\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\hat{\boldsymbol{H}}\right)\right|^2}{\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}} + \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\Theta} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)} \\ &\quad - \frac{\left|\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right)\right|^2 \cdot \operatorname{tr}\left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right)}{\operatorname{tr}^2\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}} + \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\Theta} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)} \\ &= \frac{\left|\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\right)\right|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)}{\operatorname{tr}^2\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}} + \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\Theta} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)}. \end{aligned} \tag{5.9}$$

Thus, we conclude that

$$\lambda \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) = |g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)$$

and therefore $\lambda = \frac{|g|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{\operatorname{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}})} > 0$ if the trivial solution $\boldsymbol{F} = \boldsymbol{0}$ is not allowed. As a consequence, the transmit energy constraint is an equality [see Eq. (5.7)], i.e.

$$\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) = E_{\mathrm{tx}}.$$

Next, let us substitute $\lambda = \frac{|g|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\mathrm{tx}}}$ into the first KKT condition:

$$-g^* \hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}} + |g|^2 \hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}} + |g|^2 \boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}} + \frac{|g|^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)}{E_{\mathrm{tx}}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}} = \boldsymbol{0}.$$

Solving this equation for $\boldsymbol{F}$ leads to the following expression for the robust linear precoder

$$\boldsymbol{F} = \frac{1}{g}\left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}\right)^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}} \tag{5.10}$$

where we introduced the quantity $\xi = \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})/E_{\mathrm{tx}}$ already defined in Eq. (3.22). Note that the matrix inside the parentheses is always positive definite since $\frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\mathrm{tx}}} > 0$, and therefore the inverse always exists.

Now, bearing in mind Eq. (5.10), the following equality holds

$$|g|^2 \boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}} = \left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}\right)^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\hat{\boldsymbol{H}}\left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}\right)^{-1}.$$

Applying the trace operator to the above equation, and considering that $\operatorname{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) = E_{\mathrm{tx}}$, we finally have that the optimum value for the receive weight g must obey

$$|g|^2 = \frac{\operatorname{tr}\left(\left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}\right)^{-2}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{u}}\hat{\boldsymbol{H}}\right)}{E_{\mathrm{tx}}}. \tag{5.11}$$

Finally, if we assume that g is real and positive, the optimum value for g is unique and we arrive at

$$g_{\text{Rlin}} = \sqrt{\frac{\operatorname{tr}\left(\left(\hat{\boldsymbol{H}}^{\text{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}\right)^{-2}\hat{\boldsymbol{H}}^{\text{H}}\boldsymbol{C}_{\boldsymbol{u}}\hat{\boldsymbol{H}}\right)}{E_{\text{tx}}}}. \tag{5.12}$$

Therefore, the optimum robust linear precoder that solves the optimization problem in Eq. (5.3) reads as

$$\boldsymbol{F}_{\text{Rlin}} = \frac{1}{g_{\text{Rlin}}}\left(\hat{\boldsymbol{H}}^{\text{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}_N\right)^{-1}\hat{\boldsymbol{H}}^{\text{H}}. \tag{5.13}$$

It is interesting to compare Eq. (5.13) for the optimum robust linear precoder with Eq. (3.24) corresponding to the conventional linear precoder. We also see that the statistical structure and the magnitude of CSI errors have a strong influence on the final precoder. For very small errors, i.e. $\boldsymbol{C}_{\boldsymbol{\Theta}} \rightarrow \mathbf{0}$, we obtain the classical linear MMSE precoder (TxWF) as in Eq. (3.24) and for very large CSI errors, we get $\boldsymbol{F}_{\text{Rlin}} \rightarrow \frac{1}{g_{\text{Rlin}}}(\boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}_N)^{-1}\hat{\boldsymbol{H}}^{\text{H}}$, i.e. $\boldsymbol{F}_{\text{Rlin}}$ acts in a similar manner to a matched filter (TxMF) as in Eq. (3.31). Note that the regularization with $\boldsymbol{C}_{\boldsymbol{\Theta}}$ due to the robust design is not necessarily diagonal, since no assumption was made that the entries of the error $\boldsymbol{\Theta}[q]$ in Eq. (5.1) are uncorrelated. Thus, not only the amount of error but also the structural properties of the error have an impact on the precoder.

The optimum robust linear precoding parameters $\boldsymbol{F}_{\text{Rlin}}$ and g_{Rlin} can be expressed in a more compact way if we define the regularization matrix

$$\boldsymbol{T} = \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi\mathbf{I}_N \in \mathbb{C}^{N\times N} \tag{5.14}$$

and the positive definite matrix

$$\boldsymbol{\Phi} = \left(\hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}} + \mathbf{I}_K\right)^{-1} \in \mathbb{C}^{K\times K}. \tag{5.15}$$

Indeed, applying the matrix inversion lemma (see Appendix B.1) to Eq. (5.13) we have that

$$\begin{aligned}\left(\hat{\boldsymbol{H}}^{\text{H}}\hat{\boldsymbol{H}} + \boldsymbol{T}\right)^{-1}\hat{\boldsymbol{H}}^{\text{H}} &= \left(\boldsymbol{T}^{-1} - \boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\left(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\right)^{-1}\hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\right)\hat{\boldsymbol{H}}^{\text{H}} \\ &= \boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\left(\boldsymbol{I} - \left(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\right)^{-1}\hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\right) \\ &= \boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\left(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\right)^{-1} = \boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\boldsymbol{\Phi}.\end{aligned} \tag{5.16}$$

Thus, Eqs. (5.12) and (5.13) corresponding to the scalar weight g_{Rlin} and the linear robust

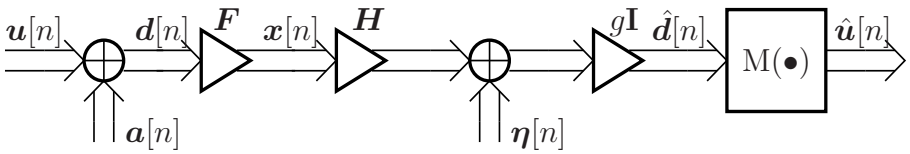


Figure 5.2: MU–MISO System with Vector Precoding.

precoder $\boldsymbol{F}_{\text{Rlin}}$, respectively, can be rewritten as

$$\boxed{\begin{aligned}\boldsymbol{F}_{\text{Rlin}} &= \frac{1}{g_{\text{Rlin}}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\boldsymbol{\Phi} \\ g_{\text{Rlin}} &= \sqrt{\frac{\operatorname{tr}\left(\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\text{H}}\boldsymbol{\Phi}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{\Phi}\hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\right)}{E_{\text{tx}}}}.\end{aligned}} \tag{5.17}$$

5.3 MU-MISO Robust Wiener Vector Precoding

Fig. 5.2 plots the block diagram of a MU–MISO system with Vector Precoding (VP). In vector precoding, the transmitter adds a perturbation signal $\boldsymbol{a}[n] \in \tau\mathbb{Z}^K + \mathrm{j}\,\tau\mathbb{Z}^K$ to the data signal $\boldsymbol{u}[n] \in \mathbb{C}^K$ prior to linear transformation with the filter $\boldsymbol{F} \in \mathbb{C}^{N\times K}$. At the receiver, the symbols are scaled with the common weight g and then passed through a modulo operator. This modulo operator enables the addition of the perturbation signal by the transmitter, since the same output can be generated by different inputs for the modulo operator. The constant associated with the modulo operator is denoted by τ. We consider the transmission of one block of data symbols of length N_{B} during which the scaling factor g is constant. The data symbols of the block $u[1], \ldots, u[N_{\text{B}}]$ are known at the transmitter.

Following similar steps as in Subsection 3.3.1, the freedom of adding $\boldsymbol{a}[n]$ is optimally exploited by VP whose robust MMSE optimization reads as

$$\begin{aligned}\{\boldsymbol{a}_{\text{RVP}}[n], \boldsymbol{x}_{\text{RVP}}[n], g_{\text{RVP}}\} = \underset{\{\boldsymbol{a}[n],\boldsymbol{x}[n],g\}}{\operatorname{argmin}} \; & \mathrm{E}_{\boldsymbol{\Theta}}\left[\frac{1}{N_{\text{B}}}\sum_{n=1}^{N_{\text{B}}}\mathrm{E}\left[\left\|\boldsymbol{d}[n]-\hat{\boldsymbol{d}}[n]\right\|_2^2 |\boldsymbol{u}[n]\right]\right] \\ \text{s.t.:} \quad & \frac{1}{N_{\text{B}}}\sum_{n=1}^{N_{\text{B}}}\|\boldsymbol{x}[n]\|_2^2 \le E_{\text{tx}}\end{aligned} \tag{5.18}$$

where $\hat{\boldsymbol{d}}[n]$ are the scaled received symbols at the users given by

$$\hat{\boldsymbol{d}}[n] = g\boldsymbol{H}\boldsymbol{x}[n] + g\boldsymbol{\eta}[n]. \tag{5.19}$$

The aim of VP is to choose the *virtual desired symbols*

$$\boldsymbol{d}[n] = \boldsymbol{u}[n] + \boldsymbol{a}[n] \tag{5.20}$$

for $n = 1, \ldots, N_{\mathrm{B}}$, so that $\hat{\boldsymbol{d}}[n]$ approximates $\boldsymbol{d}[n]$ according to the MMSE criterion, as expressed in Eq. (5.18). Note that the expectation is conditioned on the full knowledge of the symbols $\boldsymbol{u}[n]$ at the transmitter. However, since the statistics of the perturbation vector $\boldsymbol{a}[n]$ are unknown, we average the symbol MSE over the whole block. And to robustify the optimization, an expectation over the CSI errors $\boldsymbol{\Theta}$ was introduced [cf. Eq. (5.18) with Eq. (3.64)].

The MSE in Eq. (5.18) can be expressed as

$$\begin{aligned}\varepsilon\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) &= \mathrm{E}_{\boldsymbol{\Theta}}\left[\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\mathrm{E}\left[\left\|\boldsymbol{d}[n]-\hat{\boldsymbol{d}}[n]\right\|_2^2 \middle| \boldsymbol{u}[n]\right]\right] \\ &= \frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\big(\, \boldsymbol{d}^{\mathrm{H}}[n]\boldsymbol{d}[n] - g^*\boldsymbol{x}^{\mathrm{H}}[n]\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{d}[n] - g\boldsymbol{d}^{\mathrm{H}}[n]\hat{\boldsymbol{H}}\boldsymbol{x}[n] \\ &\quad + |g|^2\,\boldsymbol{x}^{\mathrm{H}}[n]\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{x}[n] + |g|^2\,\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{x}[n] + |g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\big) \end{aligned} \tag{5.21}$$

where $\mathrm{E}[\|\boldsymbol{d}[n]\|_2^2\,|\boldsymbol{u}[n]] = \|\boldsymbol{d}[n]\|_2^2$ and $\mathrm{E}[\|\boldsymbol{x}[n]\|_2^2\,|\boldsymbol{u}[n]] = \|\boldsymbol{x}[n]\|_2^2$ were applied.

We form the Lagrangian function as

$$L\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g, \lambda\right) = \varepsilon\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) + \lambda\left(\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{x}[n] - E_{\mathrm{tx}}\right) \tag{5.22}$$

where $\lambda \in \mathbb{R}^{0,+}$. Now, we set its derivative with respect to $\boldsymbol{x}^*[n], n = 1, \ldots, N_{\mathrm{B}}$, and g^* to zero, which leads to the KKT conditions

$$\begin{aligned}\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{x}^*[n]} &= \frac{1}{N_{\mathrm{B}}}\left(-g^*\boldsymbol{H}^{\mathrm{H}}\boldsymbol{d}[n] + |g|^2\,\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{x}[n] + |g|^2\,\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{x}[n]\right) \\ &\quad + \frac{\lambda}{N_{\mathrm{B}}}\boldsymbol{x}[n] = \boldsymbol{0} \end{aligned} \tag{5.23}$$

$$\begin{aligned}\frac{\partial L\left(\bullet\right)}{\partial g^*} &= \frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\Big(-\boldsymbol{x}^{\mathrm{H}}[n]\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{d}[n] + g\boldsymbol{x}^{\mathrm{H}}[n]\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{x}[n] \\ &\quad + g\,\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{x}[n]\Big) + g\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) = 0 \end{aligned} \tag{5.24}$$

$$\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{x}[n] \leq E_{\mathrm{tx}}$$

$$\lambda\left(\frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\boldsymbol{x}^{\mathrm{H}}[n]\boldsymbol{x}[n] - E_{\mathrm{tx}}\right) = 0 \quad \text{with } \lambda \geq 0. \tag{5.25}$$

From the first KKT condition, we conclude that the transmit symbols are given by

$$\boldsymbol{x}[n] = \frac{1}{g}\left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \frac{\lambda}{|g|^2}\mathbf{I}\right)^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{d}[n]. \tag{5.26}$$

Then, we have to show that $\lambda > 0$, i.e. the power constraint is active. We can rewrite the second KKT condition in Eq. (5.24) by equating to zero $\frac{\partial L(\bullet)}{\partial g^*} g^*$. This produces the result

$$|g|^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}}) = \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \left(g^* \boldsymbol{x}^{\mathrm{H}}[n] \hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{d}[n] - |g|^2 \boldsymbol{x}^{\mathrm{H}}[n] \hat{\boldsymbol{H}}^{\mathrm{H}} \hat{\boldsymbol{H}} \boldsymbol{x}[n] \right.$$
$$\left. - |g|^2 \boldsymbol{x}^{\mathrm{H}}[n] \boldsymbol{C}_{\boldsymbol{\Theta}} \boldsymbol{x}[n] \right). \tag{5.27}$$

On the other hand, multiplying Eq. (5.23) by $\boldsymbol{x}^{\mathrm{H}}[n]$ from the left and summing over $n = 1, \ldots, N_{\mathrm{B}}$, yields

$$\frac{\lambda}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{x}^{\mathrm{H}}[n] \boldsymbol{x}[n] = \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \left(g^* \boldsymbol{x}^{\mathrm{H}}[n] \hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{d}[n] - |g|^2 \boldsymbol{x}^{\mathrm{H}}[n] \hat{\boldsymbol{H}}^{\mathrm{H}} \hat{\boldsymbol{H}} \boldsymbol{x}[n] \right.$$
$$\left. - |g|^2 \boldsymbol{x}^{\mathrm{H}}[n] \boldsymbol{C}_{\boldsymbol{\Theta}} \boldsymbol{x}[n] \right). \tag{5.28}$$

By combining Eqs. (5.27) and (5.28), we obtain that the value for the Lagrangian multiplier

$$\lambda = |g|^2 \frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{\frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{x}^{\mathrm{H}}[n] \boldsymbol{x}[n]}.$$

Therefore, it becomes clear that $\lambda > 0$ for the non–trivial case that $\exists n : \boldsymbol{x}[n] \neq \boldsymbol{0}$. Hence, the transmit energy constraint is always active and we define

$$\xi = \frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\mathrm{tx}}}$$

and correspondingly, $\lambda = |g|^2 \xi$.

Bearing in mind that the transmit energy constraint in Eq. (5.18) is active and taking into account the expressions for the transmit symbols in Eq. (5.26), we reach the following solution for the robust WF-VP:

$$\boldsymbol{x}_{\mathrm{RVP}}[n] = \frac{1}{g_{\mathrm{RVP}}} \left(\hat{\boldsymbol{H}}^{\mathrm{H}} \hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi \mathbf{I} \right)^{-1} \hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{d}[n]$$
$$g_{\mathrm{RVP}} = \sqrt{\frac{\sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{d}^{\mathrm{H}}[\boldsymbol{n}] \hat{\boldsymbol{H}} \left(\hat{\boldsymbol{H}}^{\mathrm{H}} \hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \xi \mathbf{I} \right)^{-2} \hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{d}[n]}{E_{\mathrm{tx}} N_{\mathrm{B}}}}. \tag{5.29}$$

Equivalently, we can reformulate the above equations using the matrices $\boldsymbol{T}$ and $\boldsymbol{\Phi}$ already defined in Eqs. (5.14) and (5.15), so Eq. (5.29) can be rewritten as

$$\boxed{\begin{aligned} \boldsymbol{x}_{\mathrm{RVP}}[n] &= \frac{1}{g_{\mathrm{RVP}}} \boldsymbol{T}^{-1} \hat{\boldsymbol{H}}^{\mathrm{H}} \boldsymbol{\Phi} \boldsymbol{d}[n] \\ g_{\mathrm{RVP}} &= \sqrt{\frac{\sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{d}^{\mathrm{H}}[n] \boldsymbol{\Phi} \hat{\boldsymbol{H}} \boldsymbol{T}^{-2} \hat{\boldsymbol{H}} \boldsymbol{\Phi} \boldsymbol{d}[n]}{E_{\mathrm{tx}} N_{\mathrm{B}}}}. \end{aligned}} \tag{5.30}$$

It is important to note that the scalar gain g_{RVP} is only chosen once in each block.

Next, we proceed to find the perturbation signal $\boldsymbol{a}[n]$. First, we plug the optimum transmit vectors $\boldsymbol{x}_{\text{RVP}}[n]$ and the optimum gain g_{RVP} into the MSE expression. Following similar steps as those taken to arrive at Eq. (3.73), we get

$$\varepsilon\left(\boldsymbol{a}[n], \boldsymbol{x}[n], g\right) = \frac{1}{N_{\text{B}}} \sum_{n=1}^{N_{\text{B}}} \boldsymbol{d}[n]^{\text{H}} \boldsymbol{\Phi} \boldsymbol{d}[n]. \tag{5.31}$$

Note that when $\boldsymbol{C}_{\boldsymbol{\Theta}} \rightarrow \mathbf{0}$ and, therefore, $\hat{\boldsymbol{H}} \rightarrow \boldsymbol{H}$, we obtain that $\varepsilon(\boldsymbol{a}[n], \boldsymbol{x}[n], g) \rightarrow \frac{\xi}{N_{\text{B}}} \sum_{n=1}^{N_{\text{B}}} \boldsymbol{d}^{\text{H}}[n](\boldsymbol{H}\boldsymbol{H}^{\text{H}} + \xi \mathbf{I})^{-1} \boldsymbol{d}[n]$, which corresponds to the MSE for the non–robust WF-VP approach shown in Subsection 3.3.1 for the case of perfect CSI available at the transmitter.

$\boldsymbol{\Phi} \leftarrow (\hat{\boldsymbol{H}} \boldsymbol{T}^{-1} \hat{\boldsymbol{H}}^{\text{H}} + \mathbf{I}_K)^{-1}$
factorize: $\boldsymbol{\Phi} \leftarrow \boldsymbol{L}^{\text{H}} \boldsymbol{D} \boldsymbol{L}$
for $n = 1, \ldots, N_{\text{B}}$:
  $a_{\text{RVP}}[n] \leftarrow \text{argmin}_{\boldsymbol{a}[n] \in \tau \mathbb{Z}^K + \text{j}\, \tau \mathbb{Z}^K} ||\boldsymbol{D}^{1/2} \boldsymbol{L} (\boldsymbol{u}[n] + \boldsymbol{a}[n])||_2^2$
  $\boldsymbol{x}[n] \leftarrow \boldsymbol{T}^{-1} \hat{\boldsymbol{H}}^{\text{H}} \boldsymbol{\Phi} (\boldsymbol{u}[n] + \boldsymbol{a}_{\text{RVP}}[n])$
$g_{\text{RVP}} \leftarrow \sqrt{\frac{1}{E_{\text{tx}} N_{\text{B}}} \sum_{n=1}^{N_{\text{B}}} \boldsymbol{x}^{\text{H}}[n] \boldsymbol{x}[n]}$
for $n = 1, \ldots, N_{\text{B}}$:
  $\boldsymbol{x}[n] \leftarrow g_{\text{RVP}}^{-1} \boldsymbol{x}[n]$

Table 5.1: Calculation of WF-VP Robust Filters.

Finally, taking into account that every summand of Eq. (5.31) can be minimized separately, the perturbation signal can be found via the closest point search in a lattice

$$\begin{aligned} \boldsymbol{a}_{\text{RVP}}[n] &= \underset{\boldsymbol{a}[n] \in \tau \mathbb{Z}^K + \text{j}\, \tau \mathbb{Z}^K}{\text{argmin}} (\boldsymbol{u}[n] + \boldsymbol{a}[n])^{\text{H}} \boldsymbol{\Phi} (\boldsymbol{u}[n] + \boldsymbol{a}[n]) \\ &= \underset{\boldsymbol{a}[n] \in \tau \mathbb{Z}^K + \text{j}\, \tau \mathbb{Z}^K}{\text{argmin}} ||\boldsymbol{D}^{1/2} \boldsymbol{L} (\boldsymbol{u}[n] + \boldsymbol{a}[n])||_2^2 \end{aligned} \tag{5.32}$$

where the second line in Eq. (5.32) is obtained after introducing the Cholesky decomposition $\boldsymbol{\Phi} = \boldsymbol{L}^{\text{H}} \boldsymbol{D} \boldsymbol{L}$, where $\boldsymbol{L}$ is a unit lower triangular matrix and $\boldsymbol{D}$ is a diagonal matrix. We will minimize the cost function in Eq. (5.32) without another constraint using the Schnorr–Euchner algorithm [87, 88]. Note that when large errors occur, $\boldsymbol{\Phi}$ is the identity matrix leading to $\boldsymbol{a}_{\text{RVP}}[n] = \mathbf{0}$, i.e. robust VP converges to robust linear precoding.

To summarize, in order to calculate the robust WF-VP we do the following: first, we factorize $\boldsymbol{\Phi} = (\hat{\boldsymbol{H}} \boldsymbol{T}^{-1} \hat{\boldsymbol{H}}^{\text{H}} + \mathbf{I}_K)^{-1}$ to find the perturbation vectors by means of the lattice search in Eq. (5.32); second, the unscaled transmit vectors are computed by means

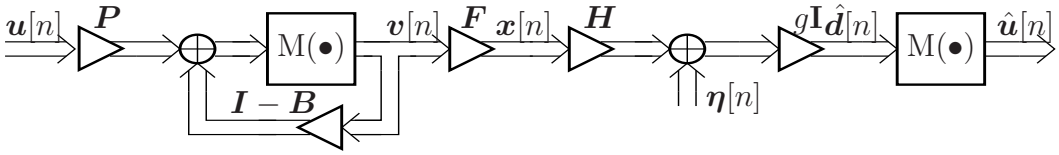


Figure 5.3: MU–MISO System with Tomlinson Harashima Precoding.

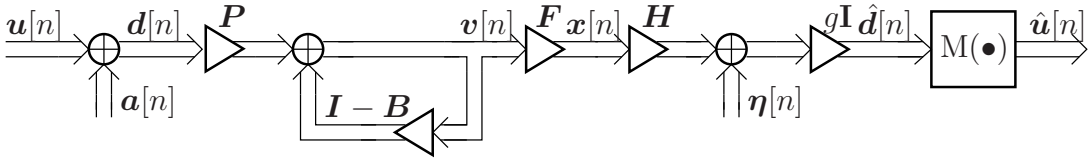


Figure 5.4: Linear Representation of Tomlinson Harashima Precoding.

of linear filtering; and third, the whole block is scaled with g_{RVP}. The pseudo code for the robust WF-VP is given in Table 5.1.

5.4 MU–MISO Robust Wiener Tomlinson–Harashima Precoding

Tomlinson–Harashima Precoding (THP) with partial CSI at the transmitter has been investigated by Fischer *et al.* [110], Simeone *et al.* [111], and Dietrich *et al* [18]. As discussed in Section 3.3, in order to avoid the high complexity of the robust VP rule in Eq. (5.32) we can employ THP as depicted in Fig. 5.3, where the perturbation signal $\boldsymbol{a}[n]$ is implicitly computed in a successive manner. The standard assumption for THP design is that the covariance matrix of the modulo operator output at the transmitter is diagonal [see Eq. (3.84)], i.e.

$$\boldsymbol{C}_{\boldsymbol{v}} = \mathrm{E}[\boldsymbol{v}[n]\boldsymbol{v}^{\mathrm{H}}[n]] = \mathrm{diag}\left(\sigma_{v,1}^2, \ldots, \sigma_{v,K}^2\right).$$

Additionally, the feedback filter $\boldsymbol{I}-\boldsymbol{B}$ must be strictly lower triangular. The optimization for robust THP can be expressed as [cf. Eq. (3.88)]

$$\begin{aligned}\{\boldsymbol{F}_{\text{RTHP}}, \boldsymbol{B}_{\text{RTHP}}, g_{\text{RTHP}}, \boldsymbol{P}_{\text{RTHP}}\} = \underset{\{\boldsymbol{F},\boldsymbol{B},g,\boldsymbol{P}\}}{\operatorname{argmin}} \; & \mathrm{E}_{\boldsymbol{\Theta}}\left[\mathrm{E}\left[\left\|\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v}[n] - g\boldsymbol{H}\boldsymbol{F}\boldsymbol{v}[n] - g\boldsymbol{\eta}[n]\right\|_2^2\right]\right] \\ \text{s.t.:} \quad & \mathrm{E}\left[\|\boldsymbol{x}[n]\|_2^2\right] \leq E_{\text{tx}} \quad \text{and} \quad \boldsymbol{B} \text{ is unit lower triangular}\end{aligned} \tag{5.33}$$

where $\boldsymbol{d}[n] = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v}[n]$ is the desired value for the inputs of the modulo operators at the receivers [see Eq. (3.86)], i.e. it is the sum of the symbols $\boldsymbol{u}[n]$ and the perturbation $\boldsymbol{a}[n]$ added by the modulo operator at the transmitter, as shown in Fig. 5.4. Remember that the

permutation matrix $\boldsymbol{P} = \sum_{k=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_{k_i}^{\mathrm{T}}$ satisfies $\boldsymbol{P}^{-1} = \boldsymbol{P}^{\mathrm{T}}$ with the index k_i of the i–th data stream to be precoded.

Similarly to Eq. (3.87), we have that

$$\begin{aligned}\varepsilon\left(\boldsymbol{P}, \boldsymbol{B}, \boldsymbol{F}, g\right) &= \mathrm{E}_{\boldsymbol{\Theta}}\left[\mathrm{E}\left[\left\|\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v}[n] - g\boldsymbol{H}\boldsymbol{F}\boldsymbol{v}[n] - g\boldsymbol{\eta}[n]\right\|_2^2\right]\right] \\ &= \mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) - g^{*}\,\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\right) - g\,\mathrm{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) \\ &\quad + |g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F}\right) + |g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F}\right) + |g|^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\end{aligned} \tag{5.34}$$

where the model for errors described at the beginning of this chapter is implicitly applied.

The restriction for the unit lower triangular structure of $\boldsymbol{B}$ in Eq. (5.33) can be expressed as

$$\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i = \boldsymbol{S}_i\boldsymbol{e}_i, \quad i = 1, \ldots, K$$

where $\boldsymbol{S}_i$ is the selection matrix $\boldsymbol{S}_i = [\boldsymbol{I}_i, \boldsymbol{0}_{i\times(K-i)}]$ defined as in Eq. (3.89).

The Lagrangian function corresponding to the constrained optimization problem in Eq. (5.33) is

$$\begin{aligned}L\left(\boldsymbol{P}, \boldsymbol{B}, \boldsymbol{F}, g, \lambda, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K\right) &= \varepsilon\left(\boldsymbol{P}, \boldsymbol{B}, \boldsymbol{F}, g\right) + \lambda\left(\mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\right) - E_{\mathrm{tx}}\right) \\ &\quad + 2\Re\left(\sum_{i=1}^{K}\mathrm{tr}\left(\boldsymbol{\mu}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i - \boldsymbol{S}_i\boldsymbol{e}_i\right)\right)\right)\end{aligned} \tag{5.35}$$

with $\boldsymbol{\mu}_i \in \mathbb{C}^i, i = 1, \ldots, K$ and where $2\Re(\sum_{i=1}^{K}\mathrm{tr}(\boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i - \boldsymbol{S}_i\boldsymbol{e}_i))$ comes from the restriction concerning the unit lower triangular structure of feedback matrix $\boldsymbol{B}$.

The solution to Eq. (5.33) can be obtained by setting the derivatives of the Lagrangian $L(\boldsymbol{P}, \boldsymbol{B}, \boldsymbol{F}, g, \lambda, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K)$ with respect to $\boldsymbol{B}^{*}$, $\boldsymbol{F}^{*}$, and g to zero. The first necessary KKT condition is obtained when we equate the derivative with respect to $\boldsymbol{F}^{*}$ to zero, i.e.

$$\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{F}^{*}} = -g^{*}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} + |g|^2\,\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}} + |g|^2\,\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}} + \lambda\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}} = \boldsymbol{0} \tag{5.36}$$

with $\mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}) \leq E_{\mathrm{tx}}$ and $\lambda(\mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}) - E_{\mathrm{tx}}) = 0$ with $\lambda \geq 0$. The resulting optimum value for $\boldsymbol{F}$ is the following

$$\boldsymbol{F} = \frac{1}{g}\left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} + \frac{\lambda}{|g|^2}\boldsymbol{I}\right)^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}.$$

Next, let us demonstrate that the inequality constraint in Eq. (5.35) is always active, i.e. $\lambda > 0$. To this end, the derivative of the Lagrangian function with respect to g is

equated to zero, i.e.

$$\frac{\partial L(\bullet)}{\partial g} = -\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right) + g^{*}\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F}\right) + g^{*}\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F}\right) + g^{*}\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) = 0$$

and the optimum value for the gain factor g^{*} is accordingly

$$g^{*} = \frac{\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right)}{\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)}. \tag{5.37}$$

Now, we multiply Eq. (5.36) by $\boldsymbol{F}^{\mathrm{H}}$ from the right and apply the trace operator to get

$$\lambda\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) = g^{*}\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\right) - |g|^{2}\operatorname{tr}\left(\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\right).$$

Plugging Eq. (5.37) into the above equation, we obtain that

$$\begin{aligned}\lambda\operatorname{tr}\left(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}\right) &= \frac{\left|\operatorname{tr}\left(\boldsymbol{H}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right)\right|^{2}}{\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)} \\ &\quad - \frac{\left|\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right)\right|^{2}\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F}\right)}{\operatorname{tr}^{2}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)} \\ &= \frac{\left|\operatorname{tr}\left(\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{P}\right)\right|^{2}}{\operatorname{tr}^{2}\left(\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{\Theta}}\boldsymbol{F} + \boldsymbol{C}_{\boldsymbol{\eta}}\right)}\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) = |g|^{2}\operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right).\end{aligned}$$

Then, it is apparent that $\lambda > 0$ and the energy restriction is active, i.e. $\operatorname{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^{\mathrm{H}}) = E_{\mathrm{tx}}$ and $\lambda = |g|^{2}\xi$ with $\xi = \frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\mathrm{tx}}}$.

Applying the matrix inversion lemma to the above expression for the feedforward filter $\boldsymbol{F}$ and considering the above result for λ, the optimum feedforward filter can be rewritten as

$$\boldsymbol{F} = \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\left(\mathbf{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\right)^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B} = \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B} \tag{5.38}$$

with the matrices $\boldsymbol{T}$ and $\boldsymbol{\Phi}$ defined in Eqs. (5.14) and (5.15), respectively.

Finally, setting the derivative of Eq. (5.35) with respect to $\boldsymbol{B}^{*}$ to zero, we have the KKT condition

$$\frac{\partial L(\bullet)}{\partial \boldsymbol{B}^{*}} = \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} - g\boldsymbol{P}\hat{\boldsymbol{H}}\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{v}} + \sum_{i=1}^{K}\boldsymbol{S}_{i}^{\mathrm{T}}\boldsymbol{\mu}_{i}^{*}\boldsymbol{e}_{i}^{\mathrm{T}} = \mathbf{0}.$$

Plugging Eq. (5.38) into the above equation we have

$$\frac{\partial L(\bullet)}{\partial \boldsymbol{B}^*} = \boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} - \boldsymbol{P}\hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\left(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\right)^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} + \sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}}$$
$$= \boldsymbol{P}\left(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\right)^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{v}} + \sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}} = \boldsymbol{0}.$$

Therefore, the feedback filter $\boldsymbol{B}$ can be expressed as

$$\boldsymbol{B} = -\boldsymbol{P}\boldsymbol{\Phi}^{-1}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\boldsymbol{e}_i^{\mathrm{T}}\sigma_{v,i}^{-2} \tag{5.39}$$

where we exploited the diagonal structure of $\boldsymbol{C}_{\boldsymbol{v}}$ [see Eq. (3.84)].

Multiplying this equation by $\boldsymbol{S}_i$ from the left and by $\boldsymbol{e}_i$ from the right, we obtain

$$\boldsymbol{S}_i\boldsymbol{B}\boldsymbol{e}_i = -\boldsymbol{S}_i\boldsymbol{P}\left(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^*\sigma_{v,i}^{-2} = \boldsymbol{S}_i\boldsymbol{e}_i$$

where we used the property that $\boldsymbol{e}_j^{\mathrm{T}}\sigma_{v,j}^{-2}\boldsymbol{e}_i = 0, j \neq i$, and $\boldsymbol{e}_j^{\mathrm{T}}\sigma_{v,i}^{-2}\boldsymbol{e}_i = \sigma_{v,i}^{-2}$, otherwise. The above equation allows us to find the Lagrangian multipliers $\boldsymbol{\mu}_i^*$, $i = 1, \ldots, K$, which are given by

$$\boldsymbol{\mu}_i^* = -\sigma_{v,i}^2\left(\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{\Phi}^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i. \tag{5.40}$$

Now, we can substitute $\boldsymbol{\mu}_i^*$ of Eq. (5.40) into the expression obtained for the feedback filter $\boldsymbol{B}$ in Eq. (5.39), and the resulting expression for $\boldsymbol{B}$ into the expression for the feedforward filter $\boldsymbol{F}$ in Eq. (5.38). We obtain

$$\boldsymbol{F} = \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{\Phi}^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}$$
$$\boldsymbol{B} = \boldsymbol{P}\boldsymbol{\Phi}^{-1}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{P}\boldsymbol{\Phi}^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}. \tag{5.41}$$

With the symmetrically permuted Cholesky factorization proposed in [92]

$$\boldsymbol{P}\boldsymbol{\Phi}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{P}\left(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\right)^{-1}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{L} \tag{5.42}$$

where $\boldsymbol{L}$ is unit lower triangular and $\boldsymbol{D}$ is non-negative diagonal, we can rewrite the

feedforward and feedback filter expressions in Eq. (5.41) as

$$\begin{aligned}
\boldsymbol{F} &= \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{L}^{-1}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{D}^{-1}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} \\
&= \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\left(\boldsymbol{S}_i\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{S}_i\boldsymbol{D}^{-1}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{S}_i\boldsymbol{L}^{-1}\boldsymbol{S}_i^{\mathrm{T}}\right)^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} \\
&= \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{D}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} \\
&= \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{D}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} = \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} \\
&= \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\sum_{i=1}^{K}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}} = \frac{1}{g}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}
\end{aligned}$$

and

$$\boldsymbol{B} = \boldsymbol{L}^{-1}\boldsymbol{D}^{-1}\boldsymbol{L}^{-\mathrm{H}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D} = \boldsymbol{L}^{-1} \tag{5.43}$$

respectively. To obtain this result, we used the properties for the selection matrix $\boldsymbol{S}_i$ given by Eq. (3.98).

In summary, we can conclude that the solution to Eq. (5.33) for a given permutation matrix $\boldsymbol{P}$ can be concisely written as

$$\boxed{\begin{aligned}\boldsymbol{F}_{\text{RTHP}} &= \frac{1}{g_{\text{RTHP}}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D} \\ \boldsymbol{B}_{\text{RTHP}} &= \boldsymbol{L}^{-1}\end{aligned}} \tag{5.44}$$

where g_{RTHP} follows from $\mathrm{tr}(\boldsymbol{F}_{\text{RTHP}}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{F}_{\text{RTHP}}^{\mathrm{H}}) = E_{\text{tx}}$, i.e.

$$\boxed{g_{\text{RTHP}} = \sqrt{\frac{\mathrm{tr}\left(\boldsymbol{P}\hat{\boldsymbol{H}}\boldsymbol{T}^{-2}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{D}\boldsymbol{L}\right)}{E_{\text{tx}}}}.} \tag{5.45}$$

Plugging the above results into Eq. (5.34), we arrive at an expression similar to Eq. (3.101), namely

$$\varepsilon\left(\boldsymbol{P},\boldsymbol{B},\boldsymbol{F},g\right) = \xi\sum_{i=1}^{K}\sigma_{v,i}^2 d_{i,i}. \tag{5.46}$$

To avoid the tough combinatorial optimization with respect to the permutation matrix $\boldsymbol{P}$, we propose a greedy optimization based on the MSE expression in Eq. (5.46). In the

i–th step, the index k_i is chosen such that the respective MSE is minimized, i.e. $d_{i,i}$ is minimized. Fortunately, this minimization can easily be included in the Cholesky factorization in Eq. (5.42). This can be seen in Table 5.2, which is very similar to Table 3.3. Note that only $\boldsymbol{\Phi}$ is defined differently and the feedforward filter computation is changed.

$$
\begin{aligned}
&\boldsymbol{\Phi} \leftarrow \left(\mathbf{I} + \hat{\boldsymbol{H}}\boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\right)^{-1} \\
&\boldsymbol{P} \leftarrow \mathbf{I}_K,\ \boldsymbol{D} \leftarrow \mathbf{0}_{K\times K} \\
&\textbf{for } i = K, \ldots, 1 \\
&\quad q \leftarrow \underset{q'=1,\ldots,i}{\operatorname{argmin}} \boldsymbol{\Phi}(q', q') \\
&\quad \boldsymbol{P}_i \leftarrow \mathbf{I}_K \text{ whose } i\text{-th and } q\text{-th rows are exchanged} \\
&\quad \boldsymbol{P} \leftarrow \boldsymbol{P}_i\boldsymbol{P} \\
&\quad \boldsymbol{\Phi} \leftarrow \boldsymbol{P}_i\boldsymbol{\Phi}\boldsymbol{P}_i^{\mathrm{T}} \\
&\quad \boldsymbol{D}(i,i) \leftarrow \boldsymbol{\Phi}(i,i) \\
&\quad \boldsymbol{\Phi}(1:i,i) \leftarrow \boldsymbol{\Phi}(1:i,i)/\boldsymbol{D}(i,i) \\
&\quad \boldsymbol{\Phi}(1:i-1,1:i-1) \leftarrow \boldsymbol{\Phi}(1:i-1,1:i-1) \\
&\qquad -\boldsymbol{\Phi}(1:i-1,i)\boldsymbol{\Phi}(1:i-1,i)^{\mathrm{H}}\boldsymbol{D}(i,i) \\
&\boldsymbol{L}^{\mathrm{H}} \leftarrow \text{upper triangular part of } \boldsymbol{\Phi} \\
&\boldsymbol{B} \leftarrow \boldsymbol{L}^{-1},\ \boldsymbol{F} \leftarrow \boldsymbol{T}^{-1}\hat{\boldsymbol{H}}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D}
\end{aligned}
$$

Table 5.2: Calculation of WF-THP Robust Filters with Ordering.

5.5 MMSE Receive Weights

We use a very simple receiver model for the precoder design where all receivers apply the same real scalar weight contrary to [18]. This assumption ensures closed-form solutions for the precoders. As was demonstrated in [18], the phase correction at the receivers is especially crucial for a system with erroneous CSI at the transmitter. In that case, the receivers must correct the wrong amplitudes and phases of the received signals due to the errors in the CSI at the transmitter. This objective is achieved by using MMSE receive coefficients. So, our system design is as follows. Based on the partial CSI, the transmitter designs the precoders under the assumption that the receivers apply the same weight and have the same errors in their CSI as the transmitter. This conservative assumption of the receivers is compensated by the application of the MMSE receive weights introduced in this section. Therefore, we have a slight mismatch between the receive weights model arising from the precoder design and the MMSE weights used instead.

In order to obtain this scalar MMSE coefficient for user k, we formulate a general

MSE $\varepsilon(g_k)$ depending on the scalar weight g_k:

$$\begin{aligned}\varepsilon\left(g_k\right) = \mathrm{E}\left[|x_k - g_k y_k|_2^2\right] &= \mathrm{E}\left[x_k x_k^* - g_k^* x_k y_k^* - g_k y_k x_k^* + g_k y_k y_k^* g_k^*\right] \\ &= c_{xx,k} - g_k^* c_{xy,k} - g_k c_{yx,k} + |g_k|^2 c_{yy,k}\end{aligned}$$

where x_k and y_k are, respectively, the desired and the received signals. The correlation coefficients are defined as $c_{xx,k} = \mathrm{E}[|x_k|^2]$ and $c_{yy,k} = \mathrm{E}[|y_k|^2]$, and the crosscorrelation between the received signal and the desired signal is obtained as $c_{xy,k} = \mathrm{E}[x_k y_k^*]$. When we compute the derivative with respect to g_k^* and set it to zero, we obtain

$$\frac{\partial \varepsilon\left(g_k\right)}{\partial g_k^*} = -c_{xy,k} + g_k c_{yy,k} = 0$$

which leads to the linear MMSE coefficient for user k given by

$$g_{\text{MMSE},k} = c_{xy,k} c_{yy,k}^{-1}. \tag{5.47}$$

The estimation of $c_{yy,k}$ is straightforward, i.e. it can be found via averaging over time, but the estimation of $c_{xy,k}$ is more delicate and depends on the type of precoder. Therefore, we distinguish the type of precoder to obtain the crosscorrelation $c_{xy,k}$ in the following subsections.

5.5.1 MMSE Weights for MU–MISO Linear Precoding

As mentioned above, MMSE receive weights are used instead of the weights directly obtained from the MSE optimization in Eq. (5.3) to correct the mismatch of the phase and the amplitude caused by the non-perfect CSI available at the transmitter. For user k the crosscorrelation $c_{xy,k}$ in Eq. (5.47) between the desired signal $u_k[n]$, where $u_k[n]$ denotes the k–th element of the transmitted symbols $\boldsymbol{u}[n]$ corresponding to the k–th user, and the received signal given by

$$y_k[n] = \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{F}_{\text{Rlin}} \boldsymbol{u}[n] + \eta_k[n]$$

is expressed as

$$c_{xy,k} = \mathrm{E}\left[u_k[n] y_k^*[n]\right] = \mathrm{E}\left[u_k[n]\left(\boldsymbol{u}^{\mathrm{H}}[n] \boldsymbol{F}_{\text{Rlin}}^{\mathrm{H}} \boldsymbol{h}_k^* + \eta_k^*[n]\right)\right] = \boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{F}_{\text{Rlin}}^{\mathrm{H}} \boldsymbol{h}_k^* \tag{5.48}$$

where we assume that $\boldsymbol{C}_{\boldsymbol{u}} = \mathrm{E}[\boldsymbol{u}[n]\boldsymbol{u}^{\mathrm{H}}[n]] = \mathbf{I}_K$ and $\boldsymbol{e}_k$ denotes the k–th column of the identity matrix $\mathbf{I}_K$.

On the other hand, the variance of the received signal $c_{yy,k}$ in Eq. (5.47) is simply given by

$$\begin{aligned}c_{yy,k} &= \mathrm{E}\left[|y_k[n]|^2\right] = \mathrm{E}\left[\left(\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{F}_{\text{Rlin}} \boldsymbol{u}[n] + \eta_k[n]\right)\left(\boldsymbol{u}^{\mathrm{H}}[n] \boldsymbol{F}_{\text{Rlin}}^{\mathrm{H}} \boldsymbol{h}_k^* + \eta_k^*[n]\right)\right] \\ &= \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{F}_{\text{Rlin}} \boldsymbol{F}_{\text{Rlin}}^{\mathrm{H}} \boldsymbol{h}_k^* + \sigma_{\eta,k}^2\end{aligned}$$

where $\sigma_{\eta,k}^2$ is the noise variance for the user k and we used $\boldsymbol{C}_{\boldsymbol{u}} = \mathbf{I}_K$.

Therefore, the MMSE receive weight is given by

$$g_{\text{MMSE-LP},k} = \boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{F}_{\text{Rlin}}^{\mathrm{H}} \boldsymbol{h}_k^* \left(\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{F}_{\text{Rlin}} \boldsymbol{F}_{\text{Rlin}}^{\mathrm{H}} \boldsymbol{h}_k^* + \sigma_{\eta,k}^2 \right)^{-1}. \tag{5.49}$$

5.5.2 MMSE Weights for MU–MISO Vector Precoding

Similarly to the case of the robust linear precoder explained in the previous subsection, it is more appropriate to use MMSE receive coefficients distinct from the common gain assumed in the VP design in Section 5.3 due to their superior performance in the presence of CSI errors at the transmitter. Remember that the desired signal for the k–th user in the case of VP is $u_k[n] + a_{\text{RVP},k}[n]$ and that the received signal is obtained as $y_k[n] = \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{x}_{\text{RVP}}[n] + \eta_k[n]$. Therefore, the crosscorrelation $c_{xy,k}$ in Eq. (5.47) is

$$\begin{aligned}
c_{xy,k} &= \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \mathrm{E}\left[\left(u_k[n] + a_{\text{RVP},k}[n] \right) y_k[n]^* \middle| \boldsymbol{u}[n] \right] \\
&= \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \mathrm{E}\left[\left(u_k[n] + a_{\text{RVP},k}[n] \right) \left(\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{x}_{\text{RVP}}^*[n] + \eta_k^*[n] \right) \middle| \boldsymbol{u}[n] \right] \\
&= \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{x}_{\text{RVP}}^*[n] \left(u_k[n] + a_{\text{RVP},k}[n] \right) \\
&= \boldsymbol{h}_k^{\mathrm{H}} \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{x}_{\text{RVP}}^*[n] \left(u_k[n] + a_{\text{RVP},k}[n] \right)
\end{aligned} \tag{5.50}$$

where we average over the whole block of N_{B} symbols since the statistics of $\boldsymbol{a}_{\text{RVP}}[n]$ are unknown. Moreover, the variance of the received signal is easily obtained as

$$\begin{aligned}
c_{yy,k} = \mathrm{E}\left[|y_k|^2 \right] &= \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \mathrm{E}\left[\left(\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{x}_{\text{RVP}}[n] + \eta_k[n] \right) \left(\boldsymbol{x}_{\text{RVP}}^{\mathrm{H}}[n] \boldsymbol{h}_k^* + \eta_k^*[n] \right) \middle| \boldsymbol{u}[n] \right] \\
&= \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \left| \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{x}_{\text{RVP}}[n] \right|^2 + \sigma_{\eta,k}^2.
\end{aligned} \tag{5.51}$$

Thus, the MMSE coefficients are given by

$$g_{\text{MMSE-VP},k} = \left(\boldsymbol{h}_k^{\mathrm{H}} \frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \boldsymbol{x}_{\text{RVP}}^*[n] d_k[n] \right) \left(\frac{1}{N_{\mathrm{B}}} \sum_{n=1}^{N_{\mathrm{B}}} \left| \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{x}_{\text{RVP}}[n] \right|^2 + \sigma_{\eta,k}^2 \right)^{-1} \tag{5.52}$$

where $d_k[n] = u_k[n] + a_{\text{RVP},k}[n]$.

5.5.3 MMSE Weights for MU–MISO Tomlinson–Harashima Precoding

Following a similar reasoning as in linear and vector robust precoding, once again it is preferable to use the MMSE coefficients at the receiver rather than the real weight g_{RTHP} given by Eq. (5.45). This is because the MMSE coefficients are capable of correcting not only the amplitude but also the phase mismatch caused by the errors in the CSI when computing the TH precoder filters. In THP, the received signal for the user k is $y_k[n] = \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{F}_{\text{RTHP}} \boldsymbol{v}[n] + \eta_k[n]$, while the desired signal is given by $\boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B}_{\text{RTHP}} \boldsymbol{v}[n]$. Therefore, the crosscorrelation $c_{xy,k}$ between these desired and received signals is expressed as follows

$$\begin{aligned} c_{xy,k} &= \mathrm{E}\left[\left(\boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B}_{\text{RTHP}} \boldsymbol{v}[n] \right) \left(\boldsymbol{v}^{\mathrm{H}}[n] \boldsymbol{F}_{\text{RTHP}}^{\mathrm{H}} \boldsymbol{h}_k^* + \eta_k^*[n] \right) \right] \\ &= \boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B}_{\text{RTHP}} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}_{\text{RTHP}}^{\mathrm{H}} \boldsymbol{h}_k^* \end{aligned} \tag{5.53}$$

whereas the variance of the received signal is

$$c_{yy,k} = \mathrm{E}\left[|y_k|^2 \right] = \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{F}_{\text{RTHP}} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}_{\text{RTHP}}^{\mathrm{H}} \boldsymbol{h}_k^* + \sigma_{\eta,k}^2.$$

Thus, the MMSE coefficients to be used together with the robust THP schemes are

$$g_{\text{MMSE-THP},k} = \left(\boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{B}_{\text{RTHP}} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}_{\text{RTHP}}^{\mathrm{H}} \boldsymbol{h}_k^* \right) \left(\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{F}_{\text{RTHP}} \boldsymbol{C}_{\boldsymbol{v}} \boldsymbol{F}_{\text{RTHP}}^{\mathrm{H}} \boldsymbol{h}_k^* + \sigma_{\eta,k}^2 \right)^{-1}. \tag{5.54}$$

5.6 Training data

It is important to point out that the proposed system with robust precoding is based on two training signals that must be sent frequently.

First, common pilot signals must be transmitted from the transmit antennas to enable an estimation of the vector channels at the single-antenna receivers (see Section 4.1). With these vector channel estimates, the receivers find the channel covariance matrices via time averaging. Since the covariance matrices only change slowly, the feedback of the information that allows us to know the channel covariance matrix at the transmitter does not cost much data rate. We assume, however, that the channel statistics are perfectly known at both the transmitter and the receiver side. Based on the fed–back and erroneous CSI, the transmitter is able to perform a robust precoder design as described in this chapter.

Second, distinct dedicated pilot signals must be sent to each receiver to allow an estimation of the overall precoder and channel filter. This estimate is necessary for the design of the MMSE receivers which correct the phase and the amplitude of the received signal [18]. Clearly, the receivers are unable to directly estimate this quantity, since they know neither the precoder nor the channel. Therefore, the training symbols are precoded

such that the total channel, i.e. the combination of the channel and the precoder for the pilot symbols, is equal to $c_{xy,k}$. For example, the vector-valued precoder for the dedicated pilot channel of robust VP is $\boldsymbol{p}_{\text{RVP,pilot},k} = \frac{1}{N_{\text{B}}} \sum_{n=1}^{N_{\text{B}}} \boldsymbol{x}_{\text{RVP}}[n](u_k^*[n] + a_{\text{RVP},k}^*[n])$, which gives, after transmission over $\boldsymbol{h}_k^{\text{T}}$, the complex–conjugate of the desired $c_{xy,k}$ in Eq. (5.50). Similarly to the VP case, the complex–conjugate of the desired $c_{xy,k}$ in Eqs. (5.48) and (5.53) is obtained with the vector–valued precoders $\boldsymbol{p}_{\text{Rlin,pilot},k} = \boldsymbol{F}_{\text{Rlin}}\boldsymbol{e}_k$ and $\boldsymbol{p}_{\text{RTHP,pilot},k} = \boldsymbol{F}_{\text{RTHP}}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{B}_{\text{RTHP}}^{\text{H}}\boldsymbol{P}\boldsymbol{e}_k$ for robust linear precoding and robust THP, respectively.

5.7 Conclusions

In this chapter, we have derived the expressions for robust linear and nonlinear Wiener precoders in order to compensate the mismatch between the true channel and the erroneous channel at the transmitter when no full channel information is available. We have mathematically derived the precoder filters for linear precoding, THP, and VP, taking into account the errors in the CSI introduced by channel estimation, rank reduction, quantization, and feedback delay discussed in Chapter 4. Basically, the robust designs incorporate an error matrix $\boldsymbol{C}_{\boldsymbol{\Theta}}$ (the so–called *regularization matrix*) into the final expressions obtained for the precoders, which enables us to avoid the enormous degradation in performance with increasing errors in CSI when non-robust schemes are used. We have also introduced receive MMSE weights different from the weights resulting from the optimizations with the objective of correcting the mismatch in amplitude and phase due to non-perfect CSI. This leads to different weights for each user, which clearly must be better than a common gain for all users.

An optimized limited feedback design combined with these robust schemes will lead to an improvement of the BER performance with limited feedback rate. This is in fact an open issue at present and is still the subject of much research. The objective of the following chapters is to find the best limited feedback design to be used together with the robust precoders of this chapter so as to make it possible to implement these limited feedback schemes in practice.

Chapter 6

Feedback Design based on CSI MSE

In the vector *broadcast channel* (BC), the centralized transmitter has more degrees of freedom than the receivers. Therefore, it is appropriate to separate the signals by applying precoding at the transmitter. To be able to design precoding, the transmitter needs knowledge about the channel states of the different receivers. In the case of *Frequency Division Duplex* (FDD) systems, this knowledge can be obtained by feedback (at least partially), where the *Channel State Information* (CSI) of the receiver is quantized to meet the limited rate conditions of the feedback channel. This feedback channel is assumed to be error-free, but it introduces a delay. CSI can be obtained by different mechanisms at the receiver side, which gives rise to a greater or lesser amount of degradation in the final information sent through the feedback channel. Each user estimates the channel and reduces it to a low-dimensional representation for data compression that is possible due to the channel correlations. Before the feedback, the CSI is quantized and only the index of the codebook entry is sent to the transmitter. Fig. 6.1 plots the block diagram of the limited feedback channel including the different steps for channel estimation, coefficient truncation, quantization, and feedback delay.

The standard assumption for feedback design is error-free CSI at the receivers [1–5]. The receivers, however, get their CSI after estimation and thus it contains errors. In this case, a feedback design based on mutual information is difficult to achieve [112, 113] and we therefore resort to a design based on the precoding MSE.

In this chapter, we propose some limited feedback designs with the objective of optimizing the quality of the CSI exploiting the low data rate of the feedback channel as efficiently as possible. These feedback designs also enable us to obtain adequate statistical characterizations of the errors in the fed–back CSI that lead to closed-form expressions for the resulting robust precoders. In particular, we provide three limited feedback designs based on a CSI–MSE metric, i.e. on the mismatch between the true channel and the erroneous channel eventually available at the transmitter. The proposals are the following.

- A preliminary design of the limited feedback that consists of a basic modeling of errors that is quite intuitive but less than optimum when it comes to minimizing the CSI-MSE metric.
- A non-Bayesian error modeling of errors based on an MSE feedback optimization. This MSE optimization based on a CSI metric does not include the quantizer parameters. For simplicity, the quantizer is the uniform quantizer explained in Chapter 4 and the errors due to the quantization process are considered as additive and independent noise that follows a uniform distribution.
- A Bayesian modeling of errors based on a joint MSE feedback optimization. Now, the MSE optimization based on the CSI also includes the parameters of the quantizer (partition cells and codebook entries), i.e. the MSE optimization involves the parameters for whole system, which leads to an optimum feedback design. Another novel issue is that the closed-form solutions obtained for the error matrices of the robust design are based on a Bayesian modeling of errors. Since the delayed channel versions fed back to the transmitter after estimation and truncation and also the quantization errors are assumed to be Gaussian distributed, unlike in the previous approaches, we can obtain the expressions for the probability density function of the channel vector conditional on the delayed, truncated, and quantized channel estimate according to a Bayesian framework. The number of bits allocated to quantize each channel vector resulting from truncation and estimation processes also has an important influence on the final performance achieved by the proposed limited feedback design. In this sense, we propose an algorithm to allocate the bits in real time so as to minimize the final MSE with a negligible increase in terms of computational complexity.

Note that all the schemes proposed above are based on separate optimizations of feedback and precoding, i.e. there is no unique and joint optimization that also includes the design of the precoder filters. Such a design is proposed in Chapter 7. Therefore, we first optimize the limited feedback and then the optimum precoders are designed taking into account the errors of the optimum CSI sent from the receivers.

6.1 Preliminary Design of Limited Feedback

We start by estimating the channels at the receivers using the observations of the pilot symbols. Then, we project the resulting channel estimation onto the eigenvectors of the channel covariance matrix to obtain the Karhunen-Loève transformation of the channel vector which optimally provides a dimensionality reduction with the smallest possible MSE (see Section 4.2). The coefficients of the truncated KL expansion are then quantized

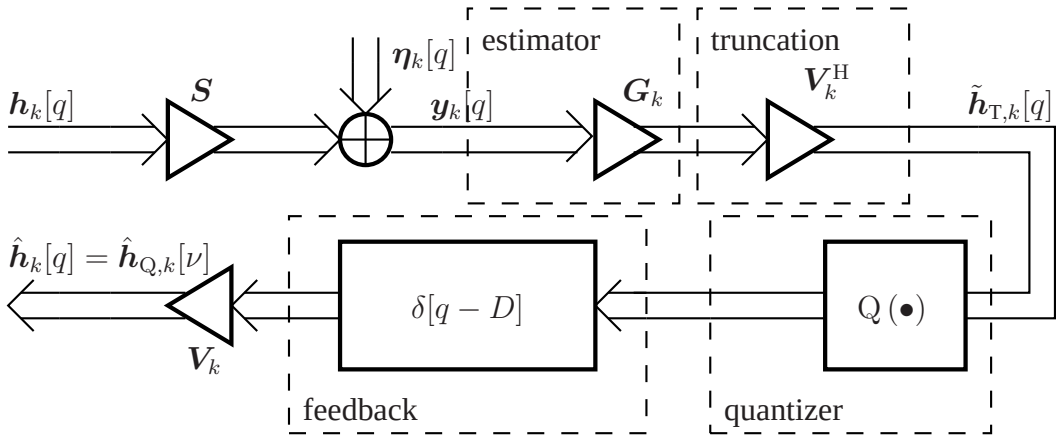


Figure 6.1: Model of Limited Feedback: Channel Estimation, Truncation, Quantization, and Feedback Delay.

prior to transmission over the feedback channel, which introduces a delay. We incorporate this delay into our model, considering a system without feedback delay but a delayed observation for the channel estimator. This partial CSI is then used at the transmitter to reconstruct the channel vector and to design the precoders. The proposed limited feedback design models the errors in a quite intuitive way but it is obvious that the MSE is not minimized, since no MSE optimization is performed in order to optimize the fed–back CSI. Nevertheless, this approach is quite illustrative due to its simplicity and its acceptable performance in terms of BER, as shown later by means of computer simulations.

In the following subsections, we describe this process in more detail and obtain the statistical description of the errors incurred at each step. Throughout this section we will assume that the signals and errors are uncorrelated.

6.1.1 Estimator and Rank Reduction Designs

As shown in [114, 115], we use linear estimators at the receiver based on $N_{\mathrm{tr}} \times N$ pilot symbols per time slot q to enable the channel vector estimation for the k-th user. We use the least-squares estimator explained in Subsection 4.1.1 of Chapter 4. According to Eq. (4.4), the least-squares channel estimate is obtained when we consider the estimator $\boldsymbol{G}_{\text{LS-estim},k} = \boldsymbol{S}^{\dagger} = (\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S})^{-1}\boldsymbol{S}^{\mathrm{H}}$. Therefore, the channel estimate is given by

$$\hat{\boldsymbol{h}}_{\mathrm{es},k}[q] = \boldsymbol{S}^{\dagger}\boldsymbol{y}_k[q] = \boldsymbol{h}_k[q] + \boldsymbol{S}^{\dagger}\boldsymbol{\eta}_k[q] = \boldsymbol{h}_k[q] + \boldsymbol{\eta}_{\mathrm{es},k}[q] \tag{6.1}$$

where [see Eq. (4.1)]

$$\boldsymbol{y}_k[q] = \boldsymbol{S}\boldsymbol{h}_k[q] + \boldsymbol{\eta}_k[q] \in \mathbb{C}^{N_{\mathrm{tr}}}$$

with $\boldsymbol{S} \in \mathbb{C}^{N_{\text{tr}} \times N}$ containing the training symbols for all users, $\boldsymbol{h}_k[q] \in \mathbb{C}^N$ as the channel vector for user k, and $\boldsymbol{\eta}_k[q] \in \mathbb{C}^{N_{\text{tr}}}$ being the AWGN with variance σ_η^2. We also have that [cf. Eq. (4.6)]

$$\boldsymbol{\eta}_{\text{es},k}[q] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma_\eta^2 (\boldsymbol{S}^{\text{H}} \boldsymbol{S})^{-1}) \tag{6.2}$$

since we assume $\boldsymbol{\eta}_k[q] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma_\eta^2 \mathbf{I})$.

The transmission over the feedback channel, however, introduces a delay of $D = q - \nu$ slots. This delay can equivalently be modeled as follows. The estimator gets outdated training data, i.e. the observation of the estimator is delayed by D slots. Then, the respective feedback channel has no delay. In other words, the precoder is designed during the time slot q and the channel estimate is obtained during the time slot $\nu = q - D$. Thus, the channel estimate for delayed training data reads as

$$\hat{\boldsymbol{h}}_{\text{es},k}[\nu] = \boldsymbol{h}_k[\nu] + \boldsymbol{\eta}_{\text{es},k}[\nu] \tag{6.3}$$

where $\boldsymbol{\eta}_{\text{es},k}[\nu]$ has the statistical properties described in Eq. (6.2). Clearly, $\hat{\boldsymbol{h}}_{\text{es},k}[\nu]$ can be rewritten as

$$\hat{\boldsymbol{h}}_{\text{es},k}[\nu] = \boldsymbol{h}_k[q] + \boldsymbol{h}_k[\nu] - \boldsymbol{h}_k[q] + \boldsymbol{\eta}_{\text{es},k}[\nu] = \boldsymbol{h}_k[q] + \boldsymbol{\eta}'_{\text{es},k}[\nu] \tag{6.4}$$

being $\boldsymbol{\eta}'_{\text{es},k}[\nu] = \boldsymbol{h}_k[\nu] - \boldsymbol{h}_k[q] + \boldsymbol{n}_{\text{es},k}[\nu]$. With the temporal correlation properties of $\boldsymbol{h}_k[q]$ [see Eq. (4.24)], remember that

$$\boldsymbol{C}_{\boldsymbol{h},k}[D] = E[\boldsymbol{h}_k[q] \boldsymbol{h}_k^{\text{H}}[\nu]] = \text{J}_0(\alpha_k D)\, \boldsymbol{C}_{\boldsymbol{h},k} \tag{6.5}$$

with $\alpha_k = 2\pi f_{\text{D,max},k} / f_{\text{slot}}$, where $\text{J}_0(\bullet)$ denotes the zero–th order Bessel function of the first kind, $f_{\text{D,max},k}$ is the *maximum* Doppler frequency for user k, and f_{slot} is the slot rate. Thus, we obtain with $\boldsymbol{e}_D = \boldsymbol{h}_k[\nu] - \boldsymbol{h}_k[q]$

$$E[\boldsymbol{e}_D \boldsymbol{e}_D^{\text{H}}] = 2\boldsymbol{C}_{\boldsymbol{h},k} - E[\boldsymbol{h}_k[\nu] \boldsymbol{h}_k^{\text{H}}[q]] - E[\boldsymbol{h}_k[q] \boldsymbol{h}_k^{\text{H}}[\nu]] = 2\,(1 - r_k)\, \boldsymbol{C}_{\boldsymbol{h},k} \tag{6.6}$$

where $r_k = J_0(\alpha_k D)$.

Hence, the new LS estimation error has the property

$$\boldsymbol{\eta}'_{\text{es,k}}[\nu] \sim \mathcal{N}_{\mathbb{C}} \left(\boldsymbol{0}, \boldsymbol{C}' \right) \tag{6.7}$$

with $\boldsymbol{C}' = \sigma_\eta^2 (\boldsymbol{S}^{\text{H}} \boldsymbol{S})^{-1} + 2(1 - r_k) \boldsymbol{C}_{\boldsymbol{h},k}$. Note that $\boldsymbol{\eta}'_{\text{es,k}}[\nu]$ is correlated with $\boldsymbol{h}_k[q]$. Nevertheless, we will neglect this correlatedness and assume that $\boldsymbol{h}_k[q]$ and the error $\boldsymbol{\eta}'_{\text{es,k}}[\nu]$ are uncorrelated.

After channel estimation, restrictions on the data rate of the feedback channel force us to compress the CSI to be sent to the transmitter from the users through the feedback channel. The *Karhunen-Loève* (KL) decomposition that enables us to obtain the rank

reduction basis $\boldsymbol{V}_k \in \mathbb{C}^{N\times d}$ from the channel covariance matrix was described in Section 4.2. Remember that the channel coefficients are given by [see Eq. (4.13)]

$$\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu] = \boldsymbol{V}_k^{\mathrm{H}}\hat{\boldsymbol{h}}_{\mathrm{es},k}[\nu] \in \mathbb{C}^d.$$

Therefore, from Eq. (6.4) we obtain that the estimate of the channel vector after rank–reduction in the time slot $\nu = q - D$ is

$$\hat{\boldsymbol{h}}_{\mathrm{T},k}[\nu] = \boldsymbol{V}_k\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu] = \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\hat{\boldsymbol{h}}_{\mathrm{es},k}[\nu] = \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{h}_k[q] + \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\eta}_{\mathrm{es},k}^{'}[\nu] \tag{6.8}$$

with the reduction basis $\boldsymbol{V}_k \in \mathbb{C}^{N\times d}$ assumed to be known at the transmitter. Note that the noise $\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\eta}_{\mathrm{es},k}^{'}[\nu]$ and the signal $\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{h}_k[q]$ lie in the same subspace spanned by the columns of $\boldsymbol{V}_k$. Therefore, $\hat{\boldsymbol{h}}_{\mathrm{T},k}[\nu]$ gives us no information about the properties of $\boldsymbol{h}_k[q]$ lying in the orthogonal subspace $\boldsymbol{V}_k^{\perp}$. This information was lost during the rank–reduction. The resulting error contribution due to the KL truncation reads as

$$\boldsymbol{\eta}_{\mathrm{KL},k}[q] = (\boldsymbol{I} - \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}})\boldsymbol{h}_k[q] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, (\boldsymbol{I} - \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}})\boldsymbol{C}_{\boldsymbol{h},k}(\boldsymbol{I} - \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}})). \tag{6.9}$$

Note that $\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\eta}_{\mathrm{es},k}^{'}[\nu]$ is orthogonal to $\boldsymbol{\eta}_{\mathrm{KL},k}[q]$ because $\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\eta}_{\mathrm{es},k}^{'}[\nu]$ lies in the subspace spanned by the columns of $\boldsymbol{V}_k$ and $\boldsymbol{\eta}_{\mathrm{KL},k}[q]$ lies in the subspace orthogonal to the columns of $\boldsymbol{V}_k$, since the covariance matrix of $\boldsymbol{\eta}_{\mathrm{KL},k}[q]$ is $(\boldsymbol{I} - \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}})\boldsymbol{C}_{\boldsymbol{h},k}(\boldsymbol{I} - \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}})$ [107]. Thus, we have that the CSI available at the transmitter in the time slot q (neglecting the quantization) is given by $\hat{\boldsymbol{h}}_k[q]$ obtained as

$$\hat{\boldsymbol{h}}_k[q] = \hat{\boldsymbol{h}}_{\mathrm{noQ},k}[\nu] = \hat{\boldsymbol{h}}_{\mathrm{T},k}[\nu] + \boldsymbol{\eta}_{\mathrm{KL},k}[q] \tag{6.10}$$

with

$$\hat{\boldsymbol{h}}_{\mathrm{T},k}[\nu] = \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{h}_k[q] + \boldsymbol{\eta}_{\mathrm{T},k}[\nu] \tag{6.11}$$

where [see Eq. (6.8)]

$$\boldsymbol{\eta}_{\mathrm{T},k}[\nu] = \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\eta}_{\mathrm{es},k}^{'}[\nu] \sim \mathcal{N}_{\mathbb{C}}\left(\boldsymbol{0}, \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\left(\sigma_\eta^2(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S})^{-1} + 2(1-r_k)\boldsymbol{C}_{\boldsymbol{h},k}\right)\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\right). \tag{6.12}$$

6.1.2 Quantizer Design

The uniform quantizer is the most common of the scalar quantizers. Note that even though the input is Gaussian and not uniform, we can assume that the input PDF is very smooth if the number of levels for uniform quantization is large (or equivalently, the quantizer step is very small). Therefore, the analysis of uniform quantization is simple [108] and the use of uniform quantizers gives reasonably good performance with the enormous advantage of simplicity in terms of practical implementation. The principle of the scalar

quantizer was explained in Subsection 4.3.1. The KL coefficients $\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu]$ of the rank reduced channel estimate are unfortunately not uncorrelated. Nevertheless, we assume they are uncorrelated and white (i.e. $\mathrm{E}[\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu]] = \mathbf{0}$ and $\mathrm{E}[\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu]\tilde{\boldsymbol{h}}_{\mathrm{T},k}^{\mathrm{H}}[\nu]] = \mathbf{I}$) to be able to design a uniform quantizer independently from the channel statistics. This scalar uniform quantizer based on the assumption of white coefficients has the advantage of remaining unchanged for varying channel statistics. Additionally, we assume that the input is bounded, with real and imaginary parts independently quantized and lying in the range included between $-\sqrt{2}$ and $\sqrt{2}$, so the overload region has a very low probability (≤ 0.05) of containing any input sample as long as the input is distributed as a unit Gaussian distribution. The simplicity of the proposed quantizer enables us to store initial codebooks at both the transmitter and receiver sides that need not be adapted to changing channel conditions. These codebooks might not be common to all the users since each user can use a different number of bits per coefficient to send the CSI to the transmitter. However, the scalar uniform quantizer can be computed before transmission without being recomputed in real time (just switch between look–up tables). The computer simulations in Section 6.4 show how normalizing each coefficient by the corresponding entry of the diagonal matrix resulting from the KL factorization of the channel covariance matrix in order to obtain that $\mathrm{E}[\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu]\tilde{\boldsymbol{h}}_{\mathrm{T},k}^{\mathrm{H}}[\nu]] = \mathbf{I}$ improves the performance. The improvement in terms of BER is noticeable, especially for high SNR values, since we approximately get a unit variance Gaussian distribution. Afterwards, this operation is inverted at the output of the quantizer by multiplying the quantized channel version by this scaling factor available at the receivers.

The process of quantization is as follows. Before transmission, we design uniform quantizers with representants between $-\sqrt{2}$ and $\sqrt{2}$ for each user's coefficient that have different sizes (M_i) according to the importance of the channel coefficient to be quantized. The step size for the i–th coefficient is given by $\gamma_i = \frac{2\sqrt{2}}{M_i}$, which is assumed to be the same for both the real and imaginary part. This initial codebook is stored at both the transmitter and receiver sides. The receivers perform a search to find the element in the codebook that is closest to the real or imaginary inputs corresponding to the real or imaginary parts of the KL coefficients obtained at time slot $\nu = q - D$. Then, the corresponding codebook index is fed back to the transmitter. Finally, the transmitter simply looks into its codebook and builds the precoder parameters from the selected codeword [103].

We consider the following simple model for the quantization error

$$\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu] = \tilde{\boldsymbol{h}}_{\mathrm{Q},k}[\nu] + \tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}[\nu] \in \mathbb{C}^d \tag{6.13}$$

where $\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[\nu]$ is the quantized version of the rank-reduced channel coefficients and $\tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}[\nu]$ is the additive error introduced by the quantizer. Additionally, we assume that the quantization error $\tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}[\nu]$ is uniformly distributed within the cell corresponding to a codebook entry (neglecting the different cell size for the representants $\pm\sqrt{2}$). The

resulting error variance is $\gamma_i^2/12$ for the real or imaginary part of the i–th coefficient [108]. Assuming additionally that the KL coefficients are uncorrelated with the zero–mean quantization errors, we get for the covariance matrix of the quantization noise $\tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}[\nu]$ of user k

$$\boldsymbol{C}_{\tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}} = \mathrm{E}\left[\tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}[\nu]\tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}^{\mathrm{H}}[\nu]\right] = \mathrm{diag}\left(\frac{\gamma_1^2}{6},\ldots,\frac{\gamma_d^2}{6}\right) \in \mathbb{R}^{d\times d} \tag{6.14}$$

where $\gamma_i^2/6$ with $i = 1,\ldots,d$, also comes from the fact that the quantization errors for each real and imaginary input are independent and given by $\gamma_i^2/12$ [108].

Finally, bearing in mind Eq. (6.13), we have

$$\hat{\boldsymbol{h}}_{\mathrm{T},k}[\nu] = \boldsymbol{V}_k\tilde{\boldsymbol{h}}_{\mathrm{T},k}[\nu] = \boldsymbol{V}_k\left(\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[\nu] + \tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}[\nu]\right) = \hat{\boldsymbol{h}}_{\mathrm{Q},k}[\nu] + \boldsymbol{\eta}_{\mathrm{Q},k}[\nu] \in \mathbb{C}^N. \tag{6.15}$$

For notational brevity, we have introduced $\boldsymbol{\eta}_{\mathrm{Q},k}[\nu] = \boldsymbol{V}_k\tilde{\boldsymbol{\eta}}_{\mathrm{Q},k}[\nu]$. Therefore, we get the rank deficient covariance matrix for the quantization error

$$\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q},k}} = \mathrm{E}[\boldsymbol{\eta}_{\mathrm{Q},k}[\nu]\boldsymbol{\eta}_{\mathrm{Q},k}^{\mathrm{H}}[\nu]] = \boldsymbol{V}_k\,\mathrm{diag}\left(\frac{\gamma_1^2}{6},\ldots,\frac{\gamma_d^2}{6}\right)\boldsymbol{V}_k^{\mathrm{H}}. \tag{6.16}$$

In summary, Eqs. (6.10), (6.11), and (6.15) enable us to express the CSI at the transmitter side in the time slot q as

$$\begin{aligned}\hat{\boldsymbol{h}}_k[q] &= \hat{\boldsymbol{h}}_{\mathrm{Q},k}[\nu] = \hat{\boldsymbol{h}}_{\mathrm{noQ},k}[\nu] + \boldsymbol{\eta}_{\mathrm{Q},k}[\nu] \\ &= \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{h}_k[q] + \boldsymbol{\eta}_{\mathrm{T},k}[\nu] + \boldsymbol{\eta}_{\mathrm{KL},k}[q] + \boldsymbol{\eta}_{\mathrm{Q},k}[\nu]\end{aligned} \tag{6.17}$$

where $\boldsymbol{\eta}_{\mathrm{T},k}[\nu]$ is the error due to channel estimation [see Eq. (6.12)], $\boldsymbol{\eta}_{\mathrm{KL},k}[q]$ stands for the error due to truncation [see Eq. (6.9)], and $\boldsymbol{\eta}_{\mathrm{Q},k}[\nu]$ denotes the quantization error [see Eq. (6.16)].

6.1.3 MSE Error Matrix for Robust Multi-User Precoder Design

Robust designs have been explained in Chapter 5 for implementing precoding schemes in scenarios where no perfect CSI is available at the transmitter. On the other hand, the cause of this erroneous CSI has been introduced in Chapter 4 with a brief discussion about the error sources that have an influence on the CSI. As seen in Chapter 5, we can reduce the impairments of the channel state information at the transmitter side by introducing a regularization given by an additional matrix in the design of the different precoders as shown in Eqs. (5.17), (5.30), and (5.44). Next, we spell out this matrix of regularization against errors in CSI according to the limited feedback developed in this section.

Remember that in Eq. (5.1) we introduced the channel matrix model as follows

$$\boldsymbol{H}[q] = \hat{\boldsymbol{H}}[q] + \boldsymbol{\Theta}[q].$$

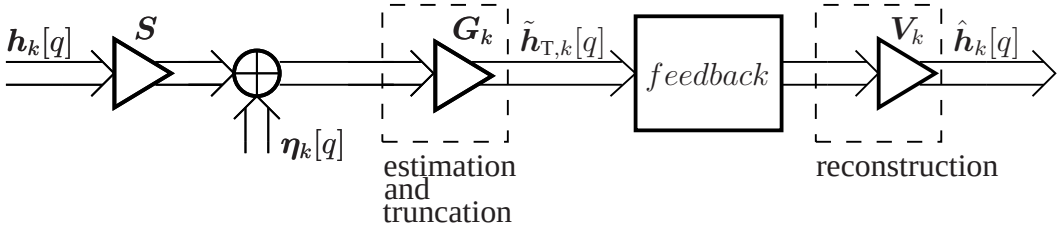


Figure 6.2: Feedback Design for a Non–Bayesian Error Modeling based on CSI MSE.

For the approach considered in this section, $\hat{\boldsymbol{H}}[q]$ is the quantized version of the channel matrix and $\boldsymbol{\Theta}[q]$ is the error matrix given by

$$\boldsymbol{C}_{\boldsymbol{\Theta}} = \sum_{k=1}^{K} \boldsymbol{C}^{*}_{\boldsymbol{\Theta},k}$$

where the error matrix for user k, which reduces the effects of the mismatch between the perfect and the imperfect CSI recovered at the transmitter given by Eq. (6.17), is expressed as

$$\begin{aligned}\boldsymbol{C}_{\boldsymbol{\Theta},k} = \boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}(\sigma_\eta^2(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S})^{-1} + 2\,(1-r_k)\,\boldsymbol{C}_{\boldsymbol{h},k})\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}} + (\boldsymbol{I}-\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}})\boldsymbol{C}_{\boldsymbol{h},k}(\boldsymbol{I}-\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}) \\ + \boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q},k}}.\end{aligned} \tag{6.18}$$

6.2 Non-Bayesian Error Modeling based on CSI MSE

In this section, we propose a joint MSE optimization of the channel estimation and the rank reduction basis, where the quantizer is modeled as a data independent additive noise source. Fig. 6.2 depicts the feedback model based on CSI MSE described in this section. Note that the quantizer is not explicitly shown since its parameters are not included into the MSE optimization. Interestingly, the resulting reduction basis is different from the eigenbasis of the channel covariance matrix (i.e. the Karhunen-Loève basis as used in the previous section). Besides the design of the components of the feedback system, the joint MSE optimization also delivers the error covariance matrix, which is necessary for a robust precoder design.

As in the last section, we start by estimating the channel at the receivers using the observations of different pilot symbols sent from the transmit antennas. Then, the estimate is reduced to a low-dimensional representation of the channel by projecting the estimate onto a basis which depends only on the statistics of the channel. The coefficients are then quantized prior to transmission over the feedback channel, which is assumed to be error-free but introduces a delay.

6.2.1 Estimator and Rank Reduction Designs

The feedback link introduces a delay that is modeled by means of estimation via an outdated version of the observation of the training channel. When only one past training period is considered, the receiver is as simple as possible, since only one observation vector has to be processed. However, the receiver could use several observation vectors to consider more than one outdated version of the channel, thereby improving the estimation quality. Accordingly, we can stack the channel information to be processed as follows [109, 116]

$$\bar{\boldsymbol{h}}_k[q] = \left[\boldsymbol{h}_k[q-D_1]^{\mathrm{T}}, \cdots, \boldsymbol{h}_k[q-D_L]^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{NL} \tag{6.19}$$

where L is the number of delayed vectors to be processed, and $D_i, i = 1, \ldots, L$, is the delay expressed as the number of slots for the i-th vector. We have that the crosscovariance matrix between the channel $\boldsymbol{h}_k[q]$ and the stacked channel $\bar{\boldsymbol{h}}_k[q]$ of Eq. (6.19) is given by [see Eq. (4.27)]

$$\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k} = \mathrm{E}\left[\bar{\boldsymbol{h}}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q]\right] = \boldsymbol{\beta}_k \otimes \boldsymbol{C}_{\boldsymbol{h},k} \in \mathbb{C}^{N\times NL} \tag{6.20}$$

where $\boldsymbol{\beta}_k = [\mathrm{J}_0(\alpha_k D_1), \ldots, \mathrm{J}_0(\alpha_k D_L)]^{\mathrm{T}} \in \mathbb{R}^L$ with $\alpha_k = 2\pi f_{\mathrm{D,max},k}/f_{\mathrm{slot}}$ (see Section 4.4). From Eq. (6.5), we have

$$\boldsymbol{C}_{\bar{\boldsymbol{h}},k} = \mathrm{E}\left[\bar{\boldsymbol{h}}_k[q]\bar{\boldsymbol{h}}_k^{\mathrm{H}}[q]\right] = \boldsymbol{C}_{\mathrm{temp},k} \otimes \boldsymbol{C}_{\boldsymbol{h},k} \in \mathbb{C}^{NL\times NL} \tag{6.21}$$

where the matrix $\boldsymbol{C}_{\mathrm{temp},k}$ comprises the temporal correlations and its i-th element in the j-th column is

$$[\boldsymbol{C}_{\mathrm{temp},k}]_{i,j} = \begin{cases} \mathrm{J}_0\left(\alpha_k\left(D_i - D_j\right)\right) & j \neq i, \\ 1 & j = i. \end{cases} \tag{6.22}$$

The final CSI recovered at the transmitter from $\bar{\boldsymbol{h}}_k[q]$ in Eq. (6.19) is given by

$$\hat{\boldsymbol{h}}_k[q] = \boldsymbol{V}_k\boldsymbol{G}_k\bar{\boldsymbol{S}}\bar{\boldsymbol{h}}_k[q] + \boldsymbol{V}_k\boldsymbol{G}_k\bar{\boldsymbol{\eta}}_k[q] + \boldsymbol{\eta}_{\mathrm{Q},k}[q] \tag{6.23}$$

where $\bar{\boldsymbol{S}} = \mathbf{I}_L \otimes \boldsymbol{S} \in \mathbb{C}^{N_{\mathrm{tr}}\times NL}$ contains the training symbols [1] and $\bar{\boldsymbol{\eta}}_k[q] \in \mathbb{C}^{N_{\mathrm{tr}}L} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\bar{\boldsymbol{\eta}},k})$ is the noise for channel estimation, being $\boldsymbol{C}_{\bar{\boldsymbol{\eta}},k} = \mathbf{I}_L \otimes \boldsymbol{C}_{\boldsymbol{\eta},k} \in \mathbb{C}^{N_{\mathrm{tr}}L\times N_{\mathrm{tr}}L}$. The zero–mean quantization error $\boldsymbol{\eta}_{\mathrm{Q},k}[q]$ is assumed to be uncorrelated with the input of the quantizer and has the covariance matrix $\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q},k}}[q]$ as given in Eq. (6.16). The filter $\boldsymbol{G}_k \in \mathbb{C}^{d\times N_{\mathrm{tr}}L}$ performs joint channel estimation and rank reduction and the rank reduction basis is given by $\boldsymbol{V}_k \in \mathbb{C}^{N\times d}$ (see Section 4.2).

The channel estimation and rank reduction with $\boldsymbol{G}_k$ together with the basis $\boldsymbol{V}_k$ are jointly optimized to end up with a channel estimate at the transmitter with minimum MSE

$$\{\boldsymbol{G}_{\mathrm{MMSE},k}, \boldsymbol{V}_{\mathrm{MMSE},k}\} = \underset{\{\boldsymbol{G}_k,\boldsymbol{V}_k\}}{\operatorname{argmin}} \mathrm{MSE}_k(\boldsymbol{G}_k, \boldsymbol{V}_k) \quad \text{s.t.: } \boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{V}_k = \mathbf{I}_d \tag{6.24}$$

[1] We assume the same training symbols in every time slot to simplify notation.

with the MSE of user k given by [see Eq. (6.23)]

$$\begin{aligned}\mathrm{MSE}_k(\boldsymbol{G}_k, \boldsymbol{V}_k) = \mathrm{E}\left[\left\|\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_k[q]\right\|_2^2\right] = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) + \mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{G}_k\boldsymbol{C}_{\bar{\boldsymbol{\eta}},k}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{V}_k^{\mathrm{H}}\right) \\ - 2\Re\left(\mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{G}_k\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}\right)\right) + \mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{G}_k\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}},k}\bar{\boldsymbol{S}}^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{V}_k^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}\right)\end{aligned} \tag{6.25}$$

where $\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}$ is the crosscovariance matrix of $\bar{\boldsymbol{h}}_k[q]$ and $\boldsymbol{h}_k[q]$ given by Eq. (6.20), $\boldsymbol{C}_{\bar{\boldsymbol{h}},k}$ is the covariance matrix of $\bar{\boldsymbol{h}}[q]$ in Eq. (6.21), and $\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}$ is the covariance matrix of the quantization error given in Eq. (6.16). In the optimization problem of Eq. (6.24) we also included the constraint for orthonormality of the columns of $\boldsymbol{V}_k$.

In order to solve the constrained optimization problem of Eq. (6.24), we construct the Lagrangian function

$$L\left(\boldsymbol{G}_k, \boldsymbol{V}_k, \boldsymbol{\Lambda}_k\right) = \mathrm{MSE}_k(\boldsymbol{G}_k, \boldsymbol{V}_k) + \mathrm{tr}\left(\boldsymbol{\Lambda}_k\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{V}_k - \mathbf{I}_d\right)\right) \tag{6.26}$$

where $\boldsymbol{\Lambda}_k \in \mathbb{C}^{d\times d}$ is the Lagrangian multiplier for the constraint. Note that $\boldsymbol{\Lambda}_k$ is Hermitian, since the constraint is Hermitian by definition.

The filter $\boldsymbol{G}_k$ is readily found by setting the derivative of the Lagrangian function in Eq. (6.26) with respect to $\boldsymbol{G}_k^*$ to zero, i.e.

$$\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{G}_k^*} = -\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}\bar{\boldsymbol{S}}^{\mathrm{H}} + \boldsymbol{G}_k\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}},k}\bar{\boldsymbol{S}}^{\mathrm{H}} + \boldsymbol{G}_k\boldsymbol{C}_{\bar{\boldsymbol{\eta}},k} = \mathbf{0}. \tag{6.27}$$

Here, we employed $\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{V}_k = \mathbf{I}$ for the last two terms. Therefore, the filter $\boldsymbol{G}_k$ is given by

$$\boldsymbol{G}_{\mathrm{MMSE},k} = \boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}^{\mathrm{H}}\bar{\boldsymbol{S}}^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}},k}\bar{\boldsymbol{S}}^{\mathrm{H}} + \boldsymbol{C}_{\bar{\boldsymbol{\eta}},k}\right)^{-1}. \tag{6.28}$$

Substituting the optimum $\boldsymbol{G}_{\mathrm{MMSE},k}$ into the cost function of Eq. (6.25) yields

$$\begin{aligned}\mathrm{MSE}_k(\boldsymbol{V}_k) = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) - \mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}^{\mathrm{H}}\bar{\boldsymbol{S}}^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}},k}\bar{\boldsymbol{S}}^{\mathrm{H}} + \boldsymbol{C}_{\bar{\boldsymbol{\eta}},k}\right)^{-1}\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\right) \\ + \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}\right).\end{aligned} \tag{6.29}$$

Now, the above optimization only depends on $\boldsymbol{V}_k$ and can be solved using Lagrangian multipliers. The Lagrangian function of Eq. (6.26) reduces to

$$\begin{aligned}L\left(\boldsymbol{V}_k, \boldsymbol{\Lambda}_k\right) = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) - \mathrm{tr}\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}^{\mathrm{H}}\bar{\boldsymbol{S}}^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}},k}\bar{\boldsymbol{S}}^{\mathrm{H}} + \boldsymbol{C}_{\bar{\boldsymbol{\eta}},k}\right)^{-1}\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}\boldsymbol{V}_k\right) \\ + \mathrm{tr}\left(\boldsymbol{\Lambda}\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{V}_k - \mathbf{I}_d\right)\right) + \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}\right)\end{aligned} \tag{6.30}$$

and by setting the derivative of the Lagrangian function with respect to $\boldsymbol{V}_k^*$ to zero we obtain

$$\frac{\partial L\left(\boldsymbol{V}_k, \boldsymbol{\Lambda}_k\right)}{\partial \boldsymbol{V}_k^*} = -\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}^{\mathrm{H}}\bar{\boldsymbol{S}}^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}},k}\bar{\boldsymbol{S}}^{\mathrm{H}} + \boldsymbol{C}_{\bar{\boldsymbol{\eta}},k}\right)^{-1}\bar{\boldsymbol{S}}\boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}\boldsymbol{V}_k + \boldsymbol{V}_k\boldsymbol{\Lambda}_k^{\mathrm{H}} = \mathbf{0}.$$

This equation can be rewritten as follows

$$\boldsymbol{W}_k \boldsymbol{V}_k = \boldsymbol{V}_k \boldsymbol{\Lambda}_k^{\mathrm{H}}. \tag{6.31}$$

where the matrix $\boldsymbol{W}_k$ is given by

$$\boldsymbol{W}_k = \boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k}^{\mathrm{H}} \bar{\boldsymbol{S}}^{\mathrm{H}} \left(\bar{\boldsymbol{S}} \boldsymbol{C}_{\bar{\boldsymbol{h}},k} \bar{\boldsymbol{S}}^{\mathrm{H}} + \boldsymbol{C}_{\bar{\boldsymbol{\eta}},k} \right)^{-1} \bar{\boldsymbol{S}} \boldsymbol{C}_{\bar{\boldsymbol{h}}\boldsymbol{h},k} \in \mathbb{C}^{N \times N_{\mathrm{tr}}}. \tag{6.32}$$

After multiplying by $\boldsymbol{V}_k^{\mathrm{H}}$ from the left, we see that $\boldsymbol{\Lambda}_k$ is not only Hermitian but also non-negative definite. Thus, the EVD (*EigenValue Decomposition*) of $\boldsymbol{\Lambda}_k$ is $\boldsymbol{\Lambda}_k = \boldsymbol{Q}_k \boldsymbol{\Phi}_k^2 \boldsymbol{Q}_k^{\mathrm{H}}$ with the unitary matrix $\boldsymbol{Q}_k \in \mathbb{C}^{d \times d}$ and the non-negative diagonal matrix $\boldsymbol{\Phi}_k^2 \in \mathbb{C}^{d \times d}$. Then, Eq. (6.31) can be rewritten as

$$\boldsymbol{W}_k \boldsymbol{V}_k' = \boldsymbol{V}_k' \boldsymbol{\Phi}_k^2 \tag{6.33}$$

where $\boldsymbol{V}_k' = \boldsymbol{V}_k \boldsymbol{Q}_k$ is a matrix with orthonormal columns as $\boldsymbol{V}_k$, since $\boldsymbol{Q}_k$ is unitary. Thus, we see that $\boldsymbol{\Lambda}_k$ in Eq. (6.31) can be replaced by a diagonal matrix $\boldsymbol{\Phi}_k^2$ without loss of generality. After multiplying Eq. (6.33) by $\boldsymbol{V}_k^{'\mathrm{H}}$ from the left, we have that

$$\boldsymbol{V}_k^{'\mathrm{H}} \boldsymbol{W}_k \boldsymbol{V}_k^{'} = \boldsymbol{\Phi}_k^2 \tag{6.34}$$

i.e. $\boldsymbol{V}_k^{'}$ is the matrix that diagonalizes $\boldsymbol{W}_k$. Thus, the columns of $\boldsymbol{V}_k^{'} = \boldsymbol{V}_k \boldsymbol{Q}_k$ are eigenvectors of $\boldsymbol{W}_k$ and not those of $\boldsymbol{C}_{\boldsymbol{h},k}$ as we intuitively used in Section 6.1. With this intermediate result for the rank reduction basis $\boldsymbol{V}_k^{'}$, the cost function of Eq. (6.24) is given by

$$\mathrm{MSE}_k = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) - \sum_{i \in \mathbb{I}} \varphi_{k,i}^2 + \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}\right) \tag{6.35}$$

where $\mathbb{I}$ denotes the set of eigenvectors indices collected in $\boldsymbol{V}_k \boldsymbol{Q}_k$ and $\varphi_{k,i}^2$ is the i-th eigenvalue of $\boldsymbol{W}_k$. Clearly, $\mathrm{MSE}_k(\boldsymbol{G}_{\mathrm{MMSE},k}, \boldsymbol{V}_k)$ is independent of $\boldsymbol{Q}_k$. Therefore, we can set $\boldsymbol{Q}_k = \mathbf{I}_d$ and $\boldsymbol{V}_k \in \mathbb{C}^{N \times d}$ contains d eigenvectors of $\boldsymbol{W}_k$. Note that since the rank reduction is focused on $\boldsymbol{V}_k^{\mathrm{H}}$, the bit allocation of each user can be decided off-line taking into account its maximum number of bits to be sent through the feedback channel. The term $\sum_{i \in \mathbb{I}} \varphi_{k,i}^2$ in Eq. (6.35) is fixed because it only depends on d. Taking into account that the former coefficient is larger than the latter, the allocation of a higher number of bits to the larger eigenvalue reduces the final MSE. Therefore, it can easily be seen that the term $\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k})$ in Eq. (6.35) is minimized when we distribute the total number of bits as uniformly as possible. For example, when we consider $d = 2$ and 12 bits per user, the best result corresponds to allocating 6 bits for each coefficient (i.e. 3 bits for each real or imaginary part), or, alternatively, when we have 10 bits per user the best choice is 6 and 4 bits for the first and second coefficient, respectively [2]. This counter–intuitive result

[2]The number of bits used for quantization must be two times a cardinal number to end up with a cardinal number of bits for the real and imaginary part.

follows from neglecting the correlations between the input and the output of the quantizer. Also note that no errors due to rank reduction are added to the channel estimation if all the eigenvectors are employed. The set of indices $\mathbb{I}$ must minimize the MSE in Eq. (6.35), i.e. the sum $\sum_{i\in\mathbb{I}} \varphi_{k,i}^2$ has to be maximized. Since the eigenvalues of $\boldsymbol{W}_k$ in Eq. (6.32) are sorted in non–increasing order, i.e. $\varphi_{k,1}^2 \geq \varphi_{k,2}^2 \geq \cdots \varphi_{k,N}^2$, the set $\mathbb{I}$ will contain the indices corresponding to the first d eigenvalues of $\boldsymbol{W}_k$.

In the following, we consider the special case with $L = 1$ and $D_1 = D$, i.e. only one observation vector in Eq. (6.19) is processed. Then, the above expressions are reduced as shown below [117]. Let q be the time slot corresponding to the design of the precoder and $\nu = q - D$ the time slot in which the outdated version of the channel estimate is obtained.

The cost function in Eq. (6.25) can be rewritten as follows

$$\begin{aligned}\mathrm{MSE}_k(\boldsymbol{G}_k, \boldsymbol{V}_k) &= \mathrm{E}\left[\left\|\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_k[q]\right\|_2^2\right] = \mathrm{E}\left[\left\|\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_{\mathrm{Q},k}[\nu]\right\|_2^2\right] \\ &= \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) + \mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{G}_k\boldsymbol{C}_{\boldsymbol{\eta},k}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{V}_k^{\mathrm{H}}\right) - 2\Re\left(\mathrm{tr}\left(r_k\boldsymbol{V}_k\boldsymbol{G}_k\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\right)\right) \\ &\quad + \mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{G}_k\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{V}_k^{\mathrm{H}}\right) + \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}\right) \end{aligned} \tag{6.36}$$

with $r_k = \mathrm{J}_0(\alpha_k D)$ as explained in Eq. (6.5). Following similar steps to those in the case $L > 1$, the filter $\boldsymbol{G}_k$ is readily found by setting the derivative of the cost function with respect to $\boldsymbol{G}_k^*$ to zero:

$$\boldsymbol{G}_{\mathrm{MMSE},k} = r_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1} = r_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{G}_{\text{MMSE-estim},k} \tag{6.37}$$

where it can be seen that $\boldsymbol{G}_{\mathrm{MMSE},k}$ has the ordinary MMSE channel estimator $\boldsymbol{G}_{\text{MMSE-estim},k}$ as the first stage. The term with the projection onto the basis $\boldsymbol{V}_k^{\mathrm{H}}$ produces uncorrelated outputs and the factor r_k is due to the inherent channel prediction. Substituting the optimum $\boldsymbol{G}_{\mathrm{MMSE},k}$ into the cost function of Eq. (6.36) yields

$$\mathrm{MSE}_k(\boldsymbol{V}_k) = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) - \mathrm{tr}\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{W}_k\boldsymbol{V}_k\right) + \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}\right) \tag{6.38}$$

with the non-negative definite matrix $\boldsymbol{W}_k$ given by

$$\boldsymbol{W}_k = r_k^2\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1}\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k} \in \mathbb{C}^{N\times N}. \tag{6.39}$$

Again, the matrix $\boldsymbol{V}_k$ diagonalizes $\boldsymbol{W}_k$ as follows

$$\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{W}_k\boldsymbol{V}_k = \boldsymbol{\Phi}_k^2 \tag{6.40}$$

which enables us to arrive at a similar result for the final MSE as obtained in Eq. (6.35), i.e.

$$\mathrm{MSE}_k = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) - \sum_{i\in\mathbb{I}}\varphi_{k,i}^2 + \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q}},k}\right) \tag{6.41}$$

although $\varphi_{k,i}^2$ is now the i-th entry of the diagonal matrix in Eq. (6.40) obtained from $\boldsymbol{W}_k$ in Eq. (6.39) and not from Eq. (6.32).

6.2.2 Quantizer Design

Our work is focused on scalar quantizers that allow us to obtain closed-form solutions for robust designs with low impact on the complexity required by the user's devices. However, we also implement a vector quantizer to be robust against errors in CSI, which has the enormous disadvantage of considerably increasing the computational complexity since it implies a larger dimensionality of the search than scalar quantizers. For scalar quantization, we simply use uniform quantizers, and for vector quantization, the approach termed *Random Vector Quantization* (RVQ) (see Chapter 4). Both quantizers as used for the design in this section are described below.

Uniform Quantization Error

The truncated channel coefficients can be quantized using the same scalar uniform quantizer as described in Subsection 6.1.2. Again, we make the assumption that the input is bounded but now, contrary to the previous approach, the rank reduced channel estimate is uncorrelated since its covariance matrix is diagonal, i.e. $\boldsymbol{C}_{\tilde{\boldsymbol{h}}_{\mathrm{T},k}} = \boldsymbol{\Phi}_k^2$. Thus, we consider that both real and imaginary parts of its i-th entry lie in the interval $[-\sqrt{2}\varphi_{k,i}, +\sqrt{2}\varphi_{k,i}]$. Each coefficient is normalized to ensure unit variance at the quantizer's input, multiplying the result of rank reduction by $\boldsymbol{\Phi}_k^{-1}$ before quantization. Multiplying the codebook entry again by $\boldsymbol{\Phi}_k$, we can fix the boundaries corresponding to a unit variance Gaussian distribution $[-\sqrt{2}, \sqrt{2}]$, as done in Subsection 6.1.2. As a result, each entry of the rank reduced channel estimate is standard Gaussian and this interval selection ensures that the overload probability is less than $5\,\%$. Then, a common uniform codebook with cell size $\gamma_i = 2\sqrt{2}/M_i$, where M_i is the size of the codebook, is stored at the transmitter and the codebooks remain unchanged throughout the transmission, even though the channel characteristics may suffer variations due to the wireless environment (see Chapter 2).

Random Vector Quantization Error

As described in Subsection 4.3.2, the delayed output $\hat{\boldsymbol{h}}_{\mathrm{es},k}[\nu = q - D]$ of the estimator is the input to the *random vector quantizer* (RVQ), where $\hat{\boldsymbol{h}}_{\mathrm{es},k}[\nu]$ is given by

$$\hat{\boldsymbol{h}}_{\mathrm{es},k}[\nu] = \boldsymbol{G}_{\text{MMSE-pred},k}\boldsymbol{S}\boldsymbol{h}_k[\nu] + \boldsymbol{G}_{\text{MMSE-pred},k}\boldsymbol{\eta}_k[\nu] \tag{6.42}$$

with $\boldsymbol{S} \in \mathbb{C}^{N_{\mathrm{tr}} \times N}$ containing the training data and $\boldsymbol{G}_{\text{MMSE-pred},k}$ being the MMSE predictor of Eq. (4.9) given by $\boldsymbol{G}_{\text{MMSE-pred},k} = r_k \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1} \in \mathbb{C}^{N \times N_{\mathrm{tr}}}$ where the factor r_k produced by the outdated estimation has been included. The random vector quantizer approximates $\hat{\boldsymbol{h}}_{\mathrm{es},k}[\nu]$ by one of the M entries $\boldsymbol{y}_{k,i}$, with $i = 1, \ldots, M$, by

minimizing the squared error as follows

$$i_{min} = \min_i \left\| \hat{\boldsymbol{h}}_{\text{es},k}[\nu] - \boldsymbol{y}_{k,i} \right\|_2^2 \tag{6.43}$$

where the codebook entries are such that

$$\boldsymbol{y}_{k,i} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{h},k}) \quad i = 1, \ldots, M. \tag{6.44}$$

Consequently, the error corresponding to the i–th codebook entry obeys

$$\boldsymbol{\epsilon}_i = \hat{\boldsymbol{h}}_{\text{es},k}[\nu] - \boldsymbol{y}_{k,i} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_{\hat{\boldsymbol{h}},k} + \boldsymbol{C}_{\boldsymbol{h},k}) \tag{6.45}$$

since the codebook generation is independent of the estimate $\hat{\boldsymbol{h}}_{\text{es},k}[\nu]$. Note that this distribution is independent of the index i. $\boldsymbol{C}_{\hat{\boldsymbol{h}},k}$ is the covariance matrix of $\hat{\boldsymbol{h}}_{\text{es},k}[\nu]$ in Eq. (6.42) given by

$$\boldsymbol{C}_{\hat{\boldsymbol{h}},k} = r_k^2 \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\text{H}} \left(\boldsymbol{S} \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\text{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1} \boldsymbol{S} \boldsymbol{C}_{\boldsymbol{h},k}. \tag{6.46}$$

Therefore, we assume that the structure of the error is given by $\boldsymbol{C}_{\hat{\boldsymbol{h}},k} + \boldsymbol{C}_{\boldsymbol{h},k}$ but weighted by a factor $\zeta_k \in \mathbb{R}^{0,+}$, i.e.

$$\boldsymbol{C}_{\text{Q},k} = \zeta_k \left(\boldsymbol{C}_{\hat{\boldsymbol{h}},k} + \boldsymbol{C}_{\boldsymbol{h},k} \right). \tag{6.47}$$

The factor ζ_k results from the selection process and is the ratio of the MSE with selection over the MSE without selection:

$$\zeta_k = \frac{\text{E}[\min_i ||\hat{\boldsymbol{h}}_{\text{es},k}[\nu] - \boldsymbol{y}_{k,i}||_2^2]}{\text{E}[||\hat{\boldsymbol{h}}_{\text{es},k}[\nu] - \boldsymbol{y}_{k,i}||_2^2]} = \frac{\text{E}[\min_i ||\hat{\boldsymbol{h}}_{\text{es},k}[\nu] - \boldsymbol{y}_{k,i}||_2^2]}{\text{tr}\left(\boldsymbol{C}_{\hat{\boldsymbol{h}},k} + \boldsymbol{C}_{\boldsymbol{h},k} \right)}. \tag{6.48}$$

Note that this model for the quantization error is conservative, i.e. the error is over–estimated, since we neglect the transmitter knowledge about the codebook entry that is selected and also about the structure of the corresponding partition cell, which is known since the codebook is stored at both the transmitter and the receivers.

In Appendix E we explain in detail how to solve the integral in the numerator of Eq. (6.48), which is a by no means trivial procedure.

6.2.3 MSE Error Matrix for Robust Multi-User Precoder Design

For the robust precoder design, we again interpret the channel as a random variable and the given fed–back CSI as deterministic, i.e.

$$\boldsymbol{H}[q] = \hat{\boldsymbol{H}}[q] + \boldsymbol{\Theta}[q]$$

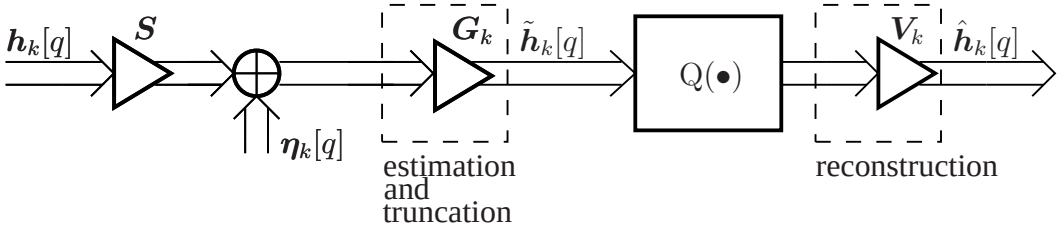


Figure 6.3: Feedback Design for a Bayesian Error Modeling based on CSI MSE.

where $\hat{\boldsymbol{H}}[q] = [\hat{\boldsymbol{h}}_1[q], \ldots, \hat{\boldsymbol{h}}_K[q]]^{\mathrm{T}} \in \mathbb{C}^{K \times N}$ comprises the channel estimates obtained from the quantized coefficients of the rank reduced channel that have been fed back. The covariance matrix of the error $\boldsymbol{\Theta}$ is

$$\boldsymbol{C}_{\boldsymbol{\Theta}} = \mathrm{E}\left[\boldsymbol{\Theta}^{\mathrm{H}}[q]\boldsymbol{\Theta}[q]\right] = \sum_{k=1}^{K} \boldsymbol{C}^{*}_{\boldsymbol{\Theta},k}$$

where $\boldsymbol{C}_{\boldsymbol{\Theta},k}$ is given by

$$\boldsymbol{C}_{\boldsymbol{\Theta},k} = \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k^2\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta}_{\mathrm{Q},k}} \tag{6.49}$$

for the uniform quantizer shown in Subsection 6.2.2. Here, the diagonal matrix $\boldsymbol{\Phi}_k^2$ contains the d dominant eigenvalues of $\boldsymbol{W}_k$ in Eqs. (6.32) and (6.39), for the cases $L > 1$ and $L = 1$, respectively, on its diagonal.

However, for the random vector quantizer also explained in Subsection 6.2.2, the MSE error matrix $\boldsymbol{C}_{\boldsymbol{\Theta},k}$ is given by

$$\boldsymbol{C}_{\boldsymbol{\Theta},k} = \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k^2\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} + \boldsymbol{C}_{\mathrm{Q},k}. \tag{6.50}$$

6.3 Bayesian Error Modeling based on Joint CSI MSE

In this section, we propose a feedback design for correlated channels that jointly considers the estimation, the rank reduction, and the quantization steps [116, 118]. Fig. 6.3 depicts the feedback model for this approach where the quantizer is explicitly included since codebook entries and partition cells are considered into the joint MSE optimization, as opposed to the optimization of Section 6.2. Therefore, the new formulation is a considerable extension to that of Section 6.2, where we only optimized the estimation and rank reduction. Our goal is the joint optimization of the orthonormal basis $\boldsymbol{V}_k$, the estimator $\boldsymbol{G}_k$, the codebook entries $\boldsymbol{y}_{k,i}$, and the partition cells $\mathcal{R}_{k,i}$, $i = 1, \ldots, M$, by

minimizing the MSE, i.e.

$$\left\{\boldsymbol{V}_k, \boldsymbol{G}_k, \{\boldsymbol{y}_{k,i}\}_{i=1}^{M}, \{\mathcal{R}_{k,i}\}_{i=1}^{M}\right\}_{\text{opt}} = \underset{\{\boldsymbol{V}_k, \boldsymbol{G}_k, \{\boldsymbol{y}_{k,i}\}_{i=1}^{M}, \{\mathcal{R}_{k,i}\}_{i=1}^{M}\}}{\operatorname{argmin}} \text{MSE}_k \quad \text{s.t.:} \quad \boldsymbol{V}_k^{\text{H}}\boldsymbol{V}_k = \mathbf{I}_d \tag{6.51}$$

with MSE_k given by

$$\text{MSE}_k\left(\boldsymbol{V}_k, \boldsymbol{G}_k, \{\boldsymbol{y}_{k,i}\}_{i=1}^{M}, \{\mathcal{R}_{k,i}\}_{i=1}^{M}\right) = \text{E}\left[\left\|\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_k[q]\right\|_2^2\right] \tag{6.52}$$

where $\hat{\boldsymbol{h}}_k[q]$ is the CSI recovered at the transmitter given by

$$\hat{\boldsymbol{h}}_k[q] = \boldsymbol{V}_k \, \text{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right) = \boldsymbol{V}_k \tilde{\boldsymbol{h}}_{\text{Q},k}[q] \in \mathbb{C}^N. \tag{6.53}$$

Remember that the feedback channel introduces a delay given by $D = q - \nu$ time slots considered as an estimation delay instead. $\tilde{\boldsymbol{h}}_k[q]$ in Eq. (6.53) collects the coefficients of the rank reduced representation as follows

$$\tilde{\boldsymbol{h}}_k[q] = \tilde{\boldsymbol{h}}_{\text{T},k}[\nu = q - D] = \boldsymbol{G}_k\left(\boldsymbol{S}\boldsymbol{h}_k[\nu] + \boldsymbol{\eta}_k[\nu]\right) \in \mathbb{C}^d \tag{6.54}$$

with the covariance matrix given by

$$\boldsymbol{C}_{\tilde{\boldsymbol{h}},k} = \text{E}\left[\tilde{\boldsymbol{h}}_k[q]\tilde{\boldsymbol{h}}_k^{\text{H}}[q]\right] = \boldsymbol{G}_k\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)\boldsymbol{G}_k^{\text{H}}. \tag{6.55}$$

Therefore, Eq. (6.53) can be rewritten as

$$\hat{\boldsymbol{h}}_k[q] = \boldsymbol{V}_k\tilde{\boldsymbol{h}}_k[q] + \boldsymbol{V}_k\tilde{\boldsymbol{\eta}}_{\text{Q},k}[q] = \boldsymbol{V}_k \, \text{Q}_k\left(\boldsymbol{G}_k\left(\boldsymbol{S}\boldsymbol{h}_k[\nu] + \boldsymbol{\eta}_k[\nu]\right)\right). \tag{6.56}$$

As we will see below, the main difficulty is the derivation of $\boldsymbol{V}_k$ and $\boldsymbol{G}_k$. The conditions for $\boldsymbol{y}_{k,i}$ and $\mathcal{R}_{k,i}$ are standard. Note that $\boldsymbol{V}_k\boldsymbol{V}_k^{\text{H}} \neq \mathbf{I}_N$ (although $\boldsymbol{V}_k^{\text{H}}\boldsymbol{V}_k = \mathbf{I}_d$), since $d \leq N$.

6.3.1 Codebook Entries

Substituting Eq. (6.53) and the definition of $\text{Q}_k(\bullet) = \sum_{i=1}^{M} \boldsymbol{y}_{k,i} \, \text{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])$ into the MSE of Eq. (6.52) we have

$$\begin{aligned}\text{MSE}_k &= \text{E}[||\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_k[q]||_2^2] = \text{E}[||\boldsymbol{h}_k[q] - \boldsymbol{V}_k \sum_{i=1}^{M} \boldsymbol{y}_{k,i} \, \text{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])||_2^2] \\ &= \text{tr}(\boldsymbol{C}_{\boldsymbol{h},k}) - 2\Re(\sum_{i=1}^{M} \boldsymbol{y}_{k,i}^{\text{H}}\boldsymbol{V}_k^{\text{H}} \, \text{E}[\text{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\boldsymbol{h}_k[q]]) + \sum_{i=1}^{M} \boldsymbol{y}_{k,i}^{\text{H}}\boldsymbol{y}_{k,i} \, \text{E}[\text{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])].\end{aligned} \tag{6.57}$$

By setting the derivative with respect to $\boldsymbol{y}_{k,i}$ to zero we obtain that

$$-\boldsymbol{V}_k^{\mathrm{H}}\,\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\boldsymbol{h}_k[q]\right]+\boldsymbol{y}_{k,i}\,\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\right]=\boldsymbol{0}$$

which leads to

$$\boldsymbol{y}_{k,i}=\left(\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\right]\right)^{-1}\boldsymbol{V}_k^{\mathrm{H}}\,\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\boldsymbol{h}_k[q]\right] \tag{6.58}$$

which is the well known centroid condition [108]. The rank-reduced estimate $\tilde{\boldsymbol{h}}_k[q]$ was previously found and is given in Eq. (6.54). Thus, the MSE expression in Eq. (6.57) can be rewritten as

$$\mathrm{MSE}_k=\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right)-\sum_{i=1}^{M}\frac{\mathrm{E}[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\boldsymbol{h}_k^{\mathrm{H}}[q]]\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\,\mathrm{E}[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\boldsymbol{h}_k[q]]}{\mathrm{E}[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])]}. \tag{6.59}$$

Remember that the channel $\boldsymbol{h}_k[q]$ and the noise $\boldsymbol{\eta}_k[q]$ are Gaussian. Therefore, $\boldsymbol{h}_k[q]$ and $\tilde{\boldsymbol{h}}_k[q]$ are jointly Gaussian because they are related through $\tilde{\boldsymbol{h}}_k[q]=\boldsymbol{G}(\boldsymbol{S}\boldsymbol{h}_k[\nu]+\boldsymbol{\eta}[\nu])$, i.e.

$$\begin{bmatrix}\boldsymbol{h}_k[q]\\ \tilde{\boldsymbol{h}}_k[q]\end{bmatrix}\sim\mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{0},\begin{bmatrix}\boldsymbol{C}_{\boldsymbol{h},k} & r_k\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\\ r_k\boldsymbol{G}_k\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k} & \boldsymbol{G}_k\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}+\boldsymbol{C}_{\boldsymbol{\eta},k}\right)\boldsymbol{G}_k^{\mathrm{H}}\end{bmatrix}\Big) \tag{6.60}$$

where $r_k=\mathrm{J}_0(\alpha_k D)$ [see Eq. (4.24)]. In addition, it is not difficult to find the mean of $\boldsymbol{h}_k[q]$ conditional on $\tilde{\boldsymbol{h}}_k[q]$ using the Theorem 10.2 of [26]. Indeed, given the zero-mean joint Gaussian vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ with covariance matrices $\boldsymbol{C}_{\boldsymbol{x}}$ and $\boldsymbol{C}_{\boldsymbol{y}}$, respectively, and the crosscovariance matrix $\boldsymbol{C}_{\boldsymbol{y}\boldsymbol{x}}=\mathrm{E}[\boldsymbol{y}\boldsymbol{x}^{\mathrm{H}}]$, the mean and the covariance matrix describing $f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{y}|\boldsymbol{x})=f_{\mathrm{G}}(\boldsymbol{y},\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}},\boldsymbol{C}_{\boldsymbol{y}|\boldsymbol{x}})$ are

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}}&=\mathrm{E}[\boldsymbol{y}|\boldsymbol{x}]=\boldsymbol{C}_{\boldsymbol{y}\boldsymbol{x}}{\boldsymbol{C}_{\boldsymbol{x}}}^{-1}\boldsymbol{x}\\ \boldsymbol{C}_{\boldsymbol{y}|\boldsymbol{x}}&=\mathrm{E}[\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}|\boldsymbol{x}]-\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}}^{\mathrm{H}}=\boldsymbol{C}_{\boldsymbol{y}}-\boldsymbol{C}_{\boldsymbol{y}\boldsymbol{x}}{\boldsymbol{C}_{\boldsymbol{x}}}^{-1}\boldsymbol{C}_{\boldsymbol{x}\boldsymbol{y}}\end{aligned} \tag{6.61}$$

respectively. Let $\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}$ be the mean of $\boldsymbol{h}_k[q]$ conditional on $\tilde{\boldsymbol{h}}_k[q]$ and $\boldsymbol{C}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}$ the covariance matrix of $\boldsymbol{h}_k[q]$ conditional on $\tilde{\boldsymbol{h}}_k[q]$. According to Eq. (6.61), we obtain that

$$\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}=\mathrm{E}\left[\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]\right]=r_k\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1}\tilde{\boldsymbol{h}}_k[q] \tag{6.62}$$

$$\begin{aligned}\boldsymbol{C}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}&=\mathrm{E}\left[\left(\boldsymbol{h}_k[q]-\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}\right)\left(\boldsymbol{h}_k[q]-\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}\right)^{\mathrm{H}}|\tilde{\boldsymbol{h}}_k[q]\right]\\ &=\boldsymbol{C}_{\boldsymbol{h},k}-r_k^2\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1}\boldsymbol{G}_k\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\end{aligned} \tag{6.63}$$

where $\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}$ can be found in Eq. (6.55).

Next, let us remember that $\tilde{\boldsymbol{h}}_k[q]$ can be modeled as $\tilde{\boldsymbol{h}}_k[q] = \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{w}$ with $\boldsymbol{w} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$ [similar to Eq. (2.23)]. Moreover, bearing in mind the equality $\mathrm{E}[\boldsymbol{h}_k[q]] = \mathrm{E}[\mathrm{E}[\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]]]$ provided by the dependence between $\boldsymbol{h}_k[q]$ and $\tilde{\boldsymbol{h}}_k[q]$ that gives us $\mathrm{E}[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\boldsymbol{h}_k[q]] = \mathrm{E}[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\,\mathrm{E}[\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]]]$, we obtain with Eq. (6.62) that

$$\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q])\boldsymbol{h}_k[q]\right] = r_k \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2}\,\mathrm{E}\left[\mathrm{S}_{k,i}\left(\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{w}\right)\boldsymbol{w}\right].$$

Substituting this result into Eq. (6.59) we obtain for the MSE

$$\mathrm{MSE}_k = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) - r_k^2\,\mathrm{tr}\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2}\boldsymbol{C}_{\mathrm{Q},k}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2}\boldsymbol{G}_k\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{V}_k\right) \tag{6.64}$$

with

$$\boldsymbol{C}_{\mathrm{Q},k} = \sum_{i=1}^{M} \frac{\mathrm{E}[\mathrm{S}_{k,i}(\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{w})\boldsymbol{w}]\,\mathrm{E}[\mathrm{S}_{k,i}(\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{w})\boldsymbol{w}^{\mathrm{H}}]}{\mathrm{E}[\mathrm{S}_{k,i}(\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{w})]}. \tag{6.65}$$

6.3.2 Estimator and Rank Reduction Designs

In this subsection, we derive the expression for the estimator $\boldsymbol{G}_k \in \mathbb{C}^{d\times N_{\mathrm{tr}}}$ that performs estimation and rank reduction at the same time. Given that the covariance matrix $\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}$ can be expressed as $\boldsymbol{G}_k(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})\boldsymbol{G}_k^{\mathrm{H}}$ [see Eq. (6.55)] and the unknown matrix $\boldsymbol{X}_k \in \mathbb{C}^{N_{\mathrm{tr}}\times d}$, which has orthonormal columns, is introduced to simplify the notation in the following derivation, we get for the estimator

$$\boldsymbol{G}_k = \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{X}_k^{\mathrm{H}}\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1/2} \in \mathbb{C}^{d\times N_{\mathrm{tr}}} \tag{6.66}$$

where it is easy to show that Eq. (6.55) is fulfilled, since $\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{X}_k = \mathbf{I}_d$.

Let us define $\boldsymbol{A}_k = r_k\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1/2} \in \mathbb{C}^{N\times N_{\mathrm{tr}}}$. We must solve

$$\{\boldsymbol{V}_{\mathrm{opt},k}, \boldsymbol{X}_{\mathrm{opt},k}\} = \underset{\{\boldsymbol{V}_k,\boldsymbol{X}_k\}}{\mathrm{argmax}}\,\mathrm{tr}\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{C}_{\mathrm{Q},k}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\right) \tag{6.67}$$

subject to $\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{V}_k = \mathbf{I}_d$ and $\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{X}_k = \mathbf{I}_d$ in order to minimize the MSE in Eq. (6.64). We construct the Lagrangian function as follows

$$\begin{aligned} L\left(\boldsymbol{V}_k, \boldsymbol{X}_k, \boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2\right) = {} & \mathrm{tr}\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{C}_{\mathrm{Q},k}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\right) \\ & - \mathrm{tr}\left(\boldsymbol{\Lambda}_1\left(\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{V}_k - \mathbf{I}_d\right)\right) - \mathrm{tr}\left(\boldsymbol{\Lambda}_2\left(\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{X}_k - \mathbf{I}_d\right)\right) \end{aligned}$$

with $\boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_1^{\mathrm{H}} \in \mathbb{C}^{d\times d}$ and $\boldsymbol{\Lambda}_2 = \boldsymbol{\Lambda}_2^{\mathrm{H}} \in \mathbb{C}^{d\times d}$. The derivative of the Lagrangian function with respect to $\boldsymbol{X}_k^*$ is

$$\frac{\partial L(\bullet)}{\partial \boldsymbol{X}_k^*} = \boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{C}_{\mathrm{Q},k} - \boldsymbol{X}_k\boldsymbol{\Lambda}_2^{\mathrm{H}} = \mathbf{0}$$

and multiplying both terms from the left by $\boldsymbol{X}_k^{\mathrm{H}}$, we have

$$\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{C}_{\mathrm{Q},k} = \boldsymbol{\Lambda}_2^{\mathrm{H}}.$$

From this result and taking into account that $\boldsymbol{\Lambda}_2 = \boldsymbol{\Lambda}_2^{\mathrm{H}}$, we conclude that

$$\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{C}_{\mathrm{Q},k} = \boldsymbol{C}_{\mathrm{Q},k}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k.$$

With the EVD $\boldsymbol{C}_{\mathrm{Q},k} = \boldsymbol{U}_k\boldsymbol{\Xi}_k\boldsymbol{U}_k^{\mathrm{H}}$, we can rewrite the above expression as follows

$$\boldsymbol{U}_k^{\mathrm{H}}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{U}_k\boldsymbol{\Xi}_k = \boldsymbol{\Xi}_k\boldsymbol{U}_k^{\mathrm{H}}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{U}_k.$$

Since $\boldsymbol{\Xi}_k$ is diagonal, $\boldsymbol{U}_k^{\mathrm{H}}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{U}_k$ must be diagonal to fulfill this equation, i.e. $\boldsymbol{U}_k^{\mathrm{H}}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{V}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{U}_k = \boldsymbol{\Phi}_k^2$ where $\boldsymbol{\Phi}_k = \boldsymbol{Q}_k\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{X}_k\boldsymbol{U}_k$ is diagonal with some unitary matrix $\boldsymbol{Q}_k \in \mathbb{C}^{d\times d}$. Therefore, the cost function reduces to

$$\mathrm{MSE}_k = \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h},k}\right) - \mathrm{tr}\left(\boldsymbol{\Xi}_k\boldsymbol{\Phi}_k^2\right)$$

which does not depend on the unitary matrix $\boldsymbol{Q}_k$. Thus, we set $\boldsymbol{Q}_k = \mathbf{I}$. In order to maximize this resulting objective $\mathrm{tr}(\boldsymbol{\Xi}_k\boldsymbol{\Phi}_k^2)$ under the assumption that the diagonal entries of $\boldsymbol{\Xi}_k$ are sorted in non–increasing order, we must choose the i-th column of $\boldsymbol{V}_k$ and $\boldsymbol{W}_k = \boldsymbol{X}_k\boldsymbol{U}_k$ to be the i-th dominant left and right singular vector of $\boldsymbol{A}_k$, respectively. Thus, $\boldsymbol{\Phi}_k$ has the d dominant singular values of $\boldsymbol{A}_k$ on its diagonal. We see that the optimal basis $\boldsymbol{V}_{\mathrm{opt},k}$ contains the d dominant left singular vectors of $\boldsymbol{A}$ or, equivalently, the d dominant eigenvectors of $\boldsymbol{A}_k\boldsymbol{A}_k^{\mathrm{H}}$. Interestingly, we also obtained this result for the case where only $\boldsymbol{V}_k$ and $\boldsymbol{G}_k$ had been optimized as in Section 6.2 [cf. Eq. (6.40)]. Note that $\boldsymbol{V}_{\mathrm{opt},k}$ and $\boldsymbol{W}_{\mathrm{opt},k}$ are fixed for given statistics $\boldsymbol{C}_{\boldsymbol{h},k}$ and $\boldsymbol{C}_{\boldsymbol{\eta},k}$. Therefore, the maximization of Eq. (6.67) is solved by $\boldsymbol{X}_{\mathrm{opt},k} = \boldsymbol{W}_{\mathrm{opt},k}\boldsymbol{U}_k^{\mathrm{H}}$, i.e. the MSE is minimized, where $\boldsymbol{U}_k$ is the modal matrix of $\boldsymbol{C}_{\mathrm{Q},k}$ of Eq. (6.65).

Since $\boldsymbol{W}_{\mathrm{opt},k} = \boldsymbol{X}_{\mathrm{opt},k}\boldsymbol{U}_k$ contains the principal right singular vectors of $\boldsymbol{A}_k$, we have that $\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}}\boldsymbol{A}_k = \boldsymbol{\Phi}_k\boldsymbol{U}_k^{\mathrm{H}}\boldsymbol{X}_{\mathrm{opt},k}^{\mathrm{H}}$. Accordingly, the estimator can be written as [cf. Eq. (6.66)]

$$\boldsymbol{G}_{\mathrm{opt},k} = \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{U}_k\boldsymbol{\Phi}_k^{-1}\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}}\boldsymbol{G}_{\text{MMSE-pred},k} \in \mathbb{C}^{d\times N_{\mathrm{tr}}} \tag{6.68}$$

where $\boldsymbol{G}_{\text{MMSE-pred},k}$ is the conventional linear MMSE predictor given by (cf. Subsection 4.1.2)

$$\boldsymbol{G}_{\text{MMSE-pred},k} = r_k\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1} \in \mathbb{C}^{N\times N_{\mathrm{tr}}}.$$

The estimator is then followed by the rank reduction performed by $\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}}$. $\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}}$ also produces decorrelation since the output of $\boldsymbol{G}_{\mathrm{opt},k}$ has the diagonal covariance matrix $\boldsymbol{\Phi}_k^2$. These two stages constitute the solution for the estimator of Eq. (6.37) for the case that

$L = 1$. However, when the quantizer is included in the optimization as in Eq. (6.51), an additional transformation with $\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{U}_k\boldsymbol{\Phi}_k^{-1}$ appears for the estimator. Note that $\boldsymbol{\Phi}_k^{-1}$ aims at normalizing every entry to unit variance, $\boldsymbol{U}_k$ applies some unitary rotation that does not affect the distribution (see Appendix D.2) and, finally, $\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}$ again introduces coloring to ensure that $\boldsymbol{C}_{\tilde{\boldsymbol{h}},k} = \boldsymbol{G}_{\text{opt},k}(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})\boldsymbol{G}_{\text{opt},k}^{\text{H}}$.

6.3.3 Quantizer Design

Having obtained the preliminary result of Eq. (6.58) for the codebook entries, we will describe in the following subsections how to obtain the quantizer parameters using the Lloyd algorithm, i.e. the codebook entries and the decision boundaries arising from the joint MSE optimization in Eq. (6.51).

Partition Cells

The MSE is the average distortion, i.e.

$$\mathrm{E}\left[\mathrm{d}_k\left(\tilde{\boldsymbol{h}}_k[q], \mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q])\right)\right] = \mathrm{E}\left[\mathrm{E}\left[\left\|\boldsymbol{h}_k[q] - \hat{\boldsymbol{h}}_k[q]\right\|_2^2 \middle| \tilde{\boldsymbol{h}}_k[q]\right]\right].$$

With $\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]} = \mathrm{E}[\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]]$ and $\boldsymbol{C}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]} = \mathrm{E}[(\boldsymbol{h}_k[q] - \boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]})(\boldsymbol{h}_k[q] - \boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]})^{\text{H}}]$, we get

$$\begin{aligned}
&\mathrm{d}_k\left(\tilde{\boldsymbol{h}}_k[q], \mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q])\right) = \mathrm{E}\left[\left\|\boldsymbol{h}_k[q] - \boldsymbol{V}_k\,\mathrm{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right)\right\|_2^2 \middle| \tilde{\boldsymbol{h}}_k[q]\right] = \mathrm{E}\left[\|\boldsymbol{h}_k[q]\|_2^2 \middle| \tilde{\boldsymbol{h}}_k[q]\right] \\
&\quad - 2\Re\left(\mathrm{E}\left[\boldsymbol{h}_k^{\text{H}}[q] \middle| \tilde{\boldsymbol{h}}_k[q]\right]\boldsymbol{V}_k\,\mathrm{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right)\right) + \left\|\boldsymbol{V}_k\,\mathrm{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right)\right\|_2^2 \\
&= \mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}\right) + \boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}^{\text{H}}\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]} - 2\Re\left(\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}^{\text{H}}\boldsymbol{V}_k\,\mathrm{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right)\right) \\
&\quad + \left\|\boldsymbol{V}_k\,\mathrm{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right)\right\|_2^2 \\
&= c_k + \left\|\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]} - \boldsymbol{V}_k\,\mathrm{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right)\right\|_2^2
\end{aligned}$$

where $c_k = \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}}\boldsymbol{G}_k^{\text{H}}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1}\boldsymbol{G}_k\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k})$ is the trace of the covariance matrix of $\boldsymbol{h}_k[q]$ conditional on $\tilde{\boldsymbol{h}}_k[q]$. Note that the term $\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}) + \boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}^{\text{H}}\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_k[q]}$ comes from $\mathrm{E}[\|\boldsymbol{h}_k[q]\|_2^2\,|\tilde{\boldsymbol{h}}_k[q]] = \mathrm{tr}(\mathrm{E}[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\text{H}}[q]|\tilde{\boldsymbol{h}}_k[q]])$. Substituting Eq. (6.62) leads to

$$\mathrm{d}_k\left(\tilde{\boldsymbol{h}}_k[q], \mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q])\right) = c_k + \left\|r_k\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}}\boldsymbol{G}_k^{\text{H}}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1}\tilde{\boldsymbol{h}}_k[q] - \boldsymbol{V}_k\,\mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q])\right\|_2^2.$$

Due to Eq. (6.68), we have that $r_k \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\mathrm{H}} \boldsymbol{G}_{\text{opt},k}^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1} = \boldsymbol{V}_{\text{opt},k} \boldsymbol{\Phi}_k \boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2}$. Hence, under the assumption that $\boldsymbol{G}_{\text{opt},k}$ and $\boldsymbol{V}_{\text{opt},k}$ are used, we get for the distortion

$$\mathrm{d}_k \left(\tilde{\boldsymbol{h}}_k[q], \mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q]) \right) = c_k + \left\| \boldsymbol{\Phi}_k \boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2} \tilde{\boldsymbol{h}}_k[q] - \mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q]) \right\|_2^2. \tag{6.69}$$

Since the MSE $\mathrm{E}[\mathrm{d}_k(\tilde{\boldsymbol{h}}_k[q], \mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q]))]$ is minimized, the partition cells must be chosen to minimize $\mathrm{d}_k(\tilde{\boldsymbol{h}}_k[q], \mathrm{Q}_k(\tilde{\boldsymbol{h}}_k[q]))$ for every $\tilde{\boldsymbol{h}}_k[q]$, i.e. $\mathcal{R}_{k,i} = \{\boldsymbol{x} \in \mathbb{C}^d | \, \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}_{k,i}) \leq \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}_{k,j}), \forall j\}$, which is the nearest neighbor condition [108].

Suggested Codebook Design

The expression in Eq. (6.69) for the distortion can be further simplified. As $\mathrm{Q}_k(\boldsymbol{y}) = \sum_{i=1}^{M} \boldsymbol{y}_{k,i} \, \mathrm{S}_{k,i}(\boldsymbol{y})$, we have to rewrite Eq. (6.58) by incorporating Eqs. (6.62) and (6.68). The i-th codebook entry then reads as

$$\begin{aligned}
\boldsymbol{y}_{k,i} &= \left(\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q]) \right] \right)^{-1} \boldsymbol{V}_k^{\mathrm{H}} \, \mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q]) \right] \mathrm{E}\left[\boldsymbol{h}_k[q] | \tilde{\boldsymbol{h}}_k[q] \right] \\
&= \left(\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q]) \right] \right)^{-1} r_k \boldsymbol{V}_{\text{opt},k}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\mathrm{H}} \boldsymbol{G}_{\text{opt},k}^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1} \, \mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q]) \tilde{\boldsymbol{h}}_k[q] \right] \\
&= \left(\mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q]) \right] \right)^{-1} \boldsymbol{\Phi}_k \boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2} \, \mathrm{E}\left[\mathrm{S}_{k,i}(\tilde{\boldsymbol{h}}_k[q]) \tilde{\boldsymbol{h}}_k[q] \right]
\end{aligned}$$

where

$$\begin{aligned}
r_k \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\mathrm{H}} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1} &= \overbrace{\boldsymbol{V}_k^{\mathrm{H}} \underbrace{r_k \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\mathrm{H}} \left(\boldsymbol{S} \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2}}_{\boldsymbol{A}_k} \boldsymbol{X}_k \boldsymbol{U}_k}^{\boldsymbol{\Phi}_k} \boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1} \\
&= \boldsymbol{\Phi}_k \boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2}.
\end{aligned}$$

It is useful to redefine the quantizer as

$$\mathrm{Q}_k(\boldsymbol{y}) = \boldsymbol{\Phi}_k \, \mathrm{Q}'_k \left(\boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2} \boldsymbol{y} \right) \tag{6.70}$$

to eliminate the uncertainty due to the first two terms of the filter $\boldsymbol{G}_{\text{opt},k}$ of Eq. (6.68). Note that $\tilde{\boldsymbol{h}}_k[q]$ has the same statistical properties as $\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2} \boldsymbol{w} = \boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2} \boldsymbol{U}_k \boldsymbol{w}$, since the unitary rotation with $\boldsymbol{U}_k$ does not change the statistical properties of $\boldsymbol{w}$ (see Appendix D.2) where $\boldsymbol{w} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I})$ as before. The redefined quantizer $\mathrm{Q}'_k(\bullet)$ is given by

$$\mathrm{Q}'_k(\boldsymbol{y}) = \sum_{i=1}^{M} \boldsymbol{y}'_{k,i} \, \mathrm{S}'_{k,i}(\boldsymbol{y}) \tag{6.71}$$

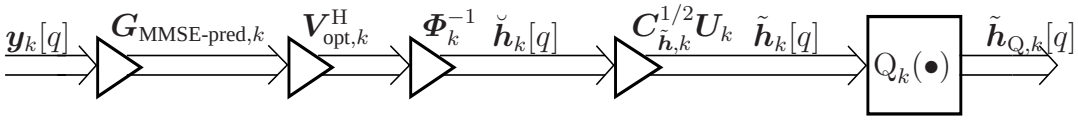


Figure 6.4: Preliminary Quantizer Design for Limited Feedback.

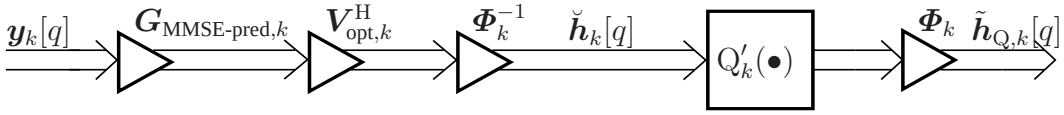


Figure 6.5: Proposed Quantizer Redesign for Limited Feedback.

where $\text{S}'_{k,i}(\boldsymbol{y}) = \text{S}_{k,i}(\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{U}_k\boldsymbol{y})$ due to Eq. (6.70). The new codebook entries are

$$\boldsymbol{y}'_{k,i} = \boldsymbol{\Phi}_k^{-1}\boldsymbol{y}_{k,i} = \left(\text{E}\left[\text{S}'_{k,i}(\boldsymbol{w})\right]\right)^{-1}\text{E}\left[\text{S}'_{k,i}(\boldsymbol{w})\,\boldsymbol{w}\right]. \tag{6.72}$$

Then, we get for the distortion [cf. Eq. (6.69)]

$$\begin{aligned}\text{d}_k\left(\tilde{\boldsymbol{h}}_k[q], \text{Q}_k\left(\tilde{\boldsymbol{h}}_k[q]\right)\right) &= \text{d}_k\left(\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{U}_k\breve{\boldsymbol{h}}_k[q], \text{Q}_k(\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{1/2}\boldsymbol{U}_k\breve{\boldsymbol{h}}_k[q])\right)\\ &= c_k + \left\|\boldsymbol{\Phi}_k\left(\breve{\boldsymbol{h}}_k[q] - \text{Q}'_k(\breve{\boldsymbol{h}}_k[q])\right)\right\|_2^2\end{aligned} \tag{6.73}$$

where $\breve{\boldsymbol{h}}_k[q] = \boldsymbol{U}_k^{\text{H}}\boldsymbol{C}_{\tilde{\boldsymbol{h}},k}^{-1/2}\tilde{\boldsymbol{h}}_k[q] \in \mathbb{C}^d$. Remember that $\boldsymbol{\Phi}_k$ is diagonal. Thus, the distortion to be minimized for the design of $\text{Q}'_k(\bullet)$ has a very simple structure. Additionally, $\breve{\boldsymbol{h}}_k[q] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \mathbf{I})$ which leads to the simple centroid condition in Eq. (6.72).

Note that we can concentrate on the design of $\text{Q}'_k(\bullet)$, because $\text{Q}_k(\tilde{\boldsymbol{h}}_k[q]) = \boldsymbol{\Phi}_k\,\text{Q}'_k(\breve{\boldsymbol{h}}_k[q])$ with the output $\breve{\boldsymbol{h}}_k[q]$ of

$$\boldsymbol{G}'_{\text{opt},k} = \boldsymbol{\Phi}_k^{-1}\boldsymbol{V}_{\text{opt},k}^{\text{H}}\boldsymbol{G}_{\text{MMSE-pred},k}. \tag{6.74}$$

Figs. 6.4 and 6.5 depict the quantizer design previously proposed in comparison with the quantizer redefinition shown in this subsection, in which the inputs to the quantizer $\text{Q}'_k(\bullet)$ are white vectors of Gaussian random variables.

Then, the resulting CSI of the transmitter is $\boldsymbol{V}_{\text{opt},k}\boldsymbol{\Phi}_k\,\text{Q}'_k(\boldsymbol{G}'_{\text{opt},k}\boldsymbol{y}_k[q])$. Also note that the estimator $\boldsymbol{G}'_{\text{opt},k}$ and the basis $\boldsymbol{V}_{\text{opt},k}$ only depend on the channel statistics. Hence, they can be computed independently of the choice for $\text{Q}'_k(\bullet)$.

To summarize, we obtained from the joint optimization in Eq. (6.51) that the received training symbols $\boldsymbol{y}_k[q]$ are passed through the ordinary MMSE predictor $\boldsymbol{G}_{\text{MMSE-pred},k}$, rank reduced with $\boldsymbol{V}_{\text{opt},k}^{\text{H}}$, and weighted with $\boldsymbol{\Phi}_k^{-1}$ to obtain uncorrelated unit-variance entries. Then, the index ℓ_k found by the quantizer $\text{Q}'_k(\bullet)$ is fed back and the CSI at the transmitter is $\boldsymbol{V}_{\text{opt},k}\boldsymbol{\Phi}_k\boldsymbol{y}'_{k,\ell}$. Note that $\boldsymbol{A}_k = r_k\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}}(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\text{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1/2}$ depends

only on the channel statistics $\boldsymbol{C}_{\boldsymbol{h},k}$ and $\boldsymbol{C}_{\boldsymbol{\eta},k}$, which change very slowly. Therefore, $\boldsymbol{V}_{\text{opt},k}$ can be communicated to the transmitter with negligible overhead and we assume a perfect knowledge of $\boldsymbol{V}_{\text{opt},k}$ and $\boldsymbol{\Phi}_k$ at the transmitter. Despite the simplicity of Eq. (6.73), we

1. Set $m = 1$
2. Initial codebook $\mathcal{C}_1 = \{y_i\}$ and partition cells $\mathcal{R}_i$ with the uniform scalar quantizer of Subsection 6.2.2
3. Set the threshold to stop the iterations ϵ_{min} and set $\epsilon = \infty$

while $\epsilon > \epsilon_{\text{min}}$ **do**

4. Given the codebook $\mathcal{C}_m$, the Lloyd algorithm gives the improved codebook $\mathcal{C}_{m+1}$

Nearest Neighbor Condition: The distortion of the scalar quantizer is given by

$$\mathrm{d}\left(x, \mathrm{Q}\left(x\right)\right) = \left(x - \mathrm{Q}\left(x\right)\right)^2 = \sum_{i=1}^{M} \left(x - y_i\right) \mathrm{S}_i\left(x\right)$$

The new partition cell is easily obtained by applying this distortion measure as

$$\mathcal{R}_i = \{\alpha_i \leq x < \beta_i\} \quad \text{with } \alpha_i = \tfrac{y_{i-1}+y_i}{2} \text{ and } \beta_i = \tfrac{y_i+y_{i+1}}{2}$$

5. *Centroid Condition*: Applying the centroid condition, the new codebook is

$$y_i = \left(\mathrm{E}\left[\mathrm{S}_i\left(x\right)\right]\right)^{-1} \mathrm{E}\left[\mathrm{S}_i\left(x\right) x\right]$$

with $x \sim \mathcal{N}_{\mathbb{R}}\left(0, \frac{1}{2}\right)$. The above expectations are obtained as

$$\begin{aligned}\mathrm{E}\left[\mathrm{S}_i\left(x\right)\right] &= \int_{\alpha_i}^{\beta_i} f_{\mathrm{G}}\left(x, 0, 1/2\right) \mathrm{d}x = \tfrac{1}{2}\left(\mathrm{erfc}\left(\alpha_i\right) - \mathrm{erfc}\left(\beta_i\right)\right) \\ &= \Phi(\sqrt{2}\alpha_i) - \Phi(\sqrt{2}\beta_i)\end{aligned}$$

and

$$\mathrm{E}\left[\mathrm{S}_i\left(x\right) x\right] = \int_{\alpha_i}^{\beta_i} x f_{\mathrm{G}}\left(x, 0, 1/2\right) \mathrm{d}x = \tfrac{1}{2\sqrt{\pi}}\left(\exp\left(-\alpha_i^2\right) - \exp\left(-\beta_i^2\right)\right)$$

where $\mathrm{erfc}\left(\bullet\right)$ is the complementary error function defined as $\mathrm{erfc}\left(x\right) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \mathrm{e}^{-t^2} \mathrm{d}t$ and $\Phi\left(x\right) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \mathrm{e}^{-t^2/2} \mathrm{d}t$. Thus, $\mathrm{erfc}\left(x\right) = 2\,\Phi\left(\sqrt{2}x\right)$

The codebook entries of $\mathcal{C}_{m+1}$ are given by

$$y_i = \tfrac{1}{\sqrt{\pi}} \tfrac{\exp\left(-\alpha_i^2\right) - \exp\left(-\beta_i^2\right)}{\mathrm{erfc}(\alpha_i) - \mathrm{erfc}(\beta_i)} = \tfrac{1}{2\sqrt{\pi}} \tfrac{\exp\left(-\alpha_i^2\right) - \exp\left(-\beta_i^2\right)}{\Phi\left(\sqrt{2}\alpha_i\right) - \Phi\left(\sqrt{2}\beta_i\right)}, \quad i = 1, \ldots, M$$

6. We compute the average distortion (MSE) for $\mathcal{C}_{m+1}$ as follows

$$\begin{aligned}\epsilon &= \mathrm{E}\left[\mathrm{d}\left(x, \mathrm{Q}\left(x\right)\right)\right] = \sum_{i=1}^{M} \mathrm{E}\left[\left(x - y_i\right)^2 \mid x \in \mathcal{R}_i\right] \mathrm{p}(x \in \mathcal{R}_i) \\ &= \tfrac{1}{2} - \sum_{i=1}^{M} \tfrac{1}{2\pi} \tfrac{\left(\exp\left(-\alpha_i^2\right) - \exp\left(-\beta_i^2\right)\right)^2}{\mathrm{erfc}(\alpha_i) - \mathrm{erfc}(\beta_i)}\end{aligned}$$

where M is the number of codebook entries of the scalar quantizer

7. $m \leftarrow m + 1$

end while

Table 6.1: Codebook Optimization of a Scalar Quantizer for a Real–Valued Gaussian Input with Variance 0.5.

suggest separating the scalar quantization for every entry (real and imaginary parts are also split), i.e. the partition cells $\mathcal{R}'_{k,i}$ are hyperrectangles (transform coding, [108]). With this restriction, the design of $\mathrm{Q}'_k(\bullet)$ is independent of $\boldsymbol{\Phi}_k$ or any other quantity related to our system. The scalar quantizer for any of the $2d$ real-valued scalars is the

MMSE optimal scalar quantizer for a real-valued Gaussian random variable with variance 0.5. Due to this property, there is no need to compute the parameters for $\mathrm{Q}'_k(\bullet)$ in real time. Instead, they can be computed in advance (with the Lloyd algorithm [108]) and stored. Moreover, the restriction of separating scalar quantization enables closed-form expressions for the conditional moments to design the precoder, as can be seen in Subsection 6.3.5.

Initial Codebook

Although the estimators and the quantizers are jointly optimized by minimizing the CSI–MSE [e.g. Eq. (6.35)], we can compute the codebook parameters in advance before the data transmission since the estimators are independent of the choice of the codebook. Therefore, we construct the initial codebook off-line to be stored at both the transmitter and receiver side with no need to recompute its parameters for varying channel statistics since it is based on a real–valued standard Gaussian distribution with variance 0.5. This codebook is much more appropriate than the codebooks of the previous subsections, since we do not have uniform inputs, but rather unit variance complex Gaussian inputs. As a consequence, this initial codebook is easily obtained by means of the Lloyd algorithm [108, 119] as shown in Table 6.1. Since its calculation is not trivial, note that the average distortion shown in step 6 of this table is obtained as follows

$$\begin{aligned}
\epsilon\left(\log_2\left(M\right)\right) &= \sum_{i=1}^{M} \mathrm{E}\left[\left(x-y_i\right)^2 \mid x \in \mathcal{R}_i\right] \mathrm{p}(x \in \mathcal{R}_i) \\
&= \sum_{i=1}^{M} \mathrm{E}\left[x^2 \mid x \in \mathcal{R}_i\right] \mathrm{p}(x \in \mathcal{R}_i) + \sum_{i=1}^{M} \left(\frac{\int_{\alpha_i}^{\beta_i} x' f_x(x') \mathrm{d}x'}{\int_{\alpha_i}^{\beta_i} f_x(x') \mathrm{d}x'} \right)^2 \int_{\alpha_i}^{\beta_i} f_x(x) \mathrm{d}x \\
&\quad - 2 \sum_{i=1}^{M} \mathrm{E}\left[x | x \in \mathcal{R}_i\right] \frac{\int_{\alpha_i}^{\beta_i} x' f_x(x') \mathrm{d}x'}{\int_{\alpha_i}^{\beta_i} f_x(x') \mathrm{d}x'} \int_{\alpha_i}^{\beta_i} f_x(x) \mathrm{d}x \\
&= \frac{1}{2} - \sum_{i=1}^{M} \left(\frac{\int_{\alpha_i}^{\beta_i} x' f_x(x') \mathrm{d}x'}{\int_{\alpha_i}^{\beta_i} f_x(x') \mathrm{d}x'} \right)^2 \int_{\alpha_i}^{\beta_i} f_x(x) \mathrm{d}x \\
&= \frac{1}{2} - \sum_{i=1}^{M} \frac{1}{2\pi} \frac{\left(\exp\left(-\alpha_i^2\right) - \exp\left(-\beta_i^2\right)\right)^2}{\mathrm{erfc}\left(\alpha_i\right) - \mathrm{erfc}\left(\beta_i\right)}
\end{aligned} \tag{6.75}$$

where $f_x(x) = \frac{1}{\sqrt{\pi}} \exp(-x^2)$, $\sum_{i=1}^{M} \mathrm{E}[x^2 \mid x \in \mathcal{R}_i] \mathrm{p}(x \in \mathcal{R}_i) = \mathrm{E}[x^2] = \frac{1}{2}$, $y_i = \frac{\mathrm{E}[x' \, \mathrm{S}_{k,i}(x')]}{\mathrm{E}[\mathrm{S}_{k,i}(x')]} = \frac{\int_{\alpha_i}^{\beta_i} x' f_x(x') \mathrm{d}x'}{\int_{\alpha_i}^{\beta_i} f_x(x') \mathrm{d}x'}$, $\mathrm{E}[x | x \in \mathcal{R}_i] = y_i$, and $\mathrm{p}(x \in \mathcal{R}_i) = \int_{\alpha_i}^{\beta_i} f_x(x) \mathrm{d}x$. This average distortion as a function of the number of iterations is plotted in Fig. 6.6.

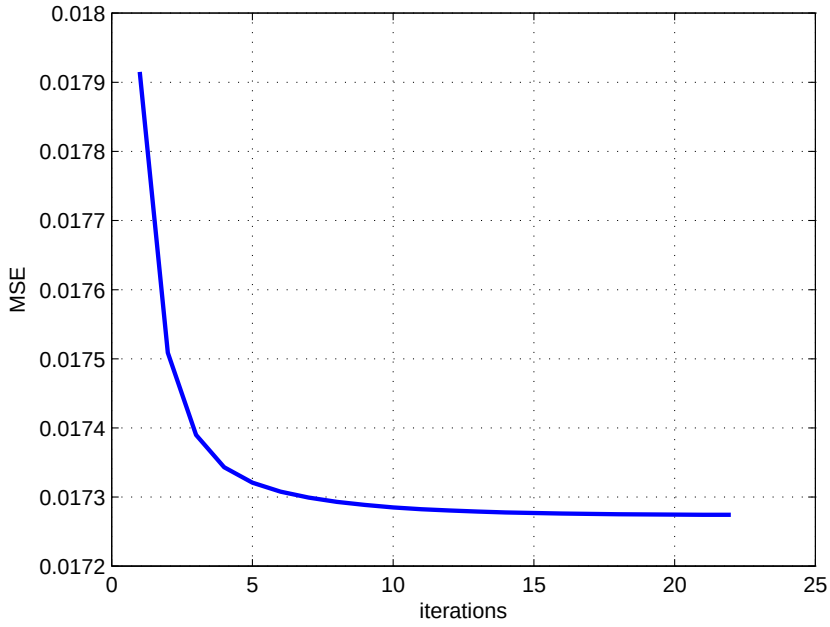


Figure 6.6: Average Distortion vs. Number of Iterations.

6.3.4 Bit Allocation

When using scalar quantization (transform coding, [108]) instead of vector quantization, the available bits have to be allocated to the different scalar coefficients. Since in real systems the bandwidth of feedback channels is very limited, the total number of bits N_{bit} should be very small and, therefore, strategies such as optimum bit allocation can greatly improve the performance with a negligible increase in computational complexity.

The average distortion or MSE is given by Eq. (6.73). Let $\breve{h}_{k,i}^{\text{Re}}[q]$ and $\breve{h}_{k,i}^{\text{Im}}[q]$ be the real or imaginary part of the i–th element of $\breve{\boldsymbol{h}}_k[q]$. With $\boldsymbol{\Phi}_k = \text{diag}(\varphi_{k,1}, \ldots, \varphi_{k,d})$ and using scalar quantizers, Eq. (6.73) can be expressed as

$$\begin{aligned}\varepsilon_k &= c_k + \text{E}\left[\sum_{i=1}^{d} \varphi_{k,i}^2 \left(\left(\breve{h}_{k,i}^{\text{Re}}[q] - \text{Q}_{k,i}^{'}\left(\breve{h}_{k,i}^{\text{Re}}[q]\right)\right)^2 + \left(\breve{h}_{k,i}^{\text{Im}}[q] - \text{Q}_{k,i}^{'}\left(\breve{h}_{k,i}^{\text{Im}}[q]\right)\right)^2\right)\right] \\ &= c_k + \sum_{i=1}^{d} \varphi_{k,i}^2 \varepsilon_{\breve{h}_{k,i}} \end{aligned} \tag{6.76}$$

where $\varepsilon_{\breve{h}_{k,i}} = \text{E}[|\breve{h}_{k,i}[q] - \text{Q}_{k,i}^{'}(\breve{h}_{k,i}^{\text{Re}}[q]) - \text{j}\,\text{Q}_{k,i}^{'}(\breve{h}_{k,i}^{\text{Im}}[q])|^2]$ is the MSE between $\breve{h}_{k,i}[q]$ and its quantized version. Remember that d is the number of coefficients resulting from the rank reduction process to be sent from each user to the transmitter. Note that $\varepsilon_{\breve{h}_{k,i}}$ is fixed

for a given number of bits $b_{k,i}$ since $\breve{\boldsymbol{h}}_k[q] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$. Therefore, for a given number of bits, we can calculate $\varepsilon_{\breve{h}_{k,i}}$ off-line to be stored at the users prior to transmission. Let $\varepsilon_{\breve{h}_{k,i}} = 2\text{MSE}(b_{k,i})$ be the function that determines the mean squared error in terms of the number of bits used to quantize the real or imaginary part of each channel coefficient [see Eq. (6.75)]. Then, the bit allocation problem can be solved by means of the optimization problem

$$\{b_{\text{opt},k,1}, \ldots, b_{\text{opt},k,d}\} = \underset{\{b_{k,1},\ldots,b_{k,d}\}}{\operatorname{argmin}} \sum_{i=1}^{d} \varphi_{k,i}^2 2\text{MSE}\left(b_{k,i}\right) \quad \text{s.t.:} \sum_{i=1}^{d} 2b_{k,i} = N_{\text{bit}} \tag{6.77}$$

where N_{bit} is the number of bits per user sent through the feedback channel. It should be mentioned that we provide the same number of bits to quantize both the real and the imaginary part of each coefficient and, therefore, it is obvious that each quantized coefficient uses an even number of bits.

In principle, we would have to test all the possible bit allocations whose total number of bits N_{bit} is fixed, which can make the search difficult when the number of bits to be allocated is high. However, the MSE of each quantizer decreases with a higher number of bits and, due to $\varphi_{k,1} \geq \ldots \geq \varphi_{k,d}$, the total MSE is always smaller when more bits are allocated to quantize the coefficients with lower indices. Thus, we only have to test bit allocations whose number of bits decreases or stays constant with the coefficient index (see Table 6.2). In the sequel, we refer to this bit allocation algorithm as *optimum bit allocation*.

6.3.5 MSE Error Matrix for Robust Multi-User Precoder Design

For the robust precoder design, we must find the conditional moments $\text{E}[\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\text{Q},k}[q]]$ and $\text{E}[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\text{H}}[q]|\tilde{\boldsymbol{h}}_{\text{Q},k}[q]]$ of the probability density function $f_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\text{Q},k}[q]}(\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\text{Q},k}[q])$, since the transmitter only knows $\tilde{\boldsymbol{h}}_{\text{Q},k}[q]$, but the cost function depends on $\boldsymbol{h}_k[q]$ [3]. The closed-form expressions will be obtained for the special case that $\text{Q}'_k(\bullet)$ performs separate scalar quantization as assumed in the previous two subsections. Remember that the transmission over the feedback channel introduces a delay of $D = q - \nu$ slots, i.e. the precoder is designed during the time slot q and the channel estimate is obtained during the time slot $\nu = q - D$. Remember also that $\breve{\boldsymbol{h}}_k[q] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_d)$ is the input vector to the quantizer given by $\breve{\boldsymbol{h}}_k[q] = \boldsymbol{G}'_{\text{opt},k}\boldsymbol{y}_k[q]$, where $\boldsymbol{G}'_{\text{opt},k}$ is the estimator that results from the quantizer redefinition and $\boldsymbol{y}_k[q]$ is the received pilot signal (see Subsection 6.3.3).

[3] For example, the precoder in Eq. (5.17) depends on $\hat{\boldsymbol{H}}$ and $\boldsymbol{T}$. The row of $\hat{\boldsymbol{H}}$ corresponding to user k is $\text{E}[\boldsymbol{h}_k^{\text{T}}[q]|\tilde{\boldsymbol{h}}_{\text{Q},k}[q]]$ and $\boldsymbol{C}_{\boldsymbol{\Theta}}$ in $\boldsymbol{T}$ [see Eq. (5.14)] contains $\text{E}[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\text{H}}[q]|\tilde{\boldsymbol{h}}_{\text{Q},k}[q]]$ in the Bayesian framework employed in this section.

Bits per user	$d = 2$	$d = 3$	$d = N = 4$
$N_{\text{bit}} = 8$ 4 for real part 4 for imaginary part	$[4,0]^{\mathrm{T}}, [3,1]^{\mathrm{T}}$ $[2,2]^{\mathrm{T}}$	$[4,0,0]^{\mathrm{T}}, [3,1,0]^{\mathrm{T}}$ $[2,2,0]^{\mathrm{T}}, [2,1,1]^{\mathrm{T}}$	$[4,0,0,0]^{\mathrm{T}}, [3,1,0,0]^{\mathrm{T}}$ $[2,2,0,0]^{\mathrm{T}}, [2,1,1,0]^{\mathrm{T}}$ $[1,1,1,1]^{\mathrm{T}}$
$N_{\text{bit}} = 12$ 6 for real part 6 for imaginary part	$[6,0]^{\mathrm{T}}, [5,1]^{\mathrm{T}}$ $[4,2]^{\mathrm{T}}, [3,3]^{\mathrm{T}}$	$[6,0,0]^{\mathrm{T}}, [5,1,0]^{\mathrm{T}}$ $[4,2,0]^{\mathrm{T}}, [4,1,1]^{\mathrm{T}}$ $[3,3,0]^{\mathrm{T}}, [3,2,1]^{\mathrm{T}}$ $[2,2,2]^{\mathrm{T}}$	$[6,0,0,0]^{\mathrm{T}}, [5,1,0,0]^{\mathrm{T}}$ $[4,2,0,0]^{\mathrm{T}}, [4,1,1,0]^{\mathrm{T}}$ $[3,3,0,0]^{\mathrm{T}}, [3,2,1,0]^{\mathrm{T}}$ $[3,1,1,1]^{\mathrm{T}}, [2,2,2,0]^{\mathrm{T}}$ $[2,2,1,1]^{\mathrm{T}}$
$N_{\text{bit}} = 16$ 8 for real part 8 for imaginary part	$[8,0]^{\mathrm{T}}, [7,1]^{\mathrm{T}}$ $[6,2]^{\mathrm{T}}, [5,3]^{\mathrm{T}}$ $[4,4]^{\mathrm{T}}$	$[8,0,0]^{\mathrm{T}}, [7,1,0]^{\mathrm{T}}$ $[6,2,0]^{\mathrm{T}}, [6,1,1]^{\mathrm{T}}$ $[5,3,0]^{\mathrm{T}}, [5,2,1]^{\mathrm{T}}$ $[4,4,0]^{\mathrm{T}}, [4,3,1]^{\mathrm{T}}$ $[4,2,2]^{\mathrm{T}}, [3,3,2]^{\mathrm{T}}$	$[8,0,0,0]^{\mathrm{T}}, [7,1,0,0]^{\mathrm{T}}$ $[6,2,0,0]^{\mathrm{T}}, [6,1,1,0]^{\mathrm{T}}$ $[5,3,0,0]^{\mathrm{T}}, [5,2,1,0]^{\mathrm{T}}$ $[5,1,1,1]^{\mathrm{T}}, [4,4,0,0]^{\mathrm{T}}$ $[4,3,1,0]^{\mathrm{T}}, [4,2,2,0]^{\mathrm{T}}$ $[4,2,1,1]^{\mathrm{T}}, [3,3,2,0]^{\mathrm{T}}$ $[3,3,1,1]^{\mathrm{T}}, [3,2,2,1]^{\mathrm{T}}$ $[2,2,2,2]^{\mathrm{T}}$

Table 6.2: Number of Bits Assigned per User Coefficient for CSI–MSE Metric.

Taking into account that the conditional moments needed for the robust design $\mathrm{E}[\boldsymbol{h}[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]$ and $\mathrm{E}[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]]$ can be further obtained as

$$\begin{aligned}\mathrm{E}\left[\boldsymbol{h}[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]\right] &= \mathrm{E}\left[\mathrm{E}\left[\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]\right]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]\right]\\ \mathrm{E}\left[\boldsymbol{h}[q]\boldsymbol{h}_k^{\mathrm{H}}[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]\right] &= \mathrm{E}\left[\mathrm{E}\left[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q]|\breve{\boldsymbol{h}}_k[q]\right]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]\right]\end{aligned} \tag{6.78}$$

and since $\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q] = \mathrm{Q}_k'(\breve{\boldsymbol{h}}_k[q])$, we start by deriving the expressions for the moments $\mathrm{E}[\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]]$ and $\mathrm{E}[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q]|\breve{\boldsymbol{h}}_k[q]]$.

Since $\boldsymbol{h}_k[q]$ and $\breve{\boldsymbol{h}}_k[q]$ are jointly Gaussian, we have with Eqs. (6.54) and (6.74) that

$$\begin{aligned}\begin{bmatrix}\boldsymbol{h}_k[q]\\ \breve{\boldsymbol{h}}_k[q]\end{bmatrix} &\sim \mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}, \begin{bmatrix}\boldsymbol{C}_{\boldsymbol{h},k} & r_k\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{G}_{\mathrm{opt},k}'^{\mathrm{H}}\\ r_k\boldsymbol{G}_{\mathrm{opt},k}'\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k} & \boldsymbol{G}_{\mathrm{opt},k}'\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}+\boldsymbol{C}_{\boldsymbol{\eta},k}\right)\boldsymbol{G}_{\mathrm{opt},k}'^{\mathrm{H}}\end{bmatrix}\Big)\\ &\sim \mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}, \begin{bmatrix}\boldsymbol{C}_{\boldsymbol{h},k} & \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k\\ \boldsymbol{\Phi}_k\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} & \mathbf{I}\end{bmatrix}\Big)\end{aligned} \tag{6.79}$$

where $r_k = \mathrm{J}_0(\alpha_k D)$ [see Eq. (4.24)]. Hence, applying Eq. (6.61) yields for the conditional mean

$$\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]} = \mathrm{E}\left[\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]\right] = \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k\breve{\boldsymbol{h}}_k[q] \tag{6.80}$$

and for the conditional covariance matrix

$$\boldsymbol{C}_{\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]} = \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k^2\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}}. \tag{6.81}$$

Therefore, the conditional moment $\mathrm{E}[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q]|\breve{\boldsymbol{h}}_k[q]]$ is given by

$$\begin{aligned}\mathrm{E}\left[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[q]|\breve{\boldsymbol{h}}_k[q]\right] &= \boldsymbol{C}_{\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]} + \boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]}\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\breve{\boldsymbol{h}}_k[q]}^{\mathrm{H}}\\ &= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k^2\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} + \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k\breve{\boldsymbol{h}}_k[q]\breve{\boldsymbol{h}}_k^{\mathrm{H}}[q]\boldsymbol{\Phi}_k\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}}.\end{aligned} \tag{6.82}$$

Thus, both the conditional mean and the conditional correlation matrix in Eq. (6.78), henceforth denoted respectively by $\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]}$ and $\boldsymbol{R}_{\boldsymbol{h}_k|[q]\tilde{\boldsymbol{h}}_k[q]}$, can be written as

$$\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]} = \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k\,\mathrm{E}\left[\breve{\boldsymbol{h}}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]\right] = \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k\boldsymbol{m}_k \tag{6.83}$$

and

$$\boldsymbol{R}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]} = \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k^2\boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} + \boldsymbol{V}_{\mathrm{opt},k}\boldsymbol{\Phi}_k\boldsymbol{M}_k\boldsymbol{\Phi}_k\boldsymbol{V}_{\mathrm{opy},k}^{\mathrm{H}} \tag{6.84}$$

respectively. Here, we introduced

$$\begin{aligned}\boldsymbol{m}_k &= \frac{1}{\kappa_k}\int_{\mathbb{S}_{\ell_k}} \boldsymbol{w} f_{\boldsymbol{w}}(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w}\\ \boldsymbol{M}_k &= \frac{1}{\kappa_k}\int_{\mathbb{S}_{\ell_k}} \boldsymbol{w}\boldsymbol{w}^{\mathrm{H}} f_{\boldsymbol{w}}(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w}\end{aligned}$$

with $\kappa_k = \int_{\mathbb{S}_{\ell_k}} f_{\boldsymbol{w}}(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w}$ (see Appendix F.1). Note that $\boldsymbol{m}_k = \boldsymbol{y}_{k,\ell_k}$ [cf. Eq. (6.72)] and $\boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]} = \hat{\boldsymbol{h}}_{\mathrm{Q},k}[q]$. The region $\mathbb{S}_{\ell_k}$ is the hyperrectangle given by

$$\mathbb{S}_{\ell_k} = \left\{\boldsymbol{x} \in \mathbb{C}^d, \forall i : \alpha_{k,i}^{\mathrm{Re}}\left(\ell_k\right) \leq \Re(x_i) \leq \beta_{k,i}^{\mathrm{Re}}\left(\ell_k\right), \alpha_{k,i}^{\mathrm{Im}}\left(\ell_k\right) \leq \Im(x_i) \leq \beta_{k,i}^{\mathrm{Im}}\left(\ell_k\right)\right\} \tag{6.85}$$

where $\alpha_{k,i}^{\mathrm{Re}}(\ell_k)$ and $\beta_{k,i}^{\mathrm{Re}}(\ell_k)$ are the lower and upper limits, respectively, of the partition cells of the quantizer $\mathrm{Q}'_{k,i}(\bullet)$ applied to the real part of $\tilde{h}_{k,i}[q]$ corresponding to the fed–back index ℓ_k of user k. Similarly, $\alpha_{k,i}^{\mathrm{Im}}(\ell_k)$ and $\beta_{k,i}^{\mathrm{Im}}(\ell_k)$ are the lower and upper limits, respectively, of the partition cells of the quantizer $\mathrm{Q}'_{k,i}(\bullet)$ applied to the imaginary part of $\tilde{h}_{k,i}[q]$ when the index ℓ_k is fed back to the transmitter by the user k. Note that $\boldsymbol{w} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$ is used instead of $\breve{\boldsymbol{h}}_k[q]$ for brevity.

Taking into account that $\boldsymbol{C}_{\breve{\boldsymbol{h}},k} = \mathbf{I}_d$, the above expressions can be written as

$$\boldsymbol{m}_k = \boldsymbol{\mu}_k^{\mathrm{Re}} + \mathrm{j}\,\boldsymbol{\mu}_k^{\mathrm{Im}} \tag{6.86}$$

$$\boldsymbol{M}_k = \boldsymbol{m}_k \boldsymbol{m}_k^{\mathrm{H}} + \boldsymbol{\Sigma}_k \tag{6.87}$$

with $\boldsymbol{\mu}_k^{\mathrm{Re}} = [\mu_{k,1}^{\mathrm{Re}}, \ldots, \mu_{k,d}^{\mathrm{Re}}]^{\mathrm{T}}$. In Appendix F.2, it is shown that

$$\mu_{k,i}^{\mathrm{Re}} = \frac{1}{2\sqrt{\pi}} \frac{\exp\left(-\alpha_{k,i}^{\mathrm{Re},2}\left(\ell_k\right)\right) - \exp\left(-\beta_{k,i}^{\mathrm{Re},2}\left(\ell_k\right)\right)}{\Phi\left(\sqrt{2}\alpha_{k,i}^{\mathrm{Re}}\left(\ell_k\right)\right) - \Phi\left(\sqrt{2}\beta_{k,i}^{\mathrm{Re}}\left(\ell_k\right)\right)}. \tag{6.88}$$

Correspondingly, $\boldsymbol{\mu}_k^{\mathrm{Im}} = [\mu_{k,1}^{\mathrm{Im}}, \ldots, \mu_{k,d}^{\mathrm{Im}}]^{\mathrm{T}}$ with

$$\mu_{k,i}^{\mathrm{Im}} = \frac{1}{2\sqrt{\pi}} \frac{\exp\left(-\alpha_{k,i}^{\mathrm{Im},2}\left(\ell_k\right)\right) - \exp\left(-\beta_{k,i}^{\mathrm{Im},2}\left(\ell_k\right)\right)}{\Phi\left(\sqrt{2}\alpha_{k,i}^{\mathrm{Im}}\left(\ell_k\right)\right) - \Phi\left(\sqrt{2}\beta_{k,i}^{\mathrm{Im}}\left(\ell_k\right)\right)}. \tag{6.89}$$

The second term of $\boldsymbol{M}_k$ in Eq. (6.87) is diagonal, i.e.

$$\boldsymbol{\Sigma}_k = \mathrm{diag}\left(\sigma_{k,1}, \ldots, \sigma_{k,d}\right)$$

whose i-th diagonal element can be expressed as (see Appendix F.3)

$$\sigma_{k,i} = \tau_{k,i}^{\mathrm{Re}} + \tau_{k,i}^{\mathrm{Im}} \tag{6.90}$$

with

$$\tau_{k,i}^{\mathrm{Re}} = \frac{1}{2} - \mu_{k,i}^{\mathrm{Re},2} + \frac{1}{2\sqrt{\pi}} \frac{\alpha_{k,i}^{\mathrm{Re}}\left(\ell_k\right)\exp\left(-\alpha_{k,i}^{\mathrm{Re},2}\left(\ell_k\right)\right) - \beta_{k,i}^{\mathrm{Re}}\left(\ell_k\right)\exp\left(-\beta_{k,i}^{\mathrm{Re},2}\left(\ell_k\right)\right)}{\Phi\left(\sqrt{2}\alpha_{k,i}^{\mathrm{Re}}\left(\ell_k\right)\right) - \Phi\left(\sqrt{2}\beta_{k,i}^{\mathrm{Re}}\left(\ell_k\right)\right)}$$

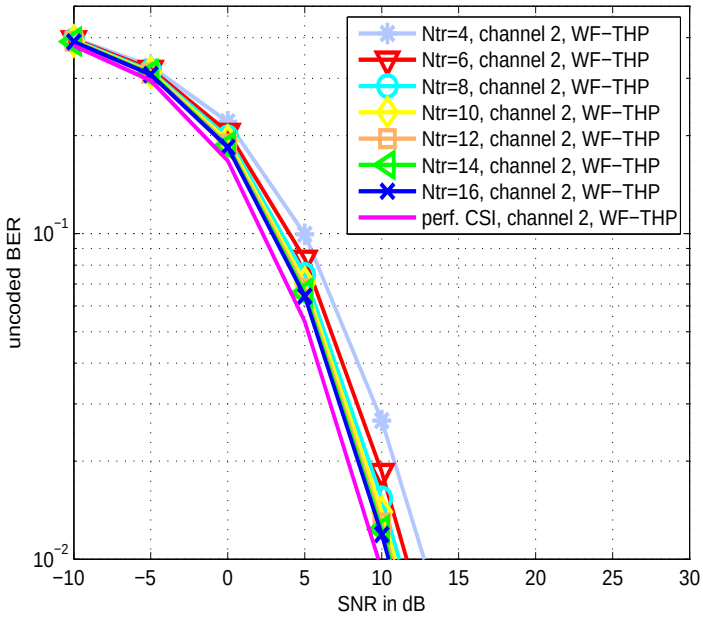


Figure 6.7: Effect of Estimation Error on the Proposed Robust WF–THP with Approach III from Section 6.3 as a Function of Different Training Lengths in an Urban Macrocell Environment.

and, correspondingly,

$$\tau_{k,i}^{\mathrm{Im}} = \frac{1}{2} - \mu_{k,i}^{\mathrm{Im},2} + \frac{1}{2\sqrt{\pi}} \frac{\alpha_{k,i}^{\mathrm{Im}}\left(\ell_k\right) \exp\left(-\alpha_{k,i}^{\mathrm{Im},2}\left(\ell_k\right)\right) - \beta_{k,i}^{\mathrm{Im}}\left(\ell_k\right) \exp\left(-\beta_{k,i}^{\mathrm{Im},2}\left(\ell_k\right)\right)}{\Phi\left(\sqrt{2}\alpha_{k,i}^{\mathrm{Im}}\left(\ell_k\right)\right) - \Phi\left(\sqrt{2}\beta_{k,i}^{\mathrm{Im}}\left(\ell_k\right)\right)}.$$

The above results enable us to compute the conditional covariance matrix

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]} &= \boldsymbol{R}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]} - \boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]} \boldsymbol{\mu}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]}^{\mathrm{H}} \\ &= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_{\mathrm{opt},k} \boldsymbol{\Phi}_k^2 \boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} + \boldsymbol{V}_{\mathrm{opt},k} \boldsymbol{\Phi}_k^2 \boldsymbol{\Sigma}_k \boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} \\ &= \boldsymbol{C}_{\boldsymbol{h},k} + \boldsymbol{V}_{\mathrm{opt},k} \boldsymbol{\Phi}_k^2 \boldsymbol{\Upsilon}_k \boldsymbol{V}_{\mathrm{opt},k}^{\mathrm{H}} \end{aligned} \tag{6.91}$$

where $\boldsymbol{\Upsilon}_k = \boldsymbol{\Sigma}_k - \mathbf{I}_d$. Note that the non-zero elements of the diagonal matrix $\boldsymbol{\Upsilon}_k \in \mathbb{R}_{0,+}^{d \times d}$ only depend on the properties of $\mathrm{Q}_k'(\bullet)$. They can therefore be computed in advance and stored as parameters of $\mathrm{Q}_k'(\bullet)$. The first and the second term in the second line of Eq. (6.91) come from the erroneous knowledge about $\boldsymbol{h}_k$, if we had $\tilde{\boldsymbol{h}}_k$. But since we only have $\tilde{\boldsymbol{h}}_{\mathrm{Q},k}$ available, the variance of the error is increased by the third term in Eq. (6.91).

As seen in this section, the uncertain knowledge about the channel at the transmitter is modeled by the conditional probability density function $f_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]}(\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q])$ whose

covariance matrix is given by Eq. (6.91). Therefore, although we consider the channel as being random we are able to exploit the statistical dependence between the channel and the fed–back information (see Chapter 4). This goal can be achieved by extending the classical precoder optimization with a mean with respect to the channel conditional on the fed–back information. The conditional mean introduces a regularization of the solution that makes it more robust to CSI errors.

When taking the conditional mean of the MSE, we always encounter the conditional mean of the channel and the conditional mean of the channel Gram [see Chapter 5, e.g. Eq. (5.10)] which can be written respectively as [see Eqs. (6.83) and (6.91)]

$$\mathrm{E}\left[\boldsymbol{H}\,\middle|\,\tilde{\boldsymbol{H}}_{\mathrm{Q}}\right] = \left[\boldsymbol{\mu}_{\boldsymbol{h}_1[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},1}[q]},\ldots,\boldsymbol{\mu}_{\boldsymbol{h}_K[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},K}[q]}\right]^{\mathrm{T}} = \hat{\boldsymbol{H}} \tag{6.92}$$

$$\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\,\middle|\,\tilde{\boldsymbol{H}}_{\mathrm{Q}}\right] = \hat{\boldsymbol{H}}^{\mathrm{H}}\hat{\boldsymbol{H}} + \boldsymbol{C}_{\boldsymbol{\Theta}} \tag{6.93}$$

where $\tilde{\boldsymbol{H}}_{\mathrm{Q}} = [\tilde{\boldsymbol{h}}_{\mathrm{Q},1}[q], \tilde{\boldsymbol{h}}_{\mathrm{Q},2}[q], \ldots, \tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]]$ and $\boldsymbol{C}_{\boldsymbol{\Theta}} = \sum_{k=1}^{K} \boldsymbol{C}^{*}_{\boldsymbol{\Theta},k} = \sum_{k=1}^{K} \boldsymbol{C}^{*}_{\boldsymbol{h}_k[q]|\tilde{\boldsymbol{h}}_{\mathrm{Q},k}[q]}$. Notice that for MMSE designs, no other conditional moments of the channel are necessary.

6.4 Simulations

This section presents the results of several computer simulations carried out to assess the proposed MU-MISO system with robust precoding and limited feedback channels as shown in Fig. 6.18. In this section, we study the BER performance achieved with the three precoding schemes depicted in this figure: robust Wiener linear precoding (*rob. WF-LP*), robust Wiener Tomlinson-Harashima precoding (*rob. WF-THP)*, and robust Wiener vector precoding (*rob. WF-VP*). Note that we only use the MMSE criterion instead of the zero-forcing criterion since zero-forcing clearly leads to suboptimum solutions, as demonstrated in Chapter 3.

We consider a MU-MISO system with $N = 4$ antennas at the transmitter and $K = 4$ single antenna users. Performance is evaluated in terms of uncoded *Bit Error Rate* (BER) versus *Signal to Noise Ratio* (SNR). The results are the mean of $5{,}000$ channel realizations with 50 QPSK modulated symbols being transmitted in each channel realization. A delay of $D = 2$ slots is considered for all the users, which are not fixed-located but moving at a given speed. The Doppler frequency is normalized with respect to the slot period and is calculated by taking into account that f_{slot} is $1{,}500$ Hz and that the center frequency is 2 GHz. We consider three different environments following the 3GPP *Spatial Channel Model* (SCM) [36]:

- channel 1 (*SCM* 1): suburban macrocell environment;

- channel 2 (*SCM* 2): urban macrocell environment;
- channel 3 (*SCM* 3): urban microcell environment.

We consider channel 2 in most of the results presented in this section due to its intermediate BER performance and diversity. The BER curves are obtained after averaging 100 channel covariance matrices. Finally, for reasons of simplicity we assume perfect CSI at the receiver for calculating the MMSE coefficients.

The first results in Figs. 6.7– 6.13 were obtained using the third feedback design discussed in Section 6.3.

First, we carry out some preliminary simulations to select the size of the training sequence. Fig. 6.7 shows the uncoded BER for robust THP over an urban macrocell environment (channel 2) and different training sequence lengths in order to illustrate the performance degradation caused by channel estimation errors. In this computer experiment, this is the only error source in the system. As a compromise between training sequence length and performance degradation, we pick the value $N_{\text{tr}} = 6$ for our subsequent simulations, which introduces a 2 dB loss at the BER operation point 10^{-2} with respect to the case of perfect CSI, as can be seen in Fig. 6.7.

The performance of robust THP schemes in channel 2 for different user speeds is plotted in Fig. 6.8. Rank reduction is applied and only $d = 2$ complex coefficients per user are transmitted through the feedback channel. These coefficients are scalarly quantized using 6 bits (3 bits per real dimension) which yields 12 bits per user. Fig. 6.8 considers the speed values of 10, 30, and 60 km/h which correspond to normalized Doppler frequencies of 0.0123, 0.0370, and 0.0741, respectively. It is apparent that, as expected, the faster the fading, the more the performance degrades. Fig. 6.8 also plots the uncoded BER when *Random Vector Quantization* (RVQ) is applied instead of scalar quantization with the same number of 12 bits per user. Each user moves at a speed of 30 km/h for the RVQ curve. Note that in RVQ the stored user's codebook contains channel vectors. As expected, the system performance is better when RVQ is used. This is because RVQ carries out a joint quantization that uses a much larger codebook ($2^{12} = 4{,}096$ entries per user) and compares an N-dimensional vector with 4,096 complex vectors in order to choose the closest one for each channel realization and each channel covariance matrix. Its computational complexity is thus much higher than that of scalar quantization, where the search is reduced to a comparison with $2^3 = 8$ scalar values for the real and imaginary parts of each fed–back coefficient. For the considered number of 12 fed–back bits per user, it is clear that the performance of RVQ for medium and high SNR must be better than that obtained with scalar quantization.

Fig. 6.9 shows the influence of the different errors sources considered throughout this work on the uncoded BER. Again, robust THP over channel 2 with a user speed of 30 km/h is considered. Obviously, each new error source adds a greater degradation in

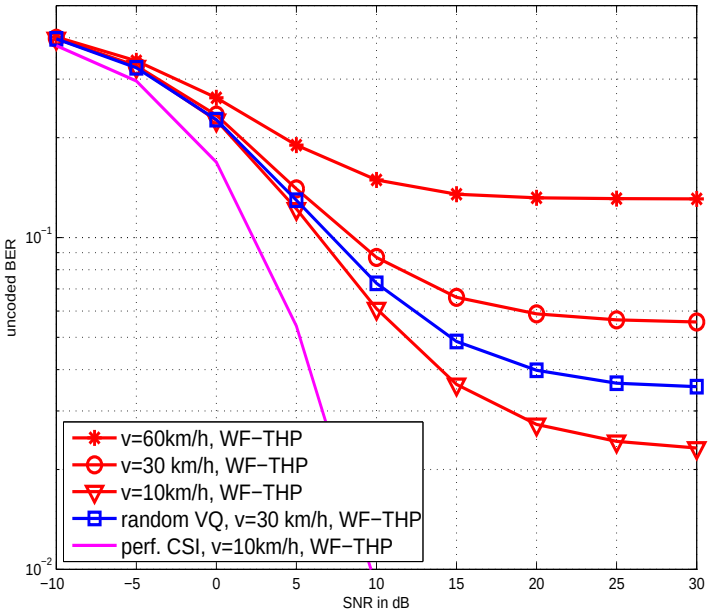


Figure 6.8: Effect of User Speed on the Proposed Robust WF–THP with Approach III from Section 6.3 in an Urban Macrocell Environment with All Errors and 12 Bits per User.

performance to the previous one. Note the performance degradation when moving from $d = 3$ to $d = 2$ truncated coefficients. Also, note the performance loss as the number of bits per user decreases. Nevertheless, truncation to $d = 2$ coefficients and $L = 12$ fed–back bits per user ensure a suitable system performance (BER is about 7×10^{-2} at an SNR of 10 dB) with the enormous advantage of noticeably reducing the feedback channel overhead. This overhead reduction becomes even more appreciable as the number of transmitting antennas increases. In the subsequent computer experiments in this section, we will use $d = 2$ and $L = 12$ as system parameters.

Fig. 6.10 plots the performance of *Linear Precoding* (LP), *Tomlinson-Harashima Precoding* (THP), and *Vector Precoding* (VP) robust schemes for the three different scenarios described in [36]. All error sources are considered, i.e. estimation, quantization, truncation, and delay errors inherent to the feedback transmission. Obviously, the performance for channel 1 (suburban macrocell) is much better than that for channel 2 (urban macrocell). And the performance for channel 2 is again better than that for channel 3 (urban microcell). This is because the spatial correlation in channel 1 is considerably larger than in channel 3 (with channel 2 in between), i.e. the third and fourth channel eigenvalues are negligible in the case of channel 1 whereas they have significant values for channel 3 and even for channel 2. Thus, performance degradation due to truncation to

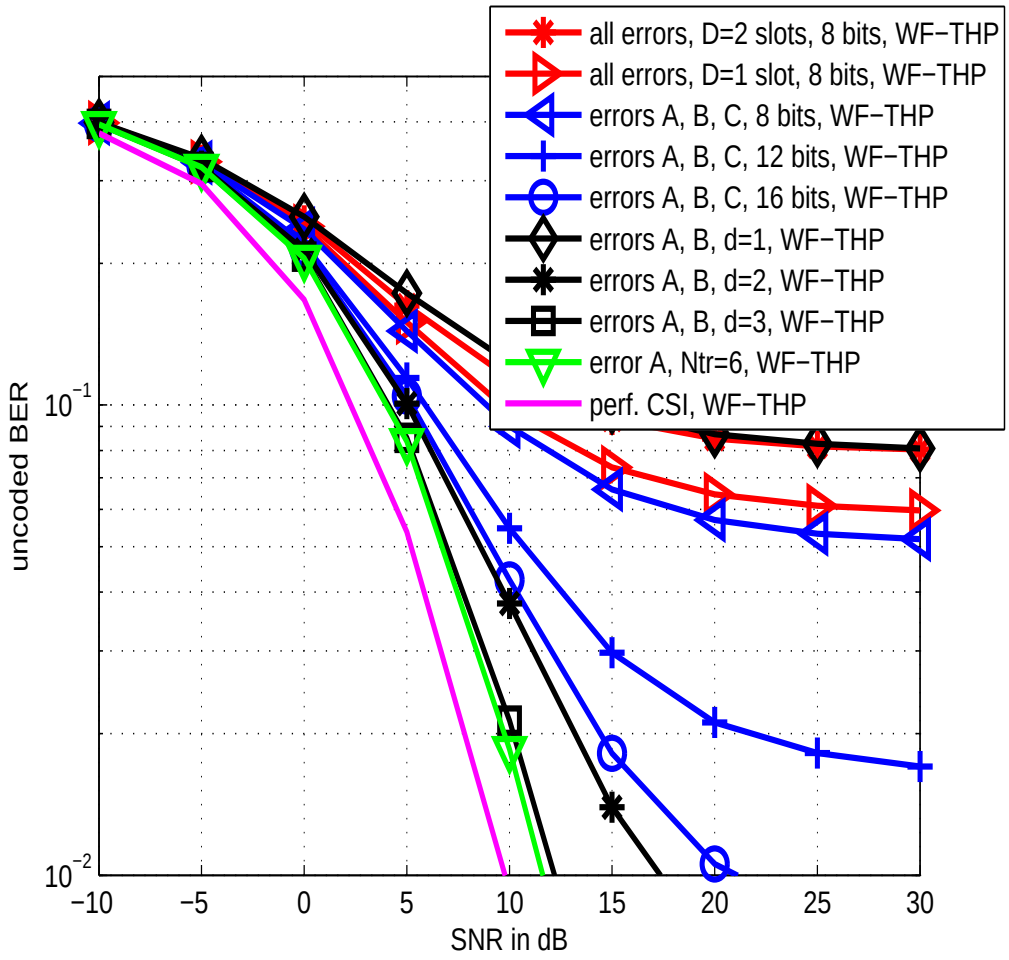


Figure 6.9: Effect of Different Types of Error on the Proposed Robust WF–THP with Approach III from Section 6.3 in an Urban Macrocell Environment. Error A: Estimation; Error B: Rank Reduction; Error C: Quantization; All Errors: Estimation, Rank Reduction, Quantization, and Delay.

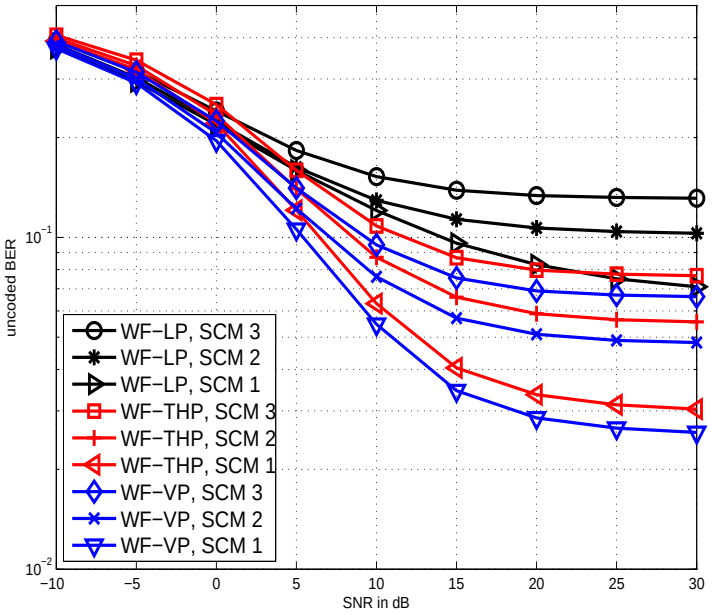


Figure 6.10: BER Performance for Different Types of 3GPP Channel Model with the Proposed Robust Precoding and Approach III from Section 6.3 with 12 Bits per User.

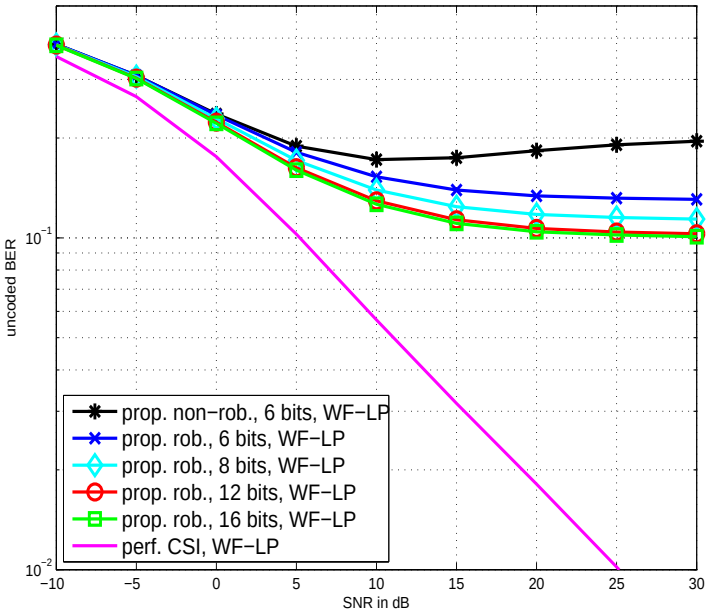


Figure 6.11: BER vs. SNR for MU–MISO Wiener Linear Precoding with Approach III from Section 6.3 in an Urban Macrocell Environment.

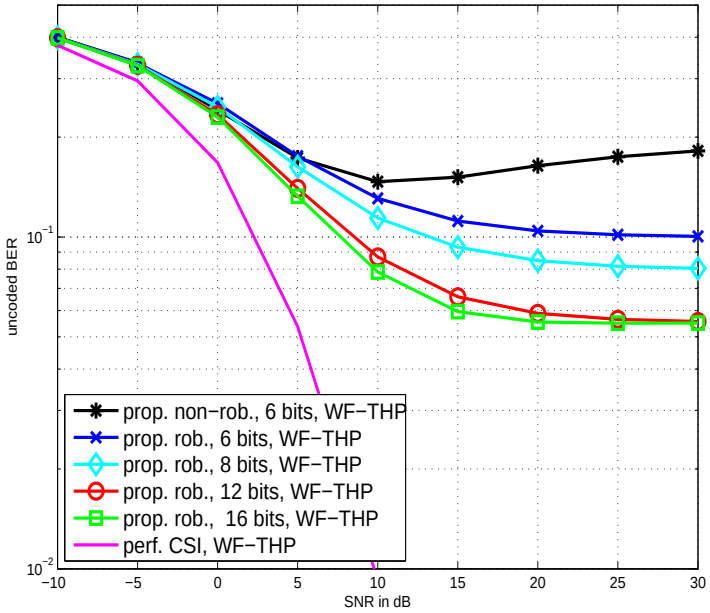


Figure 6.12: BER vs. SNR for MU–MISO Wiener THP with Approach III from Section 6.3 in an Urban Macrocell Environment.

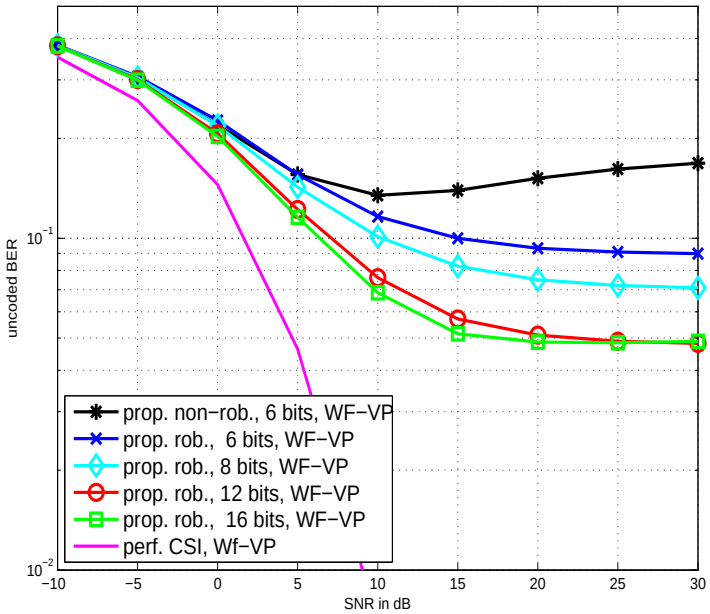


Figure 6.13: BER vs. SNR for MU–MISO Wiener VP with Approach III from Section 6.3 in an Urban Macrocell Environment.

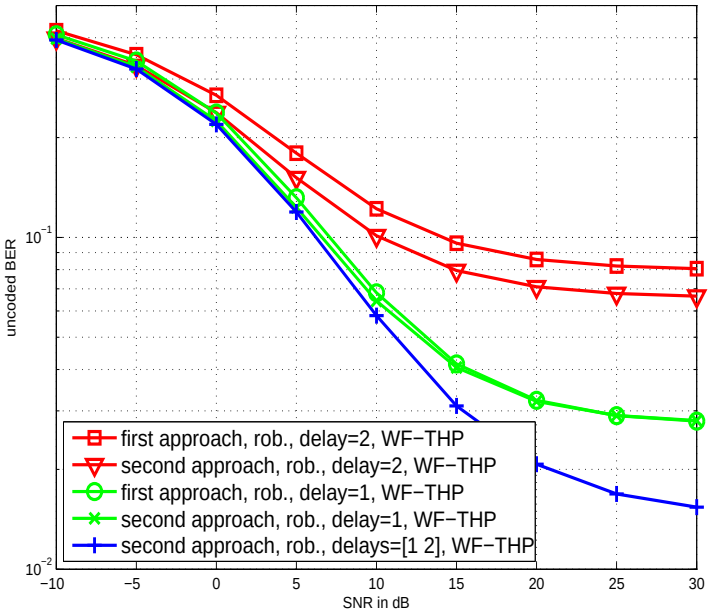


Figure 6.14: BER Performance Improvement with Approach II from Section 6.2 for Limited Feedback as a Function of the Number of Delayed Channels.

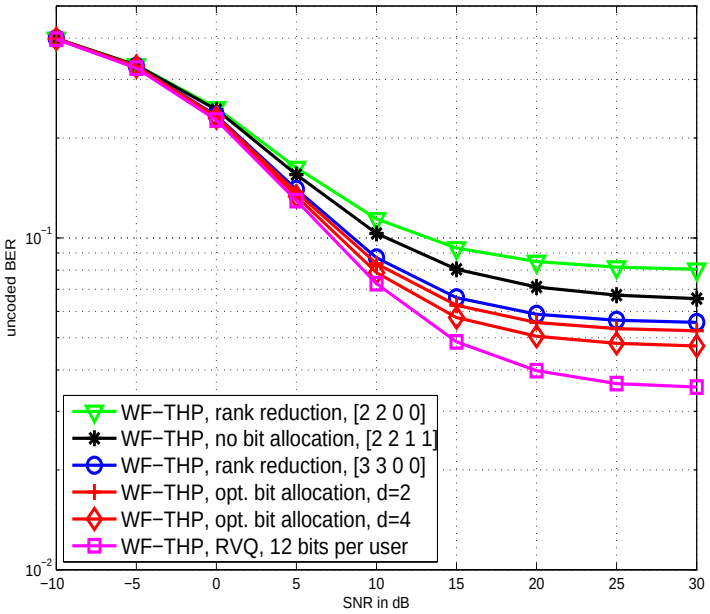


Figure 6.15: BER Performance vs. SNR with Approach III from Section 6.3 for Limited Feedback.

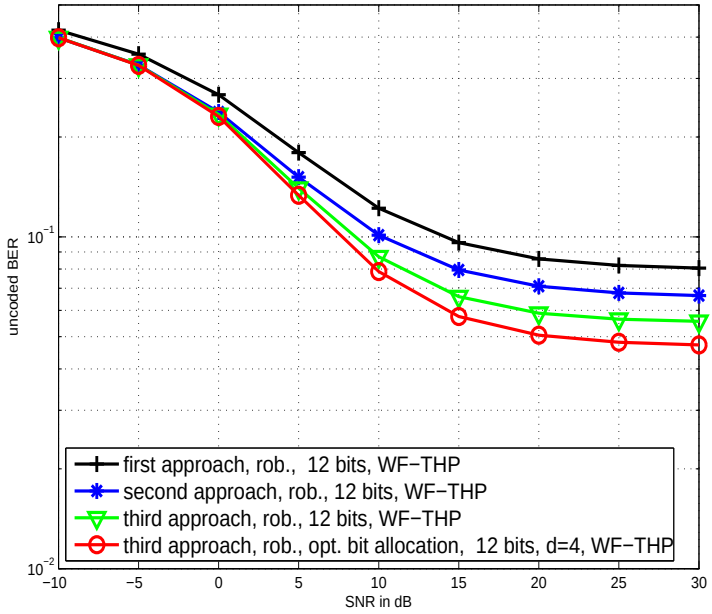


Figure 6.16: BER Comparison vs. SNR of Approaches I, II, and III in Sections 6.1, 6.2, and 6.3, respectively.

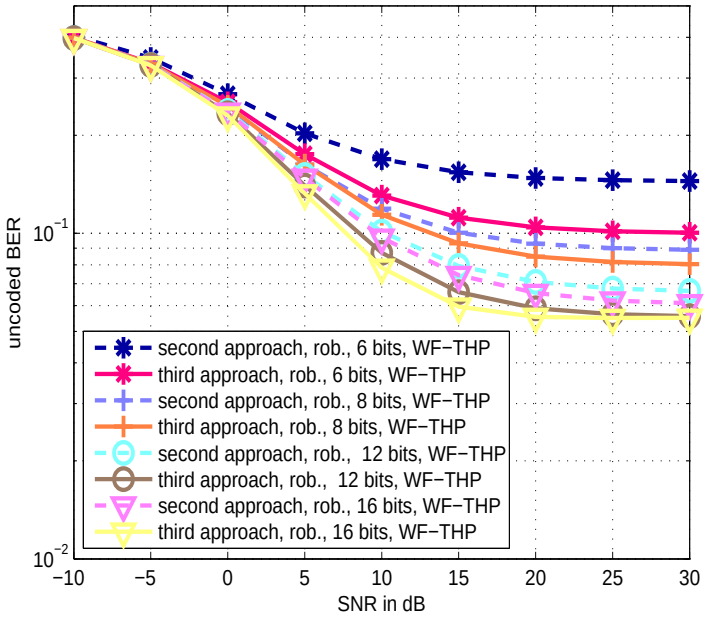


Figure 6.17: BER Comparison vs. SNR of Approaches II and III (see Sections 6.2 and 6.3).

$d = 2$ is more severe in channel 3 than in channel 1. When comparing the three precoding schemes considered, LP exhibits the worst performance for the robust design, as is also the case for perfect CSI. The achieved performance of VP is always better than that of THP, although it is quite similar. Note that the complexity of VP is considerably greater (due to the search in the lattice), which motivates the utilization of the suboptimum robust THP schemes instead.

Figs. 6.11, 6.12, and 6.13 show the improvement of our robust schemes with respect to the non–robust ones. It is apparent from these figures that the non–robust curves go up for high SNR due to the sensitivity of these schemes to imperfect CSI. The advantage of using the robust schemes, which provide a performance improvement and compensate the imperfect CSI knowledge caused by the different error sources, is also apparent. In these simulation results, different scalar codebooks (a codebook of $m = 4$ entries for coding the real and imaginary part of the first coefficient and of $m = 2$ entries for the second one, and $m = 4$, $m = 8$, or $m = 16$ codebook entries for coding each real and imaginary part of the two coefficients sent to the transmitter) have been used, i.e. we are employing 6, 8, 12, and 16 bits per user, respectively. Clearly, if the number of bits is increased, the BER reduces because the errors due to the quantization process are smaller. However, with a codebook of reasonable size, for example with only 8 entries or 12 bits per user, we already obtain a good BER performance. Moreover, the improvement in BER is almost negligible for larger codebooks, which have the enormous disadvantage of reducing the compression rate for the CSI sent through the feedback channel and at the same time considerably increasing the storage capability required at the receivers [103].

Fig. 6.14 shows the improvement in BER performance when using the second feedback design discussed in Section 6.2. Moreover, the results are shown when we consider more past channel versions for the robust design as found in Subsection 6.2.1. It is obvious that the usage of a higher amount of CSI reduces the mismatch between the true channel and the erroneous channel since the uncertainty is decreased.

Fig. 6.15 shows the BER performance corresponding to the third approach, i.e. for the joint MSE optimization that includes the quantizer parameters in the optimization (see Section 6.3). As expected, no bit allocation (the bits are spread uniformly over the coefficients) leads to worse performance than optimum bit allocation, since this latter strategy allocates the bits in the sense of minimizing the MSE. Again, the gain due to vector quantization compared to scalar quantization is apparent, but at the cost of substantially increasing the computational complexity at the user end.

Finally, Figs. 6.16 and 6.17 show a comparison related to the BER performance for the limited feedback THP approaches described in this chapter. Clearly, the two approaches that perform some MSE optimization lead to better performance than the first approach from Section 6.1. Moreover, the gains depend on the number of fed–back bits (see Fig. 6.17).

6.5 Conclusions

In this chapter, we have investigated the compression of *Channel State Information* (CSI) data in a MU-MISO system with precoding and limited feedback channel. Three different types of precoder have been considered: *Linear Precoding* (LP), *Tomlinson-Harashima Precoding* (THP), and *Vector Precoding* (VP). The fed–back CSI is employed to design a robust non-zero-forcing precoder, i.e. a robust *minimum MSE* (MMSE) precoder. Four sources of error have been considered: channel estimation, truncation for rank reduction, quantization, and feedback delay. The error modeling has allowed us to formulate robust designs for each precoding scheme with a performance considerably better than that of conventional non–robust schemes. All the designs proposed in this chapter are based on the mismatch between the full CSI knowledge and the erroneous CSI knowledge at the transmitter.

First, we investigated a very basic limited feedback design that does not minimize the final BER performance since it is not based on an MSE minimization. However, the results are excellent taking into account the simplicity of the error model developed. Next, we introduced an MSE minimization that included all the aspects of feedback with the exception of the quantizer parameters. Thus, we considered a uniform quantizer with the errors produced by the quantization process being modeled as an additive and independent noise. The results of the previous approach are improved on by this new approach, although it is still not optimum since we are not including the codebook entries and the decision boundaries into a joint MSE optimization. Finally, we focused on a Bayesian framework together with a joint MSE optimization based on a CSI metric that allows us to obtain an adequate statistical characterization of the errors on fed–back CSI. We obtained the very useful result that the optimal estimator and rank reduction only depend on the channel statistics and are independent of the quantizer used. The distortion is a diagonally weighted squared error, and thus the Lloyd algorithm can be employed to compute the quantizer. As shown with some computer simulations, this latter approach achieves better BER performance than the approach that does not include the quantizer parameters in the MSE optimization. Moreover, the trade-off between computational complexity and final performance is reasonable and the bit allocation strategy presented, resulting from the joint MSE optimization, clearly outperforms all the previous approaches.

Therefore, the simulation results show how these techniques work well in MU-MISO time-varying channels with limited feedback, given that a minimum amount of information is transmitted through the reverse channel, leading to good BER performance nevertheless. One of the major contributions of this chapter is to have found the channel vector PDF conditional on the fed–back coefficients, which is the basis of our robust precoding, i.e. to achieve a Bayesian approach for error modeling. Moreover, the advantage of the new robust design proposed is even greater if we exploit, by means

of an adequate procedure, the fact that the input is Gaussian in order to design the initial codebook according to the Lloyd algorithm.

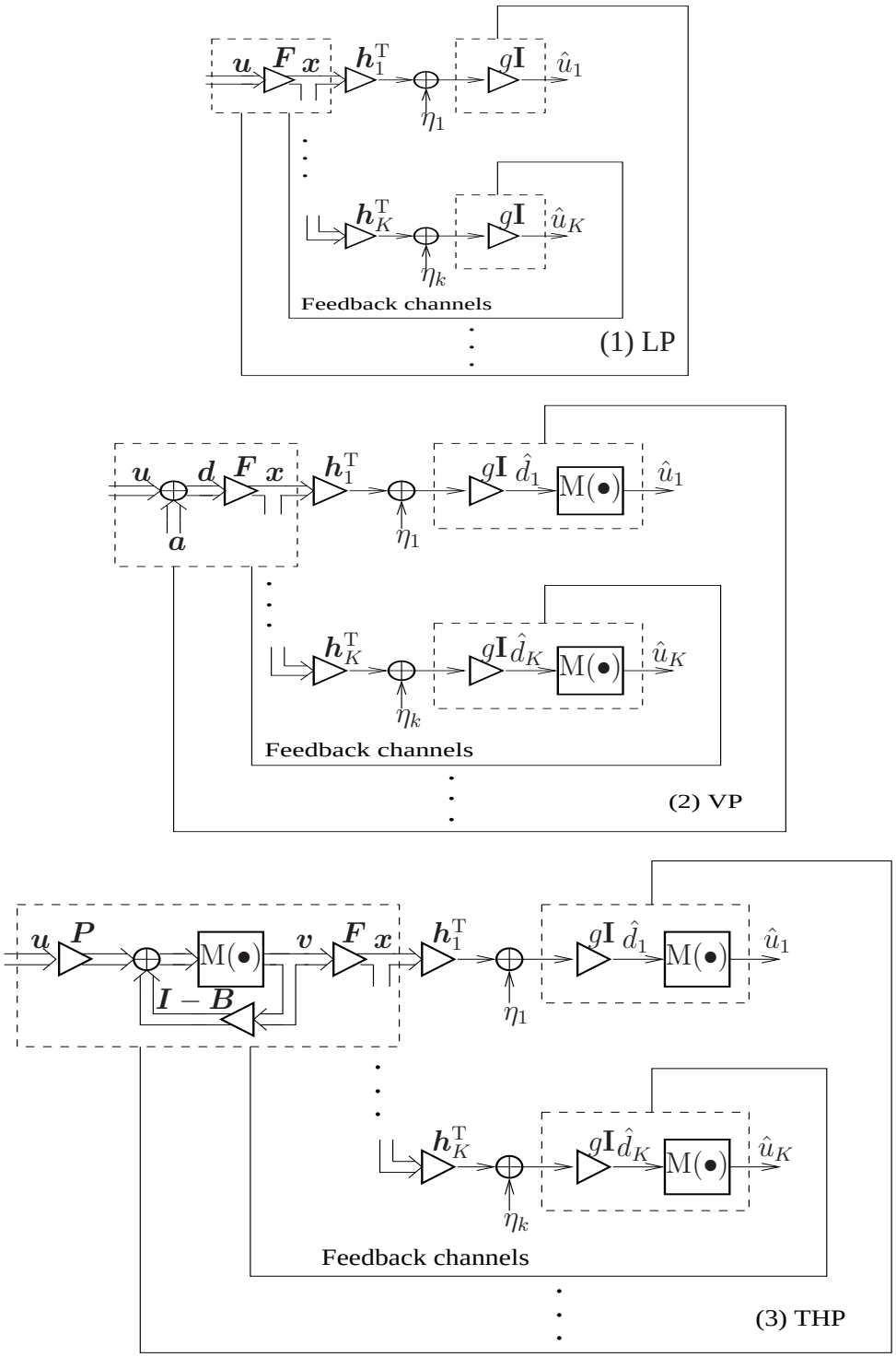


Figure 6.18: Precoding Schemes Based on Feedback Resulting in Imperfect CSIT.

Chapter 7

Feedback Design based on Precoding MSE

Signal separation in the vector *Broadcast Channel* (BC) requires some information about the channel state at the transmitter. In many cases, such as FDD systems, this *Channel State Information* (CSI) must be fed back from the receivers to the transmitter by means of a feedback channel, as already introduced in the previous chapter.

In this chapter, we jointly design the channel estimators and the quantizers at the receivers together with the precoder at the transmitter based on a precoder-centric criterion, i.e. the minimization of an MSE metric appropriate for the precoder design [120]. This is in contrast to Chapter 6, where the quantizer design was based on a CSI-MSE metric.

The procedure is as follows. First, the estimator is designed to minimize the MSE between the transmitted symbols and the symbols recovered by the users including the precoder averaged over all possible channel realizations, where a given quantizer is assumed. Interestingly, the estimators resulting from this joint formulation are independent of the codebook used and are equal to the estimators obtained previously, even though the design is no longer based on a CSI-MSE metric.

On the other hand, the codebook entries are the precoders employed. These precoders are found by minimizing the precoder-MSE conditional on the fed–back index. The use of white estimates (by dropping the coloring and the square root of the respective covariance matrix) and a restriction to rectangular regions leads to a simple computation of the conditional means necessary for the precoding design step. The most difficult part of the proposed scheme is the design of the partition cells. The cell boundaries are designed by minimizing the precoder-MSE conditional on the quantizer input. Finally, each user feeds back the index of a set of precoders and the intersection of the sets performed at the transmitter gives the appropriate precoder to be used during the transmission. Since the quantizers of the different receivers have to work separately, the metric for the

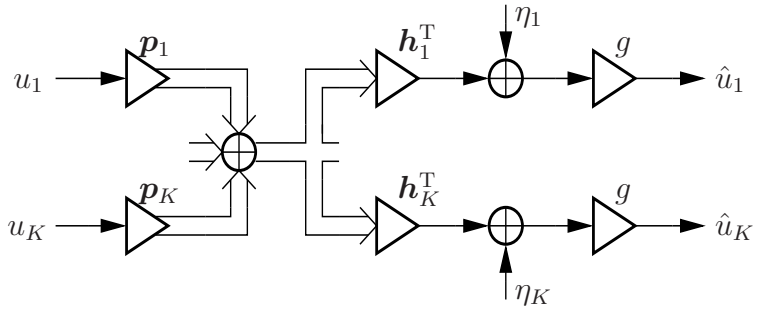


Figure 7.1: MU–MISO System Model for Linear Precoding.

computation of the partition cells cannot be expressed as a simple squared error depending on the quantizer output and its computation is quite complex as shown in this chapter.

To this end, we firstly introduce the proposed system for linear precoding and the feedback system to be optimized in order to reduce the overhead of the reverse channel. Next, we obtain the optimum estimators, partition cells, codebook entries, and linear precoders from the joint optimization based on a metric that includes the overall parameters. We also focus on how to implement bit allocation in this case, and on how we can solve the problems related to its computational complexity by means of a heuristic strategy. Finally, some computer simulations are carried out to illustrate the performance of this approach in terms of BER.

For the sake of brevity, we omit in this chapter the index q to indicate the time slot related to the block fading channel and also the index n to denote each one of the N_{B} time samples spaced with the symbol period, T_{s}, inside each slot (see Chapter 2).

7.1 System Model

Our final goal is the design of a precoder for the vector BC setup, where a centralized transmitter with N antennas communicates data to K decentralized single antenna receivers. For reasons of simplicity, we restrict ourselves to linear precoding in this chapter. Fig. 7.1 shows the block diagram of a MU–MISO system with linear precoding. Fig. 7.2 depicts the block diagram for the same system but employing a more compact form, i.e. by using a notation that results from the combination of the signals from all users. As shown in this figure, the zero-mean data signal $\boldsymbol{u} \in \mathbb{C}^K$ with unit covariance matrix, i.e. $\boldsymbol{C}_{\boldsymbol{u}} = \mathbf{I}$, is linearly transformed by the precoder $\boldsymbol{P} \in \mathbb{C}^{N \times K}$ to obtain the transmit signal $\boldsymbol{y} \in \mathbb{C}^N$, which propagates over the channel $\boldsymbol{h}_k \in \mathbb{C}^N$ to the k-th receiver and is perturbed by the additive noise η_k. The receiver applies the common receive weight $g \in \mathbb{C}$ to get the estimate $\hat{u}_k$. As shown in Fig. 7.2, combining the signals of the different

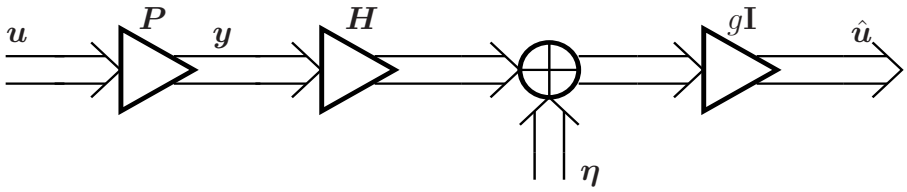


Figure 7.2: System Model for MU–MISO Linear Precoding combining Signals from All Users.

receivers yields

$$\hat{\boldsymbol{u}} = g\boldsymbol{H}\boldsymbol{P}\boldsymbol{u} + g\boldsymbol{\eta} \tag{7.1}$$

where $\hat{\boldsymbol{u}} = [\hat{u}_1, \ldots, \hat{u}_K]^{\mathrm{T}} \in \mathbb{C}^K$, $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_K] \in \mathbb{C}^K$ with $\boldsymbol{\eta} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\eta}})$, and $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]^{\mathrm{T}} \in \mathbb{C}^{K \times N}$.

The constraint for the average total transmitted energy must be satisfied, i.e.

$$\mathrm{E}\left[\|\boldsymbol{P}\boldsymbol{u}\|_2^2\right] = E_{\mathrm{tx}}.$$

The receive weight g is directly derived from this constraint taking into account that the precoder $\boldsymbol{P}$ is factorized as follows

$$\boldsymbol{P} = g^{-1}\boldsymbol{F} \quad \text{with} \quad \boldsymbol{F} \in \mathbb{C}^{N \times K}. \tag{7.2}$$

Accordingly, we have that

$$g = \sqrt{\frac{1}{E_{\mathrm{tx}}} \operatorname{tr}\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\right)}. \tag{7.3}$$

7.2 Limited Feedback Model: Channel Estimation and Quantization

Fig. 7.3 depicts the block diagram of the proposed limited feedback system. In order to obtain the information about the channel state needed to select the precoder, the centralized transmitter sends a sequence of N_{tr} pilot symbols from all transmit antennas. The received noisy pilot symbols are passed through the estimator $\boldsymbol{G}_k \in \mathbb{C}^{N \times N_{\mathrm{tr}}}$ to obtain the input

$$\boldsymbol{z}_k = \boldsymbol{G}_k\left(\boldsymbol{S}\boldsymbol{h}_k + \boldsymbol{\eta}_k\right) \in \mathbb{C}^N \tag{7.4}$$

of the quantizer $\mathrm{Q}_k(\bullet)$ of user k. Here, $\boldsymbol{S} \in \mathbb{C}^{N_{\mathrm{tr}} \times N}$ comprises the pilot symbols and $\boldsymbol{\eta}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\eta},k})$ is the noise of the pilot channel to the k-th receiver. For reasons of simplicity, the feedback channel is assumed to be error-free and without delay. The delay effect has already been studied in the previous chapter and it would be relatively easy

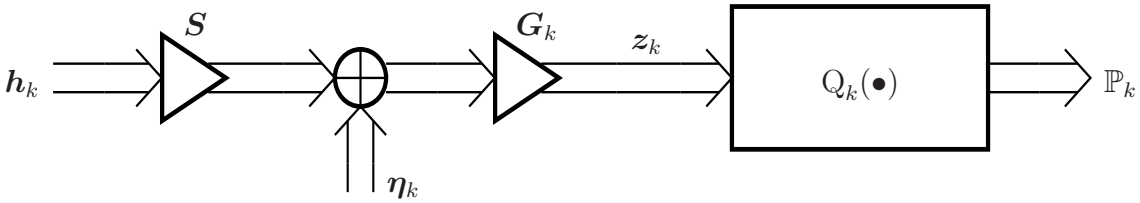


Figure 7.3: System Model for Feedback.

to correct this error, as we will see in the next section, but at the cost of unnecessarily complicating our notation. The channel $\boldsymbol{h}_k \in \mathbb{C}^N$ is assumed to be zero-mean complex Gaussian, i.e. $\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{h},k})$ with the channel covariance matrix of the k-th user $\boldsymbol{C}_{\boldsymbol{h},k} = \mathrm{E}[\boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}}] \in \mathbb{C}^{N \times N}$. Note that the rank reduction is not explicitly included in the notation. This is not necessary since a rank reduction, where a coefficient is dropped, is equivalent to setting the number of partitions corresponding to this coefficient to one. For this reason, we use $\boldsymbol{z}_k$ to denote the information recovered after the estimator in a general way, instead of $\tilde{\boldsymbol{h}}_k$ used in the previous chapter.

After estimation, it is necessary to implement some type of quantizer in order to compress all the information sent through the feedback channel, which is often limited in bandwidth. Contrary to the quantizers used in Chapter 6 where the codebook entries were white channel coefficients, the entries for the quantizers proposed in this chapter contain the precoders of Fig. 7.2, i.e. the quantizer is based on precoders and not on CSI.

Let us initially assume an ideal MU–MISO system where all the users work in a cooperative way, thus making it possible to design the quantizer jointly. In this case, the operation of the overall quantizer is

$$\mathrm{Q}\left(\boldsymbol{z}\right) = \sum_{i=1}^{M} \boldsymbol{P}_i \, \mathrm{S}_i\left(\boldsymbol{z}\right) \tag{7.5}$$

where $\boldsymbol{z} = [\boldsymbol{z}_1^{\mathrm{T}}, \ldots, \boldsymbol{z}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{KN}$ is the overall CSI and the selector function $\mathrm{S}_i(\bullet)$ is 1 if the argument lies in the partition cell $\mathcal{R}_i$, and 0 elsewhere. Each of the M codebook entries $\boldsymbol{P}_i \in \mathbb{C}^{N \times K}$ is a precoder and $\boldsymbol{P}_i$ is chosen if $\boldsymbol{z} \in \mathcal{R}_i$. However, a joint quantization is impossible, since each receiver only has access to its own CSI $\boldsymbol{z}_k$ since we have non-cooperative users in the downlink of a MU–MISO system. Therefore, the partition cell $\mathcal{R}_i \subseteq \mathbb{C}^{KN}$ must be decomposed into subregions $\mathcal{R}_{k,i} \subseteq \mathbb{C}^N$, i.e. $\mathcal{R}_i = \mathcal{R}_{1,i} \times \cdots \times \mathcal{R}_{K,i}$, where $\times$ denotes the Cartesian product defined as

$$\mathcal{R}_i = \mathcal{R}_{1,i} \times \cdots \times \mathcal{R}_{K,i} = \{(\boldsymbol{x}_{1,i}, \ldots, \boldsymbol{x}_{K,i}) \,|\, \boldsymbol{x}_{1,i} \in \mathcal{R}_{1,i}, \ldots, \boldsymbol{x}_{K,i} \in \mathcal{R}_{K,i}\}. \tag{7.6}$$

Here, $\mathcal{R}_i$ denotes the total partition cell corresponding to the i–th codebook entry and $\mathcal{R}_{k,i}$, with $k = 1, \ldots, K$ corresponds to the cell of user k for the i–th codebook entry. The quantizer $\mathrm{Q}_k(\bullet)$ of the k-th user identifies the region $\mathcal{R}_{k,i}$ in which the CSI $\boldsymbol{z}_k$ lies.

Taking into account that we will restrict ourselves to scalar quantization in a later step, we have that

$$\mathcal{R}_{k,i} = \mathcal{R}_{k,i}^{(1)} \times \cdots \times \mathcal{R}_{k,i}^{(N)}$$

with the rectangular region $\mathcal{R}_{k,i}^{(n)} \subseteq \mathbb{C}$ and N being the maximum number of coefficients sent from user k to the transmitter. When the real and imaginary part of the n-th entry $z_{k,n}$ of $\boldsymbol{z}_k$ corresponding to the k-th user's quantizer $\mathrm{Q}_k(\bullet)$ lies in the cells $\mathcal{C}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$ and $\mathcal{C}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$, respectively, the conditions $\alpha_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)} \leq \Re(z_{k,n}) < \beta_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$ or $\alpha_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)} \leq \Im(z_{k,n}) < \beta_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$ are respectively fulfilled. In that case, a set $\mathbb{P}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$ or $\mathbb{P}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$ of indices is chosen, for which it holds that

$$\mathbb{P}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)} = \left\{ i = 1, \ldots, M \mid \mathrm{Re}\left(\mathcal{R}_{k,i}^{(n)}\right) = \mathcal{C}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)} \right\}$$

and

$$\mathbb{P}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)} = \left\{ i = 1, \ldots, M \mid \mathrm{Im}\left(\mathcal{R}_{k,i}^{(n)}\right) = \mathcal{C}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)} \right\}.$$

The fed–back information of user k are the indices $j_k^{(\mathrm{Re},n)}$ and $j_k^{(\mathrm{Im},n)}$ with $n = 1, \ldots, N$. To obtain the output of the quantizer $\mathrm{Q}_k(\bullet)$, the quantized results for the different real and imaginary parts of the entries $z_{k,n}$, $n = 1, \ldots, N$, i.e. $j_k^{(\mathrm{Re},n)}$ and $j_k^{(\mathrm{Im},n)}$, must be combined by simply intersecting the sets $\mathbb{P}_{k,j_k^{(1)}}^{(1)}, \ldots, \mathbb{P}_{k,j_k^{(N)}}^{(N)}$, where $\mathbb{P}_{k,j_k^{(n)}}^{(n)} = \mathbb{P}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)} \cap \mathbb{P}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$:

$$\mathrm{Q}_k(\boldsymbol{z}_k) = \bigcap_{n=1}^{N} \mathbb{P}_{k,j_k^{(n)}}^{(n)}.$$

So, the fed–back information of user k, i.e. the output of its quantizer $\mathrm{Q}_k(\boldsymbol{z}_k)$, is equivalent to a set of indices referring to the precoders that best fit its current channel state. When collecting the fed–back information from all users, the transmitter in the BC finds the index of the final precoder by intersecting the sets of indices of all users. Therefore, the selector function of the overall quantizer in Eq. (7.5) is finally defined as

$$\mathrm{S}_i(\boldsymbol{z}) = \begin{cases} 1 & \text{for } i \in \bigcap_{k=1}^{K} \mathrm{Q}_k(\boldsymbol{z}_k) \\ 0 & \text{else.} \end{cases}$$

Note that the above intersection gives a set with cardinality one due to the properties of the Cartesian product used to split $\mathcal{R}_i$ into $\mathcal{R}_{1,i}, \ldots, \mathcal{R}_{K,i}$ [see Eq. (7.6)]. Fig. 7.4 illustrates how the precoder is selected from the indices sent by each user. This design is totally necessary since the users are not cooperative and, therefore, no single user has information about the others. Remember that the codebook entries are the precoders and the receive weights and not the CSI.

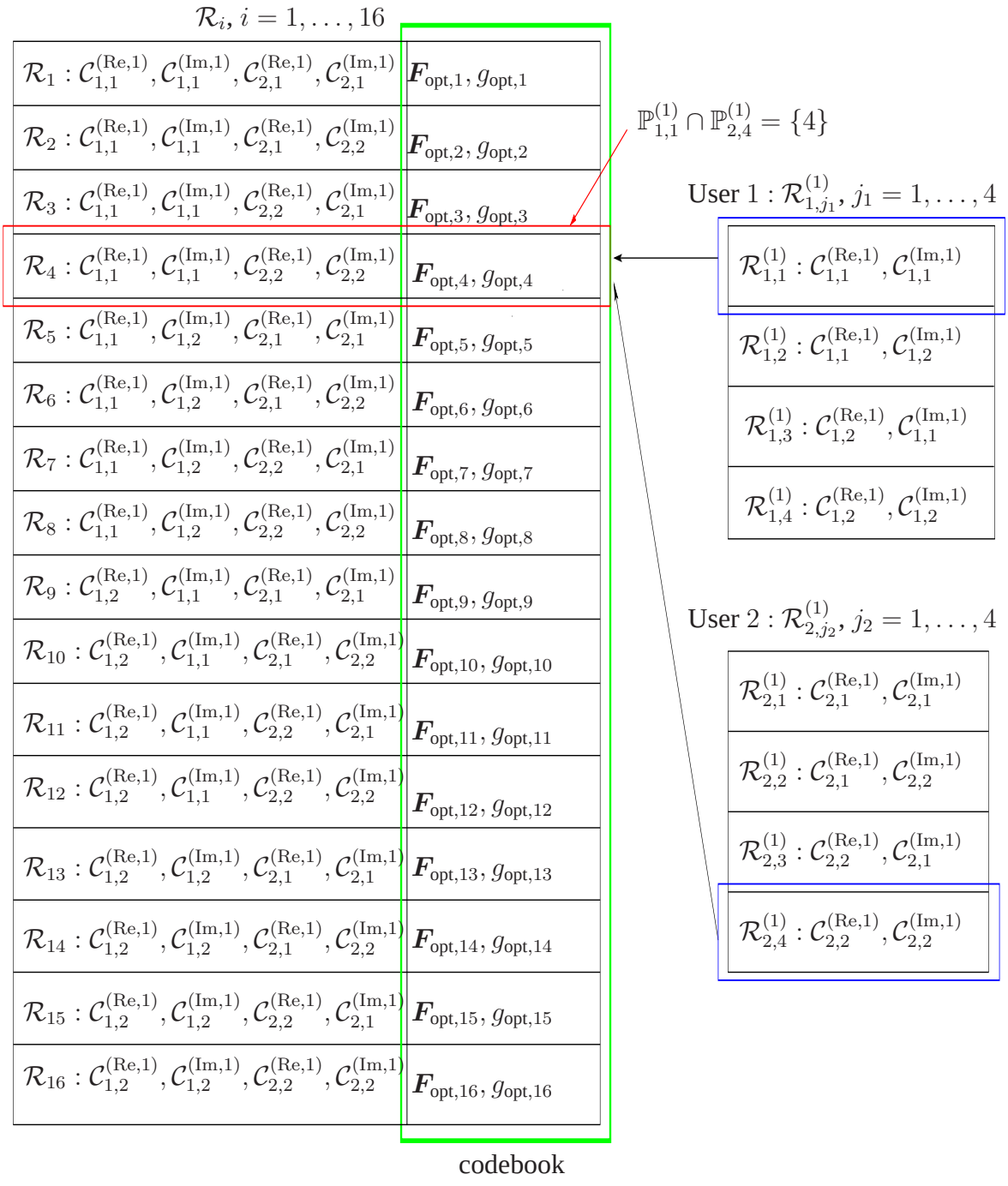


Figure 7.4: Example of Precoder Assignment with $K = 2$ Users, $d = 1$ Coefficient and $N_{\mathrm{bit}} = 2$ Bits per User. (Note that $\mathbb{P}_{1,1}^{(1)} = \mathbb{P}_{1,1}^{(\mathrm{Re},1)} \cap \mathbb{P}_{1,1}^{(\mathrm{Im},1)} = \{1, 2, 3, 4\}$ and $\mathbb{P}_{2,4}^{(1)} = \mathbb{P}_{2,4}^{(\mathrm{Re},1)} \cap \mathbb{P}_{2,4}^{(\mathrm{Im},1)} = \{4, 8, 12, 16\}$. The Number of Codebook Entries is $2^{K \times N_{\mathrm{bit}}} = 2^4 = 16$. The Index of the Overall Quantizer $\mathrm{Q}(\bullet)$ is $i = 4(j_1 - 1) + j_2$).

7.3 Proposed MMSE Optimization

All parts of the feedback system, viz. the estimators $\{\boldsymbol{G}_k\}_{k=1}^K$ and the quantizers $\{\mathrm{Q}_k(\bullet)\}_{k=1}^K$ (i.e. the partition cells $\{\mathcal{R}_i\}_{i=1}^M$ and the precoders $\{\boldsymbol{P}_i\}_{i=1}^M$ together with the weights $\{g_i\}_{i=1}^M$), minimize the precoding MSE

$$\begin{aligned}\mathrm{MSE} &= \mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2\right] = \sum_{i=1}^{M} p_i \,\mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2 \middle| \boldsymbol{z}\in\mathcal{R}_i\right] \\ &= \sum_{i=1}^{M} p_i \left(\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{u}}\right) - 2g_i\Re\left(\mathrm{tr}\left(\mathrm{E}\left[\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right]\boldsymbol{P}_i\boldsymbol{C}_{\boldsymbol{u}}\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\right. \\ &\qquad \left. + g_i^2\,\mathrm{tr}\left(\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right]\boldsymbol{P}_i\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{P}_i^{\mathrm{H}}\right)\right) \\ &= \sum_{i=1}^{M} p_i \left(K - 2g_i\Re\left(\mathrm{tr}\left(\mathrm{E}\left[\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right]\boldsymbol{P}_i\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\right. \\ &\qquad \left. + g_i^2\,\mathrm{tr}\left(\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right]\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\right)\right).\end{aligned} \tag{7.7}$$

Here, p_i denotes the probability that $\boldsymbol{z}\in\mathcal{R}_i$. We assume that $\boldsymbol{u}$ is zero–mean and uncorrelated with unit-variance, i.e. $\boldsymbol{C}_{\boldsymbol{u}}=\mathbf{I}$ for the last equality. Remember that the received signal is given by $\hat{\boldsymbol{u}} = g(\boldsymbol{H}\boldsymbol{P}\boldsymbol{u}+\boldsymbol{\eta})$ [see Eq. (7.1)], where $\boldsymbol{P}$ is the precoder obtained from the overall quantizer, i.e. $\boldsymbol{P}=\mathrm{Q}(\boldsymbol{z})=\sum_{i=1}^M \boldsymbol{P}_i\,\mathrm{S}_i(\boldsymbol{z})$. Note again that we neglect the delay of the feedback in our system model for the sake of brevity.

The optimization problem that we have to solve is

$$\{\{\boldsymbol{G}_k\}_{k=1}^K, \{\boldsymbol{P}_i\}_{i=1}^M, \{\mathcal{R}_i\}_{i=1}^M\} = \underset{\{\{\boldsymbol{G}_k\}_{k=1}^K, \{\boldsymbol{P}_i\}_{i=1}^M, \{\mathcal{R}_i\}_{i=1}^M\}}{\mathrm{argmin}} \mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2\right]. \tag{7.8}$$

Unfortunately, no closed form expressions can be obtained for both the estimators and the quantizers of the feedback systems. However, an alternating optimization can be used to minimize $\mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2\right]$, because closed form expressions for the separate minimizations are available, while the other quantities are kept fixed. Therefore, we start by fixing $\mathcal{R}_i$ and $\boldsymbol{P}_i$ and try to get the closed-form solution for the estimator $\boldsymbol{G}_k$. After that we use the Lloyd algorithm to iteratively optimize the partition cells and codebook entries of the quantizers of each user.

7.3.1 Estimators

In this subsection, the estimator $\boldsymbol{G}_k$ is optimized for given codebook entries (precoders), partition cells, and other estimators, i.e. $\boldsymbol{G}_{\mathrm{opt},k} = \mathrm{argmin}_{\boldsymbol{G}_k}\,\mathrm{MSE}$ (see Eq. (7.7)). Due to [cf. Eq. (7.4)]

$$\boldsymbol{C}_{\boldsymbol{z},k} = \mathrm{E}\left[\boldsymbol{z}_k\boldsymbol{z}_k^{\mathrm{H}}\right] = \boldsymbol{G}_k\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)\boldsymbol{G}_k^{\mathrm{H}}$$

we get the following alternative parameterization of the estimator

$$\boldsymbol{G}_k = \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{X}_k^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \tag{7.9}$$

where the unknown $\boldsymbol{X}_k \in \mathbb{C}^{N_{\mathrm{tr}} \times N}$ has orthonormal columns, i.e. $\boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{X}_k = \mathbf{I}_N$. It is very easy to see that this expression for $\boldsymbol{G}_k$ leads to $\boldsymbol{C}_{\boldsymbol{z},k}$ when we substitute into $\boldsymbol{G}_k(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})\boldsymbol{G}_k^{\mathrm{H}}$. Note that the transformation of $\boldsymbol{S}\boldsymbol{h}_k + \boldsymbol{\eta}_k$ with $(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1/2}$ leads to an uncorrelated signal with unit covariance matrix and the additional transformation with $\boldsymbol{X}_k^{\mathrm{H}}$ again gives an uncorrelated signal with unit covariance matrix irrespective of the choice for $\boldsymbol{C}_{\boldsymbol{z},k}$. Therefore, the optimization with respect to $\boldsymbol{G}_k$ can be split into an optimization with respect to $\boldsymbol{X}_k$ and a subsequent optimization with respect to $\boldsymbol{C}_{\boldsymbol{z},k}$.

Before the minimization of $\mathrm{E}\left[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2\right]$ with respect to $\boldsymbol{X}_k$ can be performed, the MSE $\mathrm{E}\left[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2\right]$ must be rewritten by using the matrix $\boldsymbol{A}_k$ defined as

$$\boldsymbol{A}_k = \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \in \mathbb{C}^{N \times N_{\mathrm{tr}}} \tag{7.10}$$

and by obtaining the conditional moments $\mathrm{E}[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]$ and $\mathrm{E}[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]$. Taking into account that $\boldsymbol{h}_k$ and $\boldsymbol{z}_k$ are jointly Gaussian, we have

$$\begin{bmatrix} \boldsymbol{h}_k \\ \boldsymbol{z}_k \end{bmatrix} \sim \mathcal{N}_{\mathbb{C}} \left(\mathbf{0}, \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{h},k} & \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}^{\mathrm{H}} \\ \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k} & \boldsymbol{C}_{\boldsymbol{z},k} \end{bmatrix} \right)$$

where $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}$ is given by

$$\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k} = \mathrm{E}\left[\boldsymbol{z}_k \boldsymbol{h}_k^{\mathrm{H}}\right] = \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{X}_k^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}. \tag{7.11}$$

Thus, the following conditional moments read as (e.g. [121])

$$\begin{aligned}
\mathrm{E}[\boldsymbol{h}_k|\boldsymbol{z}_k] &= \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{z},k}^{-1} \boldsymbol{z}_k = \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \boldsymbol{X}_k \boldsymbol{C}_{\boldsymbol{z},k}^{-1/2} \boldsymbol{z}_k \\
&= \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{C}_{\boldsymbol{z},k}^{-1/2} \boldsymbol{z}_k \\
\mathrm{E}[\boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}}|\boldsymbol{z}_k] &= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{z},k}^{-1} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k} + \mathrm{E}[\boldsymbol{h}_k|\boldsymbol{z}_k]\, \mathrm{E}[\boldsymbol{h}_k|\boldsymbol{z}_k]^{\mathrm{H}} \\
&= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}} + \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{C}_{\boldsymbol{z}_k}^{-1/2} \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{z},k}^{-1/2,\mathrm{H}} \boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}}.
\end{aligned}$$

Clearly, it holds that $\mathrm{E}[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]) = \mathrm{E}[\mathrm{E}[\boldsymbol{H}|\boldsymbol{z}]|\boldsymbol{z} \in \mathcal{R}_i]$. Therefore, taking into account that $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]^{\mathrm{T}}$, we have

$$\begin{aligned}
\mathrm{E}[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i] &= [\boldsymbol{A}_1 \boldsymbol{X}_1 \boldsymbol{\mu}_{1,i}, \ldots, \boldsymbol{A}_K \boldsymbol{X}_K \boldsymbol{\mu}_{K,i}]^{\mathrm{T}} \qquad (7.12) \\
\mathrm{E}[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i] &= \sum_{k=1}^{K} \left(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k \boldsymbol{X}_k \left(\mathbf{I} - \boldsymbol{R}_{k,i} \right) \boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}} \right)^{\mathrm{T}}
\end{aligned}$$

with

$$\boldsymbol{\mu}_{k,i} = \mathrm{E}\left[\boldsymbol{C}_{\boldsymbol{z},k}^{-1/2}\boldsymbol{z}_k \middle| \boldsymbol{z}_k \in \mathcal{R}_{k,i}\right]$$
$$\boldsymbol{R}_{k,i} = \mathrm{E}\left[\boldsymbol{C}_{\boldsymbol{z},k}^{-1/2}\boldsymbol{z}_k\boldsymbol{z}_k^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{z},k}^{-1/2,\mathrm{H}} \middle| \boldsymbol{z}_k \in \mathcal{R}_{k,i}\right].$$

Note that $\boldsymbol{\mu}_{k,i}$ and $\boldsymbol{R}_{k,i}$ only depend on the choice of $\mathcal{R}_{k,i}$ which are assumed to be given in this subsection. The above results for $\mathrm{E}[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]$ and $\mathrm{E}[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]$ can be substituted into Eq. (7.7). Thus, the MSE for the given codebook entries $\{\boldsymbol{P}_i, g_i\}_{i=1}^M$ and partition cells $\{\mathcal{R}_i\}_{i=1}^M$ is expressed as

$$\begin{aligned}\mathrm{MSE} &= \mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2\right] = \sum_{i=1}^{M} p_i\,\mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2 \middle| \boldsymbol{z} \in \mathcal{R}_i\right] \qquad (7.13)\\ &= \sum_{i=1}^{M} p_i \left(K - 2g_i\Re\left(\mathrm{tr}\left(\left[\boldsymbol{A}_1\boldsymbol{X}_1\boldsymbol{\mu}_{1,i},\ldots,\boldsymbol{A}_K\boldsymbol{X}_K\boldsymbol{\mu}_{K,i}\right]^{\mathrm{T}}\boldsymbol{P}_i\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\right.\\ &\quad \left. + g_i^2\sum_{k=1}^{K}\mathrm{tr}\left(\left(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k\boldsymbol{X}_k\left(\mathbf{I}-\boldsymbol{R}_{k,i}\right)\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\right)^{\mathrm{T}}\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\right)\right).\end{aligned}$$

As mentioned above, when introducing the alternative representation of the estimator $\boldsymbol{G}_k$ in Eq. (7.9), we first find the basis $\boldsymbol{X}_k$ by minimizing the above MSE for fixed $\boldsymbol{C}_{\boldsymbol{z},k}$, i.e.

$$\boldsymbol{X}_{\mathrm{opt},k} = \underset{\boldsymbol{X}_k}{\mathrm{argmin}}\,\mathrm{MSE} \qquad \text{s.t.:} \quad \boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{X}_k = \mathbf{I}_N.$$

The constraint ensures the sub–unitarity of $\boldsymbol{X}_k \in \mathbb{C}^{N_{\mathrm{tr}}\times N}$. The corresponding Lagrangian function reads as

$$L(\boldsymbol{X}_k, \boldsymbol{\Lambda}_k) = \mathrm{MSE} + \mathrm{tr}\left(\boldsymbol{\Lambda}_k\left(\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{X}_k - \mathbf{I}\right)\right)$$

with the Lagrangian multiplier $\boldsymbol{\Lambda}_k \in \mathbb{C}^{N\times N}$, which is Hermitian by definition, as the constraint is Hermitian. A necessary condition for optimality is that

$$\frac{\partial L(\boldsymbol{X}_k, \boldsymbol{\Lambda}_k)}{\partial \boldsymbol{X}_k^{\mathrm{T}}} = \frac{\partial \mathrm{MSE}}{\partial \boldsymbol{X}_k^{\mathrm{T}}} + \boldsymbol{\Lambda}_k\boldsymbol{X}_k^{\mathrm{H}} = \mathbf{0}.$$

From this KKT condition we obtain that [cf. Eq. (7.13)]

$$-p_i\boldsymbol{\mu}_{k,i}\boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{P}_i^{\mathrm{T}}g_i\boldsymbol{A}_k - p_i\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}g_i^2\boldsymbol{P}_i^*\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{A}_k + p_i\boldsymbol{R}_{k,i}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}g_i^2\boldsymbol{P}_i^*\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{A}_k + \boldsymbol{\Lambda}_k\boldsymbol{X}_k^{\mathrm{H}} = \mathbf{0}.$$

Since the range of the first three summands reachable for row vectors multiplied from the left is the span of the rows of $\boldsymbol{A}_k$, the space spanned by the rows of $\boldsymbol{X}_k^{\mathrm{H}}$ must be the same to fulfill the above condition. Thus,

$$\mathrm{range}\left(\boldsymbol{X}_k\right) = \mathrm{range}\left(\boldsymbol{A}_k^{\mathrm{H}}\right). \qquad (7.14)$$

By considering the *Singular Value Decomposition* (SVD) of a matrix $\boldsymbol{B} = \boldsymbol{M}\boldsymbol{D}\boldsymbol{N}^{\mathrm{H}}$, where $\boldsymbol{D}$ is a square diagonal matrix and $\boldsymbol{M}$ and $\boldsymbol{N}$ are unitary or sub–unitary, it is satisfied that the range of $\boldsymbol{B}$ is equal to the range of $\boldsymbol{M}$ [107]. With this property and the SVD of $\boldsymbol{A}_k$ given by

$$\boldsymbol{A}_k = \boldsymbol{V}_k \boldsymbol{\Phi}_k \boldsymbol{W}_k^{\mathrm{H}}$$

with unitary $\boldsymbol{V}_k \in \mathbb{C}^{N \times N}$, diagonal $\boldsymbol{\Phi}_k \in \mathbb{R}^{N \times N}$ whose diagonal elements are positive, and sub–unitary $\boldsymbol{W}_k \in \mathbb{C}^{N_{\mathrm{tr}} \times N}$, we have that range$(\boldsymbol{A}_k^{\mathrm{H}})$ = range$(\boldsymbol{W}_k)$. We can conclude that the optimal basis is expressed as

$$\boldsymbol{X}_{\mathrm{opt},k} = \boldsymbol{W}_k \boldsymbol{U}_k^{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{tr}} \times N} \tag{7.15}$$

to fulfill the condition in Eq. (7.14). The so far undefined unitary $\boldsymbol{U}_k \in \mathbb{C}^{N \times N}$ must be chosen to minimize the precoding MSE in Eq. (7.13). As $\boldsymbol{\Phi}_k \boldsymbol{W}_k^{\mathrm{H}} = \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{A}_k$, the optimal estimator must have the form [cf. Eq. (7.9)]

$$\begin{aligned} \boldsymbol{G}_{\mathrm{opt},k} &= \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{X}_k^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \\ &= \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{U}_k \boldsymbol{\Phi}_k^{-1} \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{A}_k \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \\ &= \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{U}_k \boldsymbol{\Phi}_k^{-1} \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{G}_{\text{MMSE-estim},k} \end{aligned} \tag{7.16}$$

where the conventional linear MMSE estimator is given by

$$\boldsymbol{G}_{\text{MMSE-estim},k} = \boldsymbol{C}_{\boldsymbol{h},k} \boldsymbol{S}^{\mathrm{H}} (\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1}.$$

The output of $\boldsymbol{V}_k^{\mathrm{H}}$ in Eq. (7.16) is uncorrelated and with $\boldsymbol{\Phi}_k^{-1}$, the estimate is white, i.e. with unit variance. Again, as in Chapter 6, some rotation with $\boldsymbol{U}_k$ is applied that does not change the property of unit covariance. Finally, the estimate is colored with $\boldsymbol{C}_{\boldsymbol{z},k}^{1/2}$. This result is quite surprising, since we do not optimize the mean squared error between the true channel and the channel recovered at the transmitter. Instead, the precoding MSE $\mathrm{E}\left[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2\right]$ is minimized [see Eq. (7.8)]. Additionally, the notation introduced in the previous chapter, where the MSE between the true channel and the CSI at the transmitter was minimized, explicitly included a rank reduction. We choose a different formulation now, since a rank reduction can be included by an appropriate restriction of the partition cells $\mathcal{R}_{k,i}$ (i.e. by bit allocation). Therefore, $\boldsymbol{V}_k$ in Eq. (7.16) is square and is not used for rank reduction.

We also see from Eq. (7.16) that the optimal estimator $\boldsymbol{G}_{\mathrm{opt},k}$ can be obtained in closed form except for the covariance matrix $\boldsymbol{C}_{\boldsymbol{z},k}$ and the unitary matrix $\boldsymbol{U}_k$. The optimization of these parts of the estimator is difficult and cannot be found analytically. However, the last stages of the estimator $\boldsymbol{G}_{\mathrm{opt},k}$ can be moved into the quantizer $\mathrm{Q}_k(\bullet)$ as we have

already done in Chapter 6. Thanks to this step, the partition cells $\mathcal{R}_{k,i}$ are just redefined and optimality is not spoilt. Therefore, we can set without loss of optimality that

$$\boldsymbol{G}_{\text{opt},k} = \boldsymbol{\Phi}_k^{-1} \boldsymbol{V}_k^{\text{H}} \boldsymbol{G}_{\text{MMSE-estim},k} \in \mathbb{C}^{N \times N_{\text{tr}}}. \tag{7.17}$$

Note that the optimal estimator is independent of any properties of the codebook and the other estimators. Additionally, note that the output $\boldsymbol{z}_k$ of the estimator $\boldsymbol{G}_{\text{opt},k}$ is white Gaussian. Then, we rename the output of the estimator as $\boldsymbol{w}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$. Due to the relationship of $\boldsymbol{X}_{\text{opt},k}$ and $\boldsymbol{A}_k$ [see Eq. (7.15)], we have

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{X}_k^{\text{H}} \boldsymbol{A}_k^{\text{H}} &= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k \boldsymbol{W}_k \boldsymbol{U}_k^{\text{H}} \boldsymbol{U}_k \boldsymbol{W}_k^{\text{H}} \boldsymbol{A}_k^{\text{H}} \\ &= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_k \boldsymbol{\Phi}_k \boldsymbol{W}_k^{\text{H}} \boldsymbol{W}_k \boldsymbol{W}_k^{\text{H}} \boldsymbol{W}_k \boldsymbol{\Phi}_k \boldsymbol{V}_k^{\text{H}} = \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_k \boldsymbol{\Phi}_k^2 \boldsymbol{V}_k^{\text{H}} \end{aligned}$$

and

$$\boldsymbol{A}_k \boldsymbol{W}_k \boldsymbol{U}_k^{\text{H}} \boldsymbol{R}_{k,i} \boldsymbol{U}_k \boldsymbol{W}_k^{\text{H}} \boldsymbol{A}_k^{\text{H}} = \underbrace{\boldsymbol{A}_k \boldsymbol{W}_k}_{\boldsymbol{V}_k \boldsymbol{\Phi}_k} \boldsymbol{R}_{k,i} \underbrace{\boldsymbol{W}_k^{\text{H}} \boldsymbol{A}_k^{\text{H}}}_{\boldsymbol{\Phi}_k \boldsymbol{V}_k^{\text{H}}}$$

because Gaussian distributions are invariant to unitary rotations (see Appendix D.2).

Bearing in mind the above results, the conditional moments from Eq. (7.12) can be rewritten as

$$\begin{aligned} \text{E}\left[\boldsymbol{H} | \boldsymbol{z} \in \mathcal{R}_i\right] &= \left[\boldsymbol{\mu}_{1,i}, \ldots, \boldsymbol{\mu}_{K,i}\right]^{\text{T}} \\ \text{E}\left[\boldsymbol{H}^{\text{H}} \boldsymbol{H} | \boldsymbol{z} \in \mathcal{R}_i\right] &= \sum_{k=1}^{K} \left(\left(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_k \boldsymbol{\Phi}_k^2 \boldsymbol{V}_k^{\text{H}} \right)^{\text{T}} + \boldsymbol{R}_{k,i}^{\text{T}} \right) \\ &= \sum_{k=1}^{K} \left(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_k \boldsymbol{\Phi}_k^2 \boldsymbol{V}_k^{\text{H}} + \boldsymbol{R}_{k,i} \right)^{\text{T}} \end{aligned}$$

where $\boldsymbol{\mu}_{k,i}$ and $\boldsymbol{R}_{k,i}$ are redefined as

$$\begin{aligned} \boldsymbol{\mu}_{k,i} &= \boldsymbol{V}_k \boldsymbol{\Phi}_k \, \text{E}\left[\boldsymbol{w}_k \,|\, \boldsymbol{w}_k \in \mathcal{R}_{k,i}\right] \\ \boldsymbol{R}_{k,i} &= \boldsymbol{V}_k \boldsymbol{\Phi}_k \, \text{E}\left[\boldsymbol{w}_k \boldsymbol{w}_k^{\text{H}} \,|\, \boldsymbol{w}_k \in \mathcal{R}_{k,i}\right] \boldsymbol{\Phi}_k \boldsymbol{V}_k^{\text{H}}. \end{aligned} \tag{7.18}$$

Now, $\mathrm{E}[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i]$ can be further written as

$$\begin{aligned}
\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right] &= \sum_{k=1}^{K}\left(\boldsymbol{C}_{\boldsymbol{h},k}-\boldsymbol{V}_k\boldsymbol{\Phi}_k^2\boldsymbol{V}_k^{\mathrm{H}}+\boldsymbol{V}_k\boldsymbol{\Phi}_k\,\mathrm{E}\left[\boldsymbol{w}_k\boldsymbol{w}_k^{\mathrm{H}}\,|\,\boldsymbol{w}_k\in\mathcal{R}_{k,i}\right]\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}\right)^{\mathrm{T}} \\
&= \sum_{k=1}^{K}\Big(\boldsymbol{C}_{\boldsymbol{h},k}-\boldsymbol{V}_k\boldsymbol{\Phi}_k^2\boldsymbol{V}_k^{\mathrm{H}}+\boldsymbol{\mu}_{k,i}\boldsymbol{\mu}_{k,i}^{\mathrm{H}} \\
&\quad +\boldsymbol{V}_k\boldsymbol{\Phi}_k\,\mathrm{E}\left[\left(\boldsymbol{w}_k-\boldsymbol{\Phi}_k^{-1}\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\mu}_{k,i}\right)\left(\boldsymbol{w}_k-\boldsymbol{\Phi}_k^{-1}\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\mu}_{k,i}\right)^{\mathrm{H}}\,|\,\boldsymbol{w}_k\in\mathcal{R}_{k,i}\right]\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}\Big)^{\mathrm{T}} \\
&= \sum_{k=1}^{K}\big(\underbrace{\boldsymbol{C}_{\boldsymbol{h},k}-\boldsymbol{V}_k\boldsymbol{\Phi}_k^2\boldsymbol{V}_k^{\mathrm{H}}}_{\boldsymbol{C}_{\text{estim},k}}+\boldsymbol{\mu}_{k,i}\boldsymbol{\mu}_{k,i}^{\mathrm{H}}+\underbrace{\boldsymbol{V}_k\boldsymbol{\Phi}_k\boldsymbol{C}_{\mathrm{Q},k,i}\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}}_{\boldsymbol{C}_{\text{quantize},k,i}}\big)^{\mathrm{T}} \qquad (7.19) \\
&= \sum_{k=1}^{K}\left(\boldsymbol{C}_{\boldsymbol{h},k}+\boldsymbol{\mu}_{k,i}\boldsymbol{\mu}_{k,i}^{\mathrm{H}}-\boldsymbol{V}_k\boldsymbol{\Phi}_k\boldsymbol{\Gamma}_{k,i}\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}\right)^{\mathrm{T}} \qquad (7.20)
\end{aligned}$$

where the relationship $\boldsymbol{C}_{\boldsymbol{y}|\boldsymbol{x}}=\mathrm{E}[(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}})(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}})^{\mathrm{H}}|\boldsymbol{x}]=\mathrm{E}[\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}|\boldsymbol{x}]-\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}}^{\mathrm{H}}$ is applied. $\boldsymbol{C}_{\text{estim},k}$ is the MSE error matrix due to the estimation error and $\boldsymbol{C}_{\text{quantize},k,i}$ is the error covariance matrix due to the quantization error. The matrix $\boldsymbol{\Gamma}_{k,i}=\mathbf{I}-\boldsymbol{C}_{\mathrm{Q},k,i}\in\mathbb{R}^{0,+}$ depends only on the quantizer parameters. Note that when we assume perfect channel knowledge at the receiver, i.e. when there are no errors caused by estimation, $\boldsymbol{C}_{\text{estim},k}=\mathbf{0}$, and when there is no limited rate for the feedback, i.e. no quantization errors, we have that $\boldsymbol{C}_{\text{quantize},k,i}=\mathbf{0}$. Therefore, the regularization that is introduced due to imperfect CSI at the transmitter is given by $\boldsymbol{C}_{\text{estim},k}+\boldsymbol{C}_{\text{quantize},k,i}$. Remember that the effect of feedback delay is omitted in Eq. (7.19). In the event that we assume a simple Jakes model, we would have that [cf. Eq. (6.5)]

$$\mathrm{E}\left[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[\nu]\right]=\mathrm{J}_0(2\pi f_{\text{D,max},k}D/f_{\text{slot}})\boldsymbol{C}_{\boldsymbol{h},k}=r_k\boldsymbol{C}_{\boldsymbol{h},k}$$

with the slot index q, the delay in slots $D=q-\nu$, the maximum Doppler frequency of the k-th user $f_{\text{D,max},k}$, the slot rate f_{slot}, and the zero-th order Bessel function of the first kind $\mathrm{J}_0(\bullet)$ [34]. The factor r_k in the last equality is implicitly defined. Note that the delay can be neglected by considering a speed value of $v=0$ km/h ($r_k=1$). As done in Chapter 6, the only impact is that this term r_k must be included into the expression of $\boldsymbol{A}_k$ in Eq. (7.10) since the input of the quantizer $\boldsymbol{z}_k$ given by Eq. (7.4) is obtained from outdated channel vectors and therefore $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}=r_k\boldsymbol{C}_{\boldsymbol{z},k}^{1/2}\boldsymbol{X}_k^{\mathrm{H}}\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}+\boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1/2}\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}$ [cf. Eq. (7.11)].

For the sake of notational brevity, we introduce

$$\begin{aligned}\boldsymbol{M}_i &= [\boldsymbol{\mu}_{1,i}, \ldots, \boldsymbol{\mu}_{K,i}]^{\mathrm{T}} \in \mathbb{C}^{K\times N}\\ \boldsymbol{C}_{\text{estim}} &= \sum_{k=1}^{K} \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_k \boldsymbol{\Phi}_k^2 \boldsymbol{V}_k^{\mathrm{H}} \in \mathbb{C}^{N\times N}\\ \boldsymbol{C}_{\text{quantize},i} &= \sum_{k=1}^{K} \boldsymbol{V}_k \boldsymbol{\Phi}_k \boldsymbol{C}_{\mathrm{Q},k,i} \boldsymbol{\Phi}_k \boldsymbol{V}_k^{\mathrm{H}} \in \mathbb{C}^{N\times N}. \end{aligned} \tag{7.21}$$

The precoding MSE for the optimal estimators is therefore

$$\begin{aligned}\text{MSE} = \sum_{i=1}^{M} p_i \Big(& K - 2\Re\left(\operatorname{tr}\left(\boldsymbol{M}_i g_i \boldsymbol{P}_i\right)\right) + g_i^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\\ & + g_i^2 \operatorname{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}} \boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right) \boldsymbol{P}_i \boldsymbol{P}_i^{\mathrm{H}}\right)\Big). \end{aligned} \tag{7.22}$$

In the following, we assume that the optimal estimators $\boldsymbol{G}_{\text{opt},k}$, $k = 1, \ldots, K$ are employed, i.e. the precoding MSE of Eq. (7.22) has to be minimized by the parameters of the quantizers.

Notice that the conditional moments provided by this scheme are equal to the conditional moments obtained for the joint optimization based on a CSI-metric (see Chapter 6), which is not especially remarkable since we obtained the same estimator for both approaches.

7.3.2 Codebook Entries: Precoders

The precoder $\boldsymbol{P}_i$ and the receiver weight g_i minimize the precoding MSE of Eq. (7.22) under a transmit power constraint for given partition cells $\mathcal{R}_i$, $i = 1, \ldots, M$

$$\{\boldsymbol{P}_{\text{opt},i}, g_{\text{opt},i}\} = \underset{\{\boldsymbol{P}_i, g_i\}}{\operatorname{argmin}}\, \text{MSE} \qquad \text{s.t.: } \mathrm{E}\left[\|\boldsymbol{P}_i \boldsymbol{u}\|_2^2\right] \leq E_{\text{tx}}. \tag{7.23}$$

Without destroying optimality, we make a change of variables and set $\boldsymbol{P}_i = g_i^{-1} \boldsymbol{F}_i$. Consequently, the Lagrangian function reads as

$$\begin{aligned} L\left(\boldsymbol{F}_i, g_i, \lambda\right) = \sum_{i=1}^{M} p_i \Big(& K - 2\Re\left(\operatorname{tr}\left(\boldsymbol{M}_i \boldsymbol{F}_i\right)\right) + g_i^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\\ & + \operatorname{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}} \boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right) \boldsymbol{F}_i \boldsymbol{F}_i^{\mathrm{H}}\right) + \lambda\left(g_i^{-2} \|\boldsymbol{F}_i\|_{\mathrm{F}}^2 - E_{\text{tx}}\right) \end{aligned} \tag{7.24}$$

with the Lagrangian multiplier $\lambda \in \mathbb{R}^{0,+}$.

One KKT condition is obtained by deriving with respect to g_i^*, which is assumed to be real. Equating this derivative to zero yields

$$\frac{\partial L\left(\bullet\right)}{\partial g_i^*} = 2g_i \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) - 2\lambda g_i^{-3} \left\|\boldsymbol{F}_i\right\|_{\mathrm{F}}^2 = 0$$

which leads to $\lambda = g_i^2 \frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{g_i^{-2}\|\boldsymbol{F}_i\|_{\mathrm{F}}^2} > 0$. Excluding the trivial solution, we can follow that the transmit energy constraint is active, and therefore $g_i^{-2}\|\boldsymbol{F}_i\|_{\mathrm{F}}^2 = E_{\mathrm{tx}}$ and $\lambda = g_i^2 \frac{\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\mathrm{tx}}}$.

When we set the derivative with respect to $\boldsymbol{F}_i^*$ to zero, we obtain the following KKT condition

$$\begin{aligned}\frac{\partial L\left(\bullet\right)}{\partial \boldsymbol{F}_i^*} &= -\boldsymbol{M}_i^{\mathrm{H}} + \left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)\boldsymbol{F}_i + \frac{\lambda}{g_i^2}\boldsymbol{F}_i \\ &= -\boldsymbol{M}_i^{\mathrm{H}} + \left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}} + \xi\mathbf{I}_N\right)\boldsymbol{F}_i = \mathbf{0} \end{aligned} \tag{7.25}$$

where $\xi = \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})/E_{\mathrm{tx}}$. This result together with the transmit power constraint leads to the optimal precoder (codebook entry) corresponding to the i-th partition cell $\mathcal{R}_i$ given by

$$\boxed{\begin{aligned}\boldsymbol{F}_{\text{opt},i} &= \left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}} + \xi\mathbf{I}\right)^{-1}\boldsymbol{M}_i^{\mathrm{H}} \\ g_{\text{opt},i} &= \sqrt{\frac{1}{E_{\mathrm{tx}}}\operatorname{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)^{-2}\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i\right)}.\end{aligned}} \tag{7.26}$$

This result is the *centroid condition*. Although the optimization of Eq. (7.23) gives the weight $g_{\text{opt},i}$, we use MMSE receiver weights instead to correct the phase as described in Section 5.5 and get an approximately coherent detection.

Note that the solution for the precoder is inherently robust against errors, since the respective error covariance matrices regularize the pseudo inversion in the definition of $\boldsymbol{F}_{\text{opt},i}$.

Due to the expectations $\mathrm{E}\left[\boldsymbol{w}_k \,|\, \boldsymbol{w}_k \in \mathcal{R}_{k,i}\right]$ for $k = 1, \ldots, K$ [see Eqs. (7.18) and (7.21)], the computation of the precoder $\boldsymbol{F}_i$ is difficult for general partition cells $\mathcal{R}_{1,i}, \ldots, \mathcal{R}_{K,i}$, i.e. using *vector quantization*. However, by restricting ourselves to *scalar quantization*, the integration over the rectangular regions $\mathcal{R}_{k,i}^{(n)}$ can be solved in closed form (see Appendix F, [118]). Note that this precoder is basically the same precoder as that obtained for the approach in Chapter 6, based on the CSI MSE metric. Both linear precoders are robust against errors in CSI by means of regularization terms. Contrary to the CSI MSE metric, however, where the precoder is based on already optimized and fixed partition cells that are independent from the channel statistics [1], the joint design according to the precoder MSE metric shown in this chapter optimizes the precoder and

[1] Neglecting the effect of bit allocation.

the partition cells using the Lloyd algorithm. The Lloyd algorithm switches between the precoder design and the partition cell computation and converges to locally optimum precoders and regions since every step reduces the MSE, and the MSE is lower bounded. Note that now both precoders and partition cells must be recomputed as soon as channel statistics change, which is more appropriate in the sense of MSE minimization.

7.3.3 Partition Cells

Since the other estimators' inputs are unknown to the quantizer of user ℓ, the regions $\mathcal{R}_{\ell,i}$ of the ℓ-th quantizer minimize the distortion $\mathrm{d}_\ell = \mathrm{E}[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 | \boldsymbol{z}_\ell]$ for the given codebook entries $\boldsymbol{P}_i$ and g_i, $i = 1, \ldots, M$. Motivated by the fact that $\boldsymbol{z}_\ell \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$, i.e. its entries are uncorrelated, and that the computation of the precoders is difficult for vector quantization, we restrict ourselves to scalar quantization, i.e. the entries of $\boldsymbol{z}_\ell$ are quantized separately. In this case, the partition cells $\mathcal{C}_{\ell,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$ and $\mathcal{C}_{\ell,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$ (that is, their corner coordinates $\alpha_{\ell,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$, $\beta_{\ell,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$, $\alpha_{\ell,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$, and $\beta_{\ell,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$) of the scalar quantizers for, respectively, real and imaginary parts of the n-th entry $z_{\ell,n}$ of $\boldsymbol{z}_\ell$ minimize the distortions

$$\mathrm{d}_\ell^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) = \mathrm{E}\left[\, \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 \big|\, \Re[z_{\ell,n}] \right] = \sum_{j=1}^{M_\ell^{(n)}} \mathrm{S}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])\, \mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) \quad (7.27)$$

and

$$\mathrm{d}_\ell^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) = \mathrm{E}\left[\, \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 \big|\, \Im[z_{\ell,n}] \right] = \sum_{j=1}^{M_\ell^{(n)}} \mathrm{S}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])\, \mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) \quad (7.28)$$

respectively. Here, $M_\ell^{(n)}$ is the number of codebook entries for the quantizer of $\Re[z_{\ell,n}]$ and $\Im[z_{\ell,n}]$ (in our example of Fig. 7.4: $M_\ell^{(n)} = 2$); $\mathrm{S}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ is one for $\Re[z_{\ell,n}] \in \mathcal{C}_{\ell,j}^{(\mathrm{Re},n)}$ and zero elsewhere; and $\mathrm{S}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$ is one for $\Im[z_{\ell,n}] \in \mathcal{C}_{\ell,j}^{(\mathrm{Im},n)}$ and zero elsewhere. As a result of computing these expressions for each $z_{\ell,n}$, we can obtain the indices $j_\ell^{(\mathrm{Re},n)}$ and $j_\ell^{(\mathrm{Im},n)}$ that minimize these distortions. Note that, given the n–th quantizer input of user ℓ, $z_{\ell,n}$, we assume that the other quantizer inputs $z_{k,n}$, with $k \neq \ell$, are unknown and, therefore, it is necessary to average over all the possible $z_{k,n}$. Although the other entries $z_{\ell,\nu}$ with $\nu \neq n$ are known to receiver ℓ, also over these quantities is averaged, since scalar quantizers are used. However, their corresponding cells are given since the codebook design is centralized at the transmitter and stored at both the transmitter and all the receivers.

The distortions due to the j-th codebook entry for both real and imaginary entries of the input $z_{\ell,n}$ read respectivley as

$$\begin{aligned}\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) = \sum_{i\in\mathbb{P}_{\ell,j}^{(\mathrm{Re},n)}} \frac{p_i}{p_{\ell,j}^{(\mathrm{Re},n)}} \Bigg(& K + g_i^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) - \sum_{k=1,k\neq\ell}^{K} 2\Re\left(\boldsymbol{\mu}_{k,i}^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{e}_k\right) \\ & - 2\Re\left(\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n),\mathrm{T}}\boldsymbol{F}_i\boldsymbol{e}_\ell\right) + \operatorname{tr}\left(\boldsymbol{C}_{\text{estim}}^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{H}}\right) + \sum_{k=1,k\neq\ell}^{K} \operatorname{tr}\left(\boldsymbol{R}_{k,i}^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{H}}\right) \\ & + \operatorname{tr}\left(\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n),\mathrm{T}}\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{H}}\right)\Bigg) \end{aligned} \tag{7.29}$$

and

$$\begin{aligned}\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) = \sum_{i\in\mathbb{P}_{\ell,j}^{(\mathrm{Im},n)}} \frac{p_i}{p_{\ell,j}^{(\mathrm{Im},n)}} \Bigg(& K + g_i^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) - \sum_{k=1,k\neq\ell}^{K} 2\Re\left(\boldsymbol{\mu}_{k,i}^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{e}_k\right) \\ & - 2\Re\left(\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n),\mathrm{T}}\boldsymbol{F}_i\boldsymbol{e}_\ell\right) + \operatorname{tr}\left(\boldsymbol{C}_{\text{estim}}^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{H}}\right) + \sum_{k=1,k\neq\ell}^{K} \operatorname{tr}\left(\boldsymbol{R}_{k,i}^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{H}}\right) \\ & + \operatorname{tr}\left(\boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n),\mathrm{T}}\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{H}}\right)\Bigg) \end{aligned} \tag{7.30}$$

where $\boldsymbol{F}_i = g_i\boldsymbol{P}_i$ and $\boldsymbol{e}_k$ denotes the k-th column of the $K \times K$ identity matrix. $p_{\ell,j}^{(\mathrm{Re},n)}$ and $p_{\ell,j}^{(\mathrm{Im},n)}$ are the probabilities of $\Re[z_{\ell,n}] \in \mathcal{C}_{\ell,j}^{(\mathrm{Re},n)}$ and $\Im[z_{\ell,n}] \in \mathcal{C}_{\ell,j}^{(\mathrm{Im},n)}$, respectively. Additionally, the conditional moments $\boldsymbol{\mu}_{\ell,i}$ and $\boldsymbol{R}_{\ell,i}$ under the conditions $\Re[z_{\ell,n}]$ and $\Im[z_{\ell,n}]$, denoted by $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n)}$, $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n)}$, $\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n)}$, and $\boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n)}$, can be found as follows:

$$\begin{aligned}\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n)} &= \boldsymbol{V}_\ell\boldsymbol{\Phi}_\ell\,\mathrm{E}\left[\boldsymbol{z}_\ell \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Re[z_{\ell,n}]\right] \\ \boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n)} &= \boldsymbol{V}_\ell\boldsymbol{\Phi}_\ell\,\mathrm{E}\left[\boldsymbol{z}_\ell \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Im[z_{\ell,n}]\right]\end{aligned} \tag{7.31}$$

and

$$\begin{aligned}\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n)} &= \boldsymbol{V}_\ell\boldsymbol{\Phi}_\ell\,\mathrm{E}\left[\boldsymbol{z}_\ell\boldsymbol{z}_\ell^{\mathrm{H}} \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Re[z_{\ell,n}]\right]\boldsymbol{\Phi}_\ell\boldsymbol{V}_\ell^{\mathrm{H}} \\ \boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n)} &= \boldsymbol{V}_\ell\boldsymbol{\Phi}_\ell\,\mathrm{E}\left[\boldsymbol{z}_\ell\boldsymbol{z}_\ell^{\mathrm{H}} \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Im[z_{\ell,n}]\right]\boldsymbol{\Phi}_\ell\boldsymbol{V}_\ell^{\mathrm{H}}.\end{aligned} \tag{7.32}$$

Following the *nearest neighbor condition*, the partition cells $\mathcal{C}_{\ell,j}^{(\mathrm{Re},n)}$ must be chosen such that for any input $\Re[z_{\ell,n}]$ the minimum distortion $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ is picked by the quantizer. Equivalently, for the imaginary part, the partition cells $\mathcal{C}_{\ell,j}^{(\mathrm{Im},n)}$ are chosen such that for any input $\Im[z_{\ell,n}]$ the quantizer uses the minimum distortion $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$. Since $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n)}$ and $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n)}$ are linear, and $\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n)}$ and $\boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n)}$ are quadratic functions of $\Re[z_{\ell,n}]$ and $\Im[z_{\ell,n}]$, respectively, the distortions $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$ are

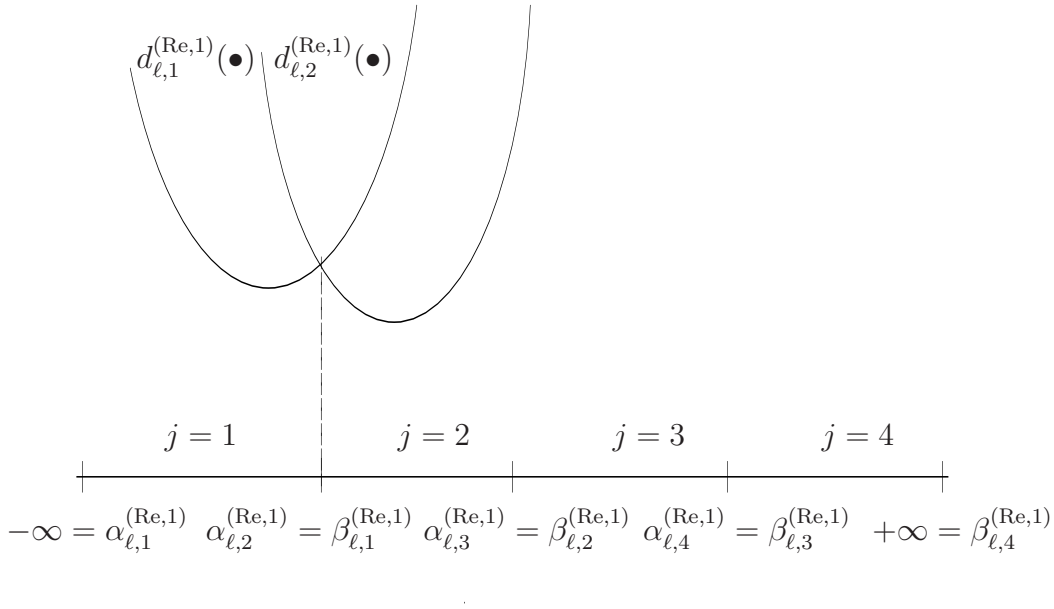


Figure 7.5: Example of Precoder Assignment with $d = 1$ Coefficient and 2 Bits per User.

also quadratic functions. Thus, for the real part of $z_{\ell,n}$ the optimal cell borders $\alpha_{\ell,j}^{(\mathrm{Re},n)}$, $\alpha_{\ell,j}^{(\mathrm{Im},n)}$, $\beta_{\ell,j}^{(\mathrm{Re},n)}$, and $\beta_{\ell,j}^{(\mathrm{Im},n)}$ are simply the roots of the quadratic polynomial equations $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$. Again, similarly for the imaginary part of $z_{\ell,n}$, the region boundaries are given by the roots of the quadratic polynomials $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$. Fig. 7.5 illustrates the method above proposed for obtaining those optimal cell borders from the roots of quadratic functions.

7.3.4 Codebook Computation

Although the estimators and the quantizers are jointly optimized by minimizing the precoding MSE in Eq. (7.7), we only have to compute the codebook parameters iteratively, since the estimators are independent of the choice of codebook and can be found in Eq. (7.17). For the computation of the codebook parameters, we use the Lloyd algorithm (e.g. [108,119]), i.e. we alternately optimize the precoders by using the centroid condition in Eq. (7.26) and optimize the partition cells following the nearest neighbor condition as discussed in the previous subsection. Since the MSE in Eq. (7.22) is reduced in every step and the MSE is non-negative, the iteration converges.

The Lloyd algorithm is initialized with the solution of Chapter 6, where the quantization was designed by minimizing the CSI MSE and whose quantizers are based on codebooks appropriate for unit variance complex Gaussian inputs. Therefore, the

parameters of these scalar quantizers can be stored and do not have to be recomputed for varying channel statistics. As a consequence, the initialization of the proposed feedback scheme based on the precoding MSE of Eq. (7.22) is very cheap.

Table 7.1 summarizes the overall design procedure for computing the codebook, which is basically a modified version of the Lloyd algorithm. Note that this new codebook has to be recomputed each time that the channel statistics vary.

1. Set $m = 1$
2. Initial codebook $\mathcal{C}_1$ and regions $\{\mathcal{R}_i\}_{i=1}^{M}$ (obtained as in Chapter 6)
3. Set the threshold to stop the iterations ϵ_{min} and set $\epsilon = \infty$

while $\epsilon > \epsilon_{\text{min}}$ **do**

4. obtain the quadratic functions:
   $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$
   $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$
5. (*Nearest Neighbor Condition*) solve the quadratic functions:
   $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) = 0$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) = 0$
   $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) = 0$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) = 0$
   to get the new partition regions $\{\mathcal{R}_i\}_{i=1}^{M}$
6. compute the new conditional channel moments:
   $\mathrm{E}\left[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i\right]$
   $\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i\right]$
7. (*Centroid condition*) compute the new precoders $\{\boldsymbol{P}_i\}_{i=1}^{M}$
8. compute the precoder MSE metric for the new codebook (precoders) $\{\boldsymbol{P}_i\}_{i=1}^{M}$ and the new partition regions $\{\mathcal{R}_i\}_{i=1}^{M}$
9. $m \leftarrow m + 1$

end while

Table 7.1: Codebook Optimization.

7.3.5 Bit Allocation

When using scalar quantization (transform coding, [108]) instead of vector quantization, the available bits have to be allocated to the different scalar coefficients. Contrary to the case of CSI MSE based feedback as in Eq. (6.76), the distortion function obtained for the case that the precoders are included in the optimization given by

$$\begin{aligned}\text{MSE} = \sum_{i=1}^{M} p_i \Big(& K - 2\Re\left(\operatorname{tr}\left(\boldsymbol{M}_i g_i \boldsymbol{P}_i\right)\right) + g_i^2 \operatorname{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) \\ & + g_i^2 \operatorname{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\right)\Big) \end{aligned} \tag{7.33}$$

has a very complicated structure since all the parameters are mixed together and it is impossible to separate the influence relative to each user and each scalar quantizer, thus making an efficient optimum bit allocation very difficult to find. We can therefore decide the optimum bit allocation by trying out all the possible bit allocation combinations and taking as a result the best one in terms of minimizing the MSE in Eq. (7.33).

The bit allocation optimization is expressed as

$$\boldsymbol{B} = \underset{\boldsymbol{B}}{\operatorname{argmin}}\, \mathrm{MSE}\,(\boldsymbol{B}) \quad \text{s.t.:} \quad \boldsymbol{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K] \in \mathbb{B}^{d\times K}, \boldsymbol{b}_k = [b_{k,1}, \ldots, b_{k,d}]^{\mathrm{T}}$$
$$\text{with} \quad \mathbb{B} = 0, 2, 4, ... \quad \text{and} \quad \sum_{n=1}^{d} b_{k,n} = N_{\text{bit}} \tag{7.34}$$

where $\boldsymbol{B}$ is the matrix that determines the bit allocation corresponding to the coefficients of each user, N_{bit} is the number of bits available for each user, and $d \leq N$ is the number of coefficients under consideration that directly implies the rank reduction that was introduced in the previous chapter by means of the rank reduction basis $\boldsymbol{V}_k^{\mathrm{H}}$. Notice that only an even number of bits is used to quantize each coefficient, since both real and imaginary parts of each coefficient make use of the same number of bits. Initially, we use the scalar quantizers (codebook entries and partition cells) obtained from the CSI metric as explained in Subsection 6.3.3 for a unit-variance input.

When the number of bits is low, there are no serious problems arising from the computational complexity, but the search for optimum bit allocation becomes infeasible as the number of bits increases. Therefore, we propose a heuristic solution to the problem by reducing the number of combinations to be tested on the MSE. It seems that a uniform distribution over all the coefficients without implementing rank reduction is the most likely allocation in the sense of minimizing the MSE. Thus, a first trial consists of distributing the bits over all the coefficients as uniformly as possible. On the other hand, it is obvious that the coefficients with more energy, i.e. those whose eigenvalues are larger, have more impact on the final MSE performance and, therefore, we must tend to allocate more bits to the first coefficients in order to minimize the MSE. Bearing this fact in mind, successive combinations will move the bits from the initial bit allocation to the coefficients with larger eigenvalues. Therefore, the MSE of Eq. (7.33) is sequentially computed by following this ordering for bit allocation so the process is stopped when, given a certain bit allocation, the MSE is greater than the previous one in the list. This will be termed *optimum bit allocation*.

To illustrate this idea, suppose that we have to distribute 8 bits for each user (see Table 7.2). According to the heuristic bit allocation shown above, the chain of possible bit allocations is given by $[2,2,2,2]^{\mathrm{T}} \rightarrow [4,2,2,0]^{\mathrm{T}} \rightarrow [4,4,0,0]^{\mathrm{T}} \rightarrow [6,2,0,0]^{\mathrm{T}} \rightarrow [8,0,0,0]^{\mathrm{T}}$. Imagine the combination given by $[4,2,2,0]^{\mathrm{T}}$ gives us less MSE than $[2,2,2,2]^{\mathrm{T}}$. In that case, we have to test the result when $[4,4,0,0]^{\mathrm{T}}$ is considered. As

Bits per user	No bit allocation	Rank reduction	Bit allocation
$N_{\text{bit}} = 6$ 3 for real part 3 for imaginary part	$[2, 2, 2, 0]^{\text{T}}$	$[4, 2, 0, 0]^{\text{T}}$	Select the optimal from: $[2, 2, 2, 0]^{\text{T}}$, $[4, 2, 0, 0]^{\text{T}}$ $[6, 0, 0, 0]^{\text{T}}$
$N_{\text{bit}} = 8$ 4 for real part 4 for imaginary part	$[2, 2, 2, 2]^{\text{T}}$	$[4, 4, 0, 0]^{\text{T}}$	Select the optimal from: $[2, 2, 2, 2]^{\text{T}}$, $[4, 2, 2, 0]^{\text{T}}$ $[4, 4, 0, 0]^{\text{T}}$, $[6, 2, 0, 0]^{\text{T}}$ $[8, 0, 0, 0]^{\text{T}}$
$N_{\text{bit}} = 10$ 5 for real part 5 for imaginary part	$[4, 2, 2, 2]^{\text{T}}$	$[4, 4, 2, 0]^{\text{T}}$	Select the optimal from: $[4, 2, 2, 2]^{\text{T}}$, $[4, 4, 2, 0]^{\text{T}}$ $[6, 4, 0, 0]^{\text{T}}$, $[8, 2, 0, 0]^{\text{T}}$ $[10, 0, 0, 0]^{\text{T}}$

Table 7.2: Number of Bits Assigned per User's Coefficient for Precoding MSE Metric.

long as the new MSE obtained is less than the previous one, we have to continue with the search until the last possibility embodied by $[8, 0, 0, 0]^{\text{T}}$. If not, we choose $[4, 2, 2, 0]^{\text{T}}$ as the optimum bit allocation for our joint approach based on precoder MSE metric. This heuristic solution significantly reduces the computational complexity of the search with negligible loss in performance.

7.4 Simulation Results

Given the enormous computational complexity due to the calculation of the distortions in Subsection 7.3.3, we consider a system with a transmitter equipped with $N = 4$ antennas that serves $K = 2$ users using QPSK modulation. We use the urban micro *Spatial Channel Model* (SCM) described in Chapter 2, which of the three spatial channel models introduced in that chapter is the most difficult for precoding, because the second and the third channel eigenvalues have a non-negligible magnitude. The results for the CSI metric are the mean of 100 channel realizations with $1{,}000$ symbols being transmitted per channel realization. The number of averaged channel settings or channel covariance matrices is 100. Due to the high complexity, these quantities are reduced for the approach shown in this chapter in the sense that only 10 channel settings are averaged. The training sequence has $N_{\text{tr}} = 16$ symbols. In the figures, the number of bits per user is given. We use the MMSE weights shown in Chapter 5 at the receiver instead of common weights arising from the optimization, which allows us to get better performance with

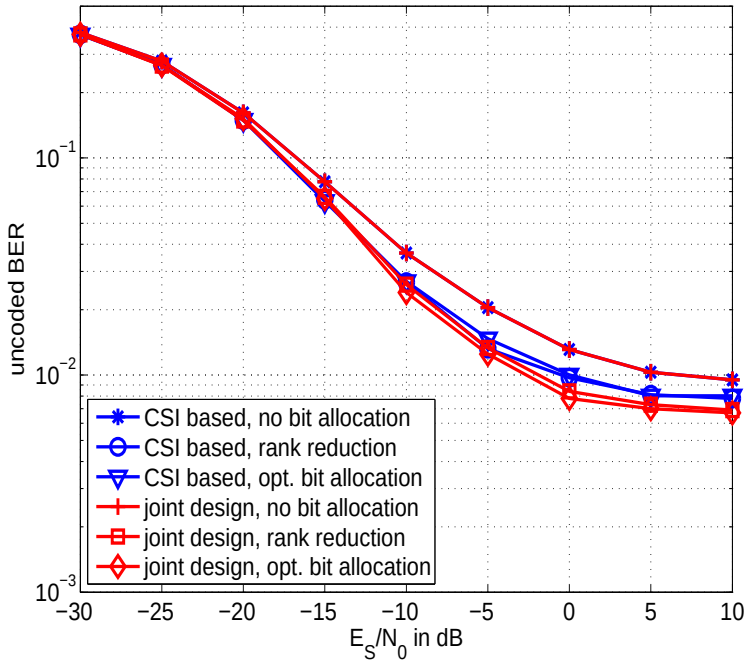


Figure 7.6: MU–MISO System with Robust Linear Precoding, $N = 4$ Antennas, $K = 2$ Users, and 8 Bits per User.

appropriate correction of the phase and amplitude mismatch caused by imperfect CSI at the transmitter.

In the simulations, we use $\boldsymbol{V}_k$ to reduce the rank of the estimated signal in order to decrease the number of possible combinations at the input of the quantizers, for reasons of complexity. We also implement three different types of bit allocation. First, *no bit allocation*, which tries to spread the bits as uniformly as possible (in the event that any bits are left over, e.g. with 10 bits for 4 dimensions, the dimension corresponding to the largest $\phi_{k,i}$ gets an additional bit). Second, *rank reduction*, which allocates as evenly as possible the bits to the first d dimensions. And third, the *optimum bit allocation*, which tries out different bit allocations and takes the result of the best one. Remember that we do not try all the possible combinations and the heuristic search explained in Subsection 7.3.5 is performed instead. To illustrate the different strategies, Table 7.2 summarizes the bit allocation strategies for different number of bits per user.

In Fig. 7.6, the feedback design based on CSI discussed in Chapter 6 is compared to the scheme proposed in this chapter for 8 bits fed back per user. As expected, bit allocation has a considerable impact on the BER performance and the feedback design based on the precoder MSE outperforms the CSI MSE feedback.

As demonstrated in Fig. 7.7, we obtained similar results for a higher and lower number

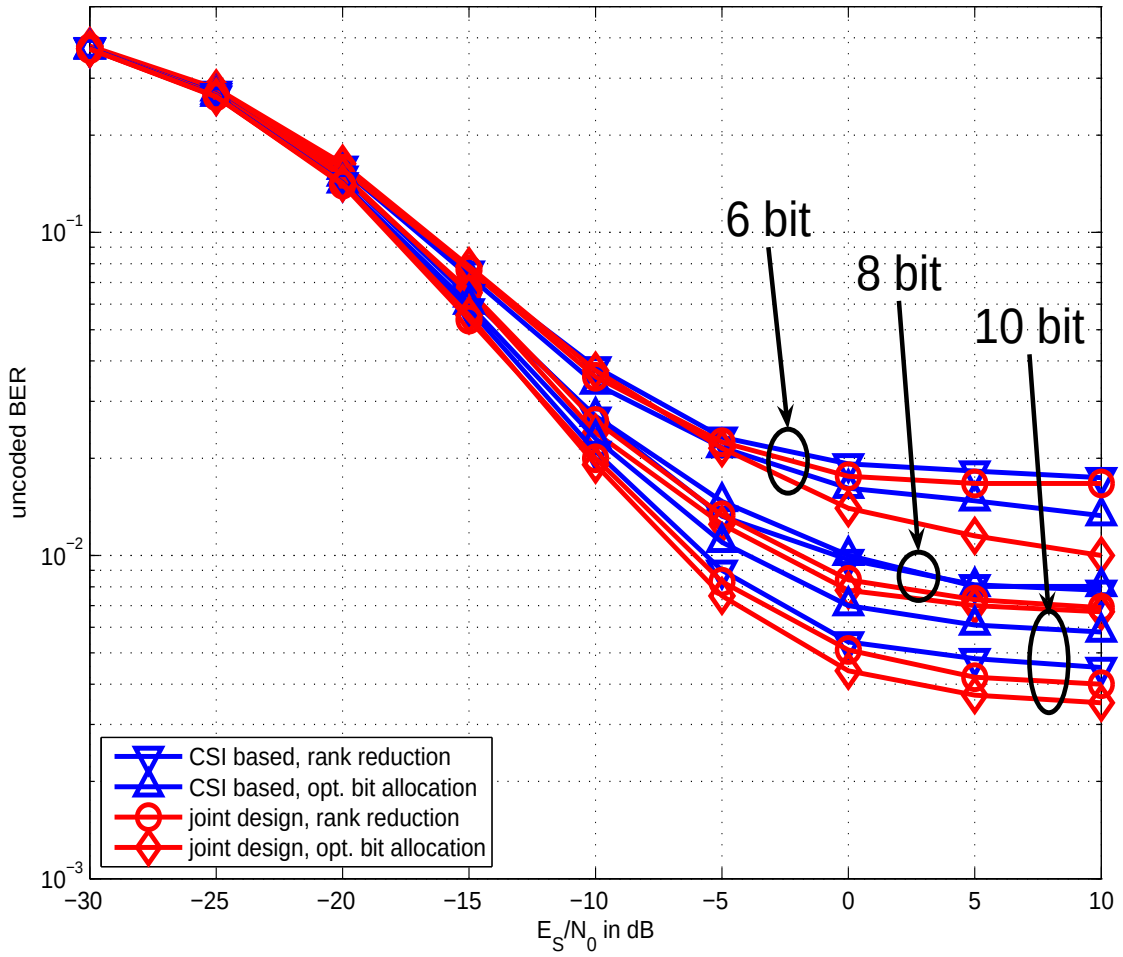


Figure 7.7: MU-MISO System with Robust Linear Precoding, $N = 4$ Antennas, $K = 2$ Users with Different Number of Bits per User.

of bits per user. Not surprisingly, a higher number of bits per user improves the BER performance of all schemes. Additionally, it seems that the advantage of the precoder MSE based design compared to the CSI MSE based design becomes more pronounced for a higher number of bits as the degrees of freedom increase.

Note that, independently from the number of bits fed back per user, rank reduction always shows a loss in performance with respect to optimum bit allocation since the information contained on some coefficients is dropped.

7.5 Conclusions

In this chapter, we have shown how we obtain the robust precoder parameters, the estimator, and the quantizer parameters in a joint optimization by means of a metric not based on CSI-MSE, i.e. minimizing the MSE between the transmitted symbols and the estimated symbols.

Interestingly, the estimators and precoders obtained with the metric oriented to the precoder are equal to the estimators and precoders resulting from the joint optimization based on a CSI-metric presented in Chapter 6. However, the crucial part of the scheme proposed in this chapter is the design of the partition cells of each user, which are designed by minimizing its own distortion but averaging over the quantizer inputs for the other users, since there is no cooperation between users in the downlink of a multiuser MISO system.

As a result, we get better BER performance with a negligible increase of the overhead of the feedback channel. This negligible overhead is due to the fact that each user does not feed back only one single index, which, for the CSI-metric case, was the quantized version of the reduced rank channel estimate. Instead of that, the precoding scheme developed in this chapter is based on the feedback of several indices from each user so that one additional task of the transmitter is performing the intersection of the indices received from all the users to find out the optimal entry of the quantizer that leads to the optimal precoder to be used during the transmission. It is important to note that the codebook entries are now the precoders rather than the white channel coefficients. Therefore, it is obvious that the design of the quantizer parameters (i.e. the codebook entries and the partition cells) becomes the hardest part of this new precoding approach, with the advantage of minimizing the MSE by including the precoder in the optimization. This improvement is even more significant when the number of fed–back bits per user is increased, albeit at the cost of much higher computational complexity. For this reason, we have to think about an efficient computation approach that reduces this complexity in order to make good use of the optimum performance achieved with the final proposed scheme. As also demonstrated with some computer simulations, the simple idea of an

optimum bit allocation even improves the final performance, regardless of whether the CSI-metric or the precoding-metric are used.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

The main objective of this work is to obtain a limited feedback design to be used together with MMSE robust precoder designs when no full channel knowledge is available at the transmitter. We have shown that the proposed designs clearly outperform non-robust designs, especially in the high SNR regime. At the same time, the proposed limited feedback ensures that the feedback rate is limited and therefore easy to implement in real environments.

We started by introducing the signal model and the channel characteristics for the downlink of a MU-MISO wireless communications system. We focused on one of the most widely used channel models, the *Spatial Channel Model* (SCM), which was used throughout this work instead of uncorrelated channels due to its greater similarity to real channels.

We examined and compared the different types of transmit and receive processing for MU-MISO and MU–SIMO systems, respectively, assuming that the transmitter and the receiver side have full knowledge about the channel and its second-order statistics. First, we focused on linear processing and both receive and transmit processing were compared via computer simulations. While the matched-filters outperformed the zero-forcing filters for low SNR, their behavior was clearly the worst for high SNR. The Wiener filters, however, were always superior compared to the other two filter types. We were able to observe a difference between linear receive and transmit processing due to the noise coloring at the receiver. Thus, there is a small advantage for the receive filters at low SNR and for the transmit filters at high SNR. Then, we compared receive and transmit nonlinear processing, focusing on some relevant schemes. Clearly, the nonlinear schemes outperformed the respective linear ones. The optimum nonlinear technique is *maximum likelihood* detection, which has an exponential complexity and, for this reason,

is prohibitive in many cases. Vector precoding takes the procedure of maximum likelihood detection to find the perturbation signal. As a result, a small loss in performance of vector precoding for low SNR but certain gains for high SNR scenarios can be observed compared to maximum likelihood detection, due to the modulo operator at the receiver side. Decision feedback equalizer (DFE) and Tomlinson-Harashima precoding can be seen as suboptimum nonlinear techniques of maximum likelihood detection and vector precoding, respectively. Therefore, their performance was clearly inferior to that of the other two schemes. Again THP outperformed DFE for high SNR but not for low SNR, due to the effect of the modulo operators at the transmitter and receiver. We obtained as a conclusion that transmit processing behaves in a similar way as receive processing when we base the comparison on the BER performance and, therefore, it is quite interesting to compensate the channel effects in advance at the transmitter, in order to exploit the degrees of freedom at the transmitter and to simplify the requirements of the receivers.

We introduced the error sources concerned with the CSI available at the transmitter to be considered throughout this work. Each user estimates the channel and reduces it to a low-dimensional representation for data compression. Compression is feasible thanks to the channel correlations of SCM and every realistic channel. Before the feedback, the CSI is quantized and only the index of the codebook entry is sent to the transmitter, since the data rate of the feedback channel is limited. We also considered that the feedback channel introduces some delay during the transmission.

After that, we extended the study to the situation where no full channel knowledge is available at the transmitter. We derived the robust Wiener linear and nonlinear transmit processing in order to compensate the mismatch between the true channel and the erroneous channel resulting from estimation, truncation, quantization, and feedback delay. By applying the proposed robust designs, we greatly improved the BER performance, avoiding the BER increasing effect observed for high SNR in non-robust schemes.

Then, we investigated the limited feedback design to be used together with these robust precoders in order to optimize the MSE between the true channel and the erroneous channel recovered by the transmitter. First, we developed a very simple feedback design where no MSE optimization is considered. Here, we considered LS estimation, truncation, uniform quantization, and feedback delay. The errors were modeled separately and then the resulting error covariance matrix was directly introduced into our robust designs. Next, we proposed a joint MSE optimization of the channel estimation and the rank reduction basis, where the quantizer was modeled as a data independent additive noise source. This approach, however, was improved when the quantizer was included in the MSE optimization and the Lloyd algorithm was used to construct the codebook and the partition regions. We also proposed a bit allocation algorithm to optimize the bits assigned to each coefficient in real time, enabling the performance to be increased further.

Finally, we devoted the last chapter to the joint design of the channel estimators

and the quantizers at the receivers together with the precoder at the transmitter, based on a precoder-centric criterion, i.e. to minimize an MSE metric appropriate for the precoder design. To this end, we optimized the MSE between the transmitted symbols and the recovered data for each user. This is in contrast to our previous work, where the quantizer design was based on an MSE metric based on the channel information. The proposed system based on a joint optimization clearly outperformed previous designs that separately optimized feedback and precoding at the cost of increasing the computational complexity at the transmitter. But this is not so important when we consider the downlink of wireless communications systems.

8.2 Future Work

Precoding is a technique of growing importance, e.g. it is being incorporated into recent wireless standards, such as the *Worldwide Interoperability for Microwave Access* (WiMAX) or *3GPP Long-Term Evolution* (LTE) standards. WiMAX is the name commonly used for the telecommunications technology that provides wireless transmission of data based on the IEEE 802.16 family of standards. IEEE 802.16 is an IEEE standard for wireless *Metropolitan Area Networks* (MAN). In WiMAX, a feedback loop is included based on a codebook where the mobile stations indicate to the base station the optimum precoding matrix to be used based on the entries of a predefined codebook. The information to be sent from the users to the transmitter is a quantized version of the channel so the transmitter uses this quantized MIMO channel to calculate an optimum precoding matrix. For channel sounding, the MS obtains the CSI by using a dedicated and predetermined signal. On the other hand, LTE, which is intended to be a mobile-communication system in the 2020s, uses a unitary precoding matrix selected from a predefined codebook which is known at both the transmitter and the receiver side. The mobile station estimates the radio channel and selects the optimum precoding matrix that offers maximum capacity. But neither WiMAX nor LTE standards optimize the feedback in any sense, and, obviously, the studied robust Wiener precoding has not yet been incorporated. Therefore, further research in this direction is vindicated and we propose some topics to be developed in the future.

8.2.1 Design of Capacity Approaching Codes for Precoded MU-MISO Systems

A noisy channel poses a limit on the rate at which information can be transferred through it without errors. This limit is known as channel capacity, and it was first introduced by Shannon in 1948 [122]. For many channels, their noisiness can be measured by a single

parameter (for example, the relation between the strength of the transmitted signal and the noise power, termed *Signal-to-Noise Ratio* (SNR)), and the value of this parameter uniquely determines the maximum data transfer rate that can be achieved under the constraint of error-free transmission. Correspondingly, there is also a lower bound for the channel parameter (e.g. SNR) to achieve a given transmission rate (without errors).

Since Shannon demonstrated the existence of this limit, much of the effort in the field of digital communications has been devoted to the search for practical channel codes capable of approaching it. Shannon's theorem only proved that infinitely long random codewords could achieve the limit, but using that approach to design a real code was believed to be impossible. Members of the University of A Coruña staff proposed a technique for obtaining good capacity approaching codes using *Extrinsic Information Transfer* (EXIT) functions in a novel way. Up to now, no such codes have been optimized for precoding schemes. The combination of optimized codes with precoding will result in a scheme that is able to achieve a higher performance than schemes without coding.

8.2.2 MU-MIMO Systems

We propose to extend our MU-MISO environment to a MU-MIMO one, where N transmit antennas and M receive antennas per user are considered. The results will be improved since diversity is increased with a growing number of receive antennas at the cost of complicating the design of the limited feedback channel.

8.2.3 Wideband Frequency Selective Channels

In frequency-selective fading, the coherence bandwidth of the channel is smaller than the bandwidth of the signal, which leads to *wideband* systems instead of the *narrowband* systems studied in this work. Contrary to wideband *single–carrier* systems based on *Code Division Multiple Access* (CDMA), the available frequency band can be divided into a number of subbands, each having a bandwidth lower than the coherence bandwidth of the channel, so that signals transmitted in each subband experience flat fading. This leads to wideband *multi–carrier* systems such as *Orthogonal Frequency Division Multiplexing* (OFDM). OFDM has been adopted as the downlink transmission scheme for LTE and is also used for several other radio technologies such as WiMAX. Therefore, another open issue to be studied further, even for the limited feedback design proposed in this work, is the case of OFDM precoding systems. Moreover, OFDM is quite robust against multipath, frequency-selectivity, and *Radio Frequency* (RF) interferences. While narrowband analysis can be relatively easy to develop, it is not so clear how to avoid the overhead of the limited feedback channel when we have to transmit N_{B} bits for each of the N_{M} OFDM tones. To reduce the amount of information to be sent from the receivers we

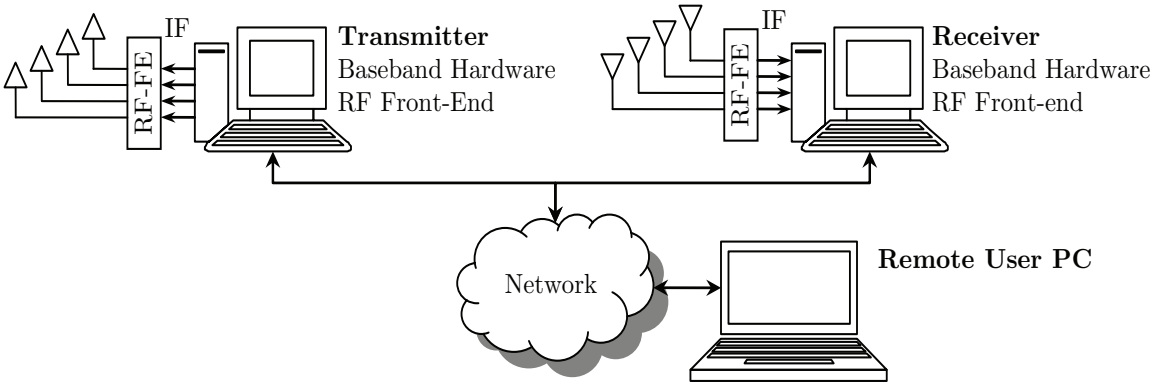


Figure 8.1: Schematic Diagram of the 4×4 MIMO Textbed.

could also think about not feeding back information related to some tones, which might be recovered at the transmitter by applying interpolation techniques.

8.2.4 Improvement of Limited Feedback Design based on the Precoding MSE Metric

Up to now, the limited feedback design involving the precoding into the MSE optimization that is described in Chapter 7 shows a good performance at the cost of increasing computational complexity and, therefore, the processing time at the signal processors. We have to work in the sense of reducing these time requirements by means of optimizing the code or moving from MATLAB to C programming in order to extend the results obtained to a greater number of transmit and receive antennas, also with larger codebooks for the quantization process.

On the other hand, in this work we have only derived the optimum limited feedback design oriented to the precoding MSE according to the linear Wiener precoding approach. It is known that the performance achieved by the nonlinear schemes, such as THP or VP, is superior to that of the linear schemes, and for this reason we could apply the ideas shown for the linear case to the other two types of filter studied, even though the derivation may be quite complex.

8.2.5 Precoding on Testbeds

We can focus on the evaluation of several of the above limited feedback schemes over realistic indoor scenarios. To this end, we could make use of a MIMO testbed which would give us an idea of the real performance of these schemes over real-world channels. In recent years, a MIMO testbed has been developed by the University of A Coruña. The

schematic diagram of this testbed is depicted in Fig. 8.1 and a picture is shown in Fig. 8.2. It is basically composed of two PCs, one for the transmitter and one for the receiver. Each PC contains the baseband hardware plus the *Radio Frequency Front-End* (RF-FE). The baseband hardware is composed of fast memories that can be accessed at the speed of the A/D and D/A converters, thus allowing the transmission and subsequent acquisition of signals in real-time while the signal processing at both transmitter and receiver sides is performed off-line. The testbed uses modern RF-FE equipment allowing wideband transmissions (up to 40 MHz) at both 2.4 and 5 GHz bands.

It is important to note that the main difference between testbeds and other types of hardware implementation, such as prototypes or demonstrators, is that only the transmissions take place in real-time, while the rest of the processing operations are carried out off-line. This could be seen as an inconvenience, since the time required for such operations is larger than in the case of real-time implementations. Moreover, this issue is especially critical in the case of precoding because the time consumed in the calculations and feedback has to be taken into account and compared to the channel coherence time. However, the off-line implementation presents major advantages such as floating point precision, high flexibility, and minimum effort needed to translate the algorithms from the simulations to the testbed. Therefore, the task of obtaining CSI at the receiver to be sent using a control link (e.g. a socket network connection) to the transmitter could be optimally performed by dedicated and powerful resources in the receiver in order to reduce the time consumption as far as possible. Then, the transmitter generates the signals to be transmitted according to that feedback information and finally, the precoded signals are sent by the transmitter hardware, acquired at the receiver side and buffered for later evaluation. With this approach only the feedback calculation at the receiver and the subsequent signal precoding at the transmitter take place in *quasi* real-time, while the rest of the operations are kept off-line.

Some preliminary trials were performed using only baseband signals while the channel was emulated by software. In that case all operations took place off-line and the evaluation of the resulting data was simplified because the channel coherence time was under control. Therefore, a major effort is still required to achieve the final objective of implementing a precoding system with limited feedback which ensures that the overall time consumption, including the calculations related to obtaining the CSI at the receiver and to building the optimum precoder at the transmitter, does not exceed the coherence time of the channel.

Figure 8.2: A Picture of the 4×4 MIMO Testbed.

8.2.6 Design of Limited Feedback based on maximizing Mutual Information

Intuitively, mutual information measures the information that X and Y share, i.e. it measures how much knowing one of these variables reduces our uncertainty about the other. For example, if X and Y are independent, then knowing X does not give any information about Y and vice versa, so their mutual information is zero. At the other extreme, if X and Y are identical then all information conveyed by X is shared with Y and, therefore, knowing X determines the value of Y and vice versa. As is known, Shannon proved that the channel capacity equals the mutual information of the channel maximized over all the possible input distributions [122], i.e.

$$C = \max_{p(x)} I\left(X;Y\right) = \max_{p(x)} \sum_{x,y} p\left(x,y\right) \log\left(\frac{p\left(x,y\right)}{p\left(x\right)p\left(y\right)}\right).$$

Therefore, we could think about optimizing the precoders for the proposed limited feedback in order to maximize the input-output mutual information and thus come closer to the channel capacity.

8.2.7 Feedback of Long-Term Channel Variations

We assume throughout this work that the second order channel statistics are known at both the receiver and the transmitter side. But this situation is not realistic. We can estimate the channel covariance matrix using supervised methods, although it is a matter of discussion how often the pilot symbols have to be transmitted or how we can detect changes in these long-term channel variations and, even more importantly, how we can feed back this information through the reverse channel without significantly increasing the amount of information sent from all the users to the transmitter. The impact of errors on second order channel statistics may strongly degrade the overall system performance and, again, it could be interesting to derive a robust precoding system against this new type of errors.

Appendix A

Spatial Channel Model (SCM)

The proposed limited feedback system is based on channel correlations and it is therefore important to obtain models for correlated channels that are as realistic as possible. In this sense, we use the 3GPP Spatial Channel Model (SCM). The 3GPP channel model includes a fixed number of $M = 6$ paths in each environment so the received signal at the mobile station consists of M-delayed multipath replicas of the transmitted signal. These M paths are defined by powers and delays randomly obtained as explained below. Each path consists of $S = 20$ subpaths with the goal of including the fading effect. All paths and subpaths are assumed to be statistically independent. Fig. A.1 depicts the most important angular parameters used to describe each SCM environment (suburban macrocell, urban macrocell or urban microcell), whose meaning is summarized in Table A.1. The parameters related to each scenario are summarized in Tables A.3, A.4, and A.5, respectively. Table A.2 includes the subpath AoD and AoA offsets for macrocell and microcell setups.

For each path m ($m = 1, 2, \ldots, 6$), the channel attenuation corresponding to the

θ_{BS}	LOS AoD direction between the BS and MS with respect to the BS antenna array orientation
$\delta_{m,\text{AoD}}$	AoD for the m–th path with respect to the LOS AoD
$\Delta_{m,s,\text{AoD}}$	Offset for the s–th subpath of the m–th path with respect to $\delta_{m,\text{AoD}}$
$\theta_{m,s,\text{AoD}}$	AoD for the s–th subpath with respect to the BS antenna array orientation
θ_{MS}	Angle between the BS–MS LOS and the MS antenna array orientation
$\delta_{m,\text{AoA}}$	AoA for the m–th path with respect to the LOS AoA
$\Delta_{m,s,\text{AoA}}$	Offset for the s–th subpath of the m–th path with respect to $\delta_{m,\text{AoA}}$
$\theta_{m,s,\text{AoA}}$	AoA for the s–th subpath of the m–th path at the MS with respect to the MS antenna array orientation

Table A.1: Angular Parameters for SCM.

Subpath s	**Offset for a** $2°$**AS at BS (Macrocell)** $\Delta_{m,s,\text{AoD}}$ (°)	**Offset for a** $5°$**AS at BS (Microcell)** $\Delta_{m,s,\text{AoD}}$ (°)	**Offset for a** $35°$**AS at MS** $\Delta_{m,s,\text{AoA}}$ (°)
$1, 2$	± 0.0894	± 0.2236	± 1.5649
$3, 4$	± 0.2826	± 0.7064	± 4.9447
$5, 6$	± 0.4984	± 0.2461	± 8.7224
$7, 8$	± 0.7431	± 1.8578	± 13.0045
$9, 10$	± 1.0257	± 2.5642	± 17.9492
$11, 12$	± 1.3594	± 3.3986	± 23.7899
$13, 14$	± 1.7688	± 4.4220	± 30.9538
$15, 16$	± 2.2961	± 5.7403	± 40.1824
$17, 18$	± 3.0389	± 7.5974	± 53.1816
$19, 20$	± 4.3101	± 10.7753	± 75.4274

Table A.2: Subpath AoD and AoA Offsets. Last Column Corresponds to Multiple Antennas at the Receiver Side.

transmit antenna t and the receive antenna r is given by

$$c_{r,t,m}(t) = A \sum_{s=1}^{S} \sqrt{G_{\text{BS}}(\theta_{m,s,\text{AoD}})\, \mathrm{e}^{\mathrm{j}(Kd_t \sin(\theta_{m,s,\text{AoD}}) + \phi_{m,s})}} \sqrt{G_{\text{MS}}(\theta_{m,s,\text{AoA}})\, \mathrm{e}^{\mathrm{j}(Kd_r \sin(\theta_{m,s,\text{AoA}}))}} \tag{A.1}$$

with $A = \sqrt{\frac{P_m \sigma_{\text{SF}}}{S}}$, where P_m is the power of the m–th path, σ_{SF} is the lognormal fading deviation, $S = 20$ is the number of subpaths per path, and $G_{\text{BS}}(\theta_{m,s,\text{AoD}})$ and $G_{\text{MS}}(\theta_{m,s,\text{AoA}})$ are the antenna gains for each BS and MS antenna depending on the angle-of-departure $\theta_{m,s,\text{AoD}}$ and the angle-of-arrival $\theta_{m,s,\text{AoA}}$, respectively. The phase of the s–th subpath of the m–th path is given by $\phi_{m,s}$, a random variable uniformly distributed over $[0, 2\pi]$. The distance d_t is the distance in meters from the BS antenna elements to the reference ($t = 1$) antenna. Note that $d_1 = 0$ for the reference antenna. The distance in meters from the MS antenna elements to the reference ($r = 1$) antenna is denoted by d_r. Again, $d_1 = 0$ for the reference antenna. The wavelength in meters is given by λ, which leads to the wave number K defined as $2\pi/\lambda$. For the MU-MISO scenario considered in this work, we have $d_r = 0$, with $N_{\text{r}} = 1$ and, therefore, the last exponential term vanishes.

Each coefficient of the N-dimensional MISO channel vector corresponding to the user k is obtained according to the expression in Eq. (A.1), which leads to the channel vector corresponding to the m-th path given by $\boldsymbol{c}_{k,m}(t) = [c_{1,1,m}(t), \ldots, c_{1,N,m}(t)]^{\text{T}}$ and, since the channel vector is the sum of the signal received through all the paths (remember that the number of paths is 6), we have the narrowband channel impulse response [cf.

Eq. (2.3)]

$$\boldsymbol{h}_{\mathrm{SCM},k}\left(t\right)=\sum_{m=1}^{6}\boldsymbol{c}_{k,m}\left(t\right)\delta\left(t-\tau_{m}(t)\right).$$

The channel covariance matrix that models the channel spatial correlations is then obtained as follows

$$\boldsymbol{C}_{\boldsymbol{h},k}=\mathrm{E}\left[\boldsymbol{h}_{\mathrm{SCM},k}\left(t\right)\boldsymbol{h}_{\mathrm{SCM},k}^{\mathrm{H}}\left(t\right)\right] \tag{A.2}$$

where we assume that the channel is stationary, and therefore that $\boldsymbol{C}_{\boldsymbol{h},k}$ is constant. We also assume that $\boldsymbol{C}_{\boldsymbol{h},k}$ is known a priori.

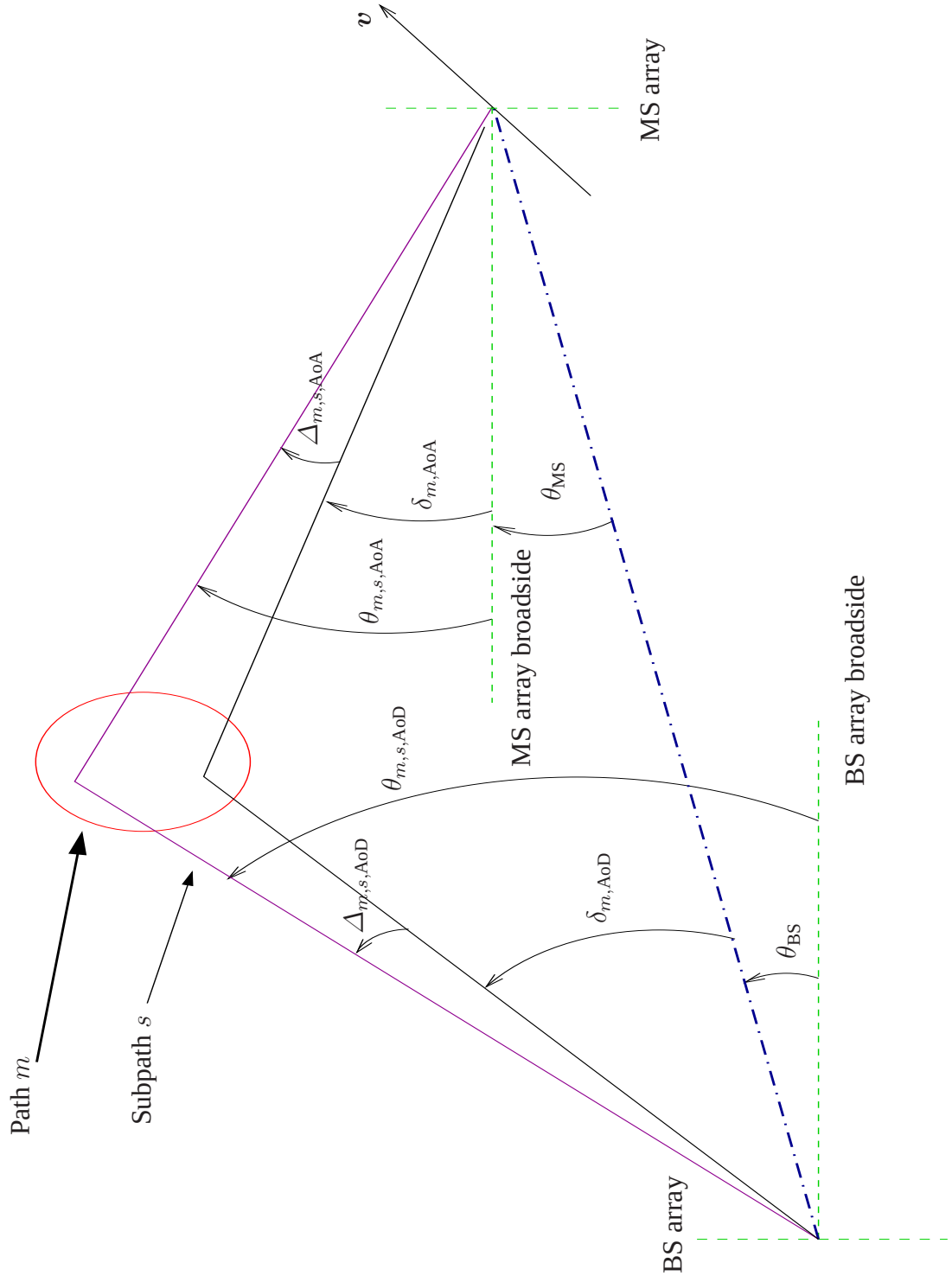


Figure A.1: BS and MS Angle Parameters.

Channel Scenario	**Suburban Macro**
Paths	M=6
Subpaths per path	S=20
Angular spread at BS $\sigma_{\text{AS}} = 10^{(\epsilon_{\text{AS}}x+\mu_{\text{AS}})}$ $x \sim \mathcal{N}(0,1)$	$\text{E}[\sigma_{\text{AS}}] = 5°$ $\epsilon_{\text{AS}} = 0.13$ $\mu_{\text{AS}} = 0.69$
$r_{\text{AS}} = \sigma_{\text{AoD}}/\sigma_{\text{AS}}$	1.2
Per-path AS at BS	2°
BS per-path AoD $\delta_{m,\text{AoD}}$	$\delta_m^{'} = \mathcal{N}(0, \sigma^2_{\text{AoD}})$ $\lvert\delta^{'}_{(1)}\rvert \leq \ldots \leq \lvert\delta^{'}_{(M)}\rvert$ $\delta_{m,\text{AoD}} = \delta^{'}_{(m)}$ $\sigma_{\text{AoD}} = r_{\text{AS}}\sigma_{\text{AS}}$
Offset of the s–th subpath $\Delta_{m,s,\text{AoD}}$	Fixed (Table A.2)
AoD of the s–th subpath	$\theta_{m,s,\text{AoD}} = \theta_{\text{BS}} + \delta_{m,\text{AoD}} + \Delta_{m,s,\text{AoD}}$
Angular spread (AS) per path at MS	35°
Delay spread $\sigma_{\text{DS}} = 10^{(\epsilon_{\text{DS}}x+\mu_{\text{DS}})}$ $x \sim \mathcal{N}(0,1)$	$\text{E}[\sigma_{\text{DS}}] = 0.17\mu s$ $\mu_{\text{DS}} = -6.8$ $\epsilon_{\text{DS}} = 0.288$
r_{DS}	1.4
Distribution for path delays	$\tau_m^{'} = -r_{\text{DS}}\sigma_{\text{DS}}\ln z_m$ $z_m \sim \mathcal{U}(0,1)$
Power of the m– path	$P_m^{'} = \text{e}^{(1-r_{\text{DS}})(\tau^{'}_{(m)}-\tau^{'}_{(1)})/r_{\text{DS}}\sigma_{\text{DS}}} \cdot 10^{-\xi_m/10}$ $\xi_m \sim \mathcal{N}(0, \sigma_{\text{RND}})$, σ_{RND} =3 dB $P_m = \frac{P_m}{\sum_{j=1}^{6} P_j}$
AoA for the m–th path	$\delta_{m,\text{AoA}} \sim \mathcal{N}(0, \sigma^2_{m,\text{AoA}})$ $\sigma_{m,\text{AoA}} = 104.12(1 - \exp(-0.2175\lvert P_m(\text{dB})\rvert))$
Offset of the s–th subpath $\Delta_{m,s,\text{AoA}}$	Fixed (Table A.2)
AoA for the s–th subpath	$\theta_{m,s,\text{AoA}} = \theta_{\text{MS}} + \delta_{m,\text{AoA}} + \Delta_{m,s,\text{AoA}}$
Lognormal shadowing deviation σ_{SF}	8 dB

Table A.3: Environment Parameters. SCM 1: Suburban Macrocell.

Channel Scenario	**Urban Macro**
Paths	M=6
Subpaths per path	S=20
Angular spread at BS $\sigma_{\text{AS}} = 10^{(\epsilon_{\text{AS}} x + \mu_{\text{AS}})}$ $x \sim \mathcal{N}(0,1)$	$\text{E}[\sigma_{\text{AS}}] = 8°, 15°$ $8°$: $\epsilon_{\text{AS}} = 0.34$, $\mu_{\text{AS}} = 0.810$ $15°$: $\epsilon_{\text{AS}} = 0.21$, $\mu_{\text{AS}} = 1.18$
$r_{\text{AS}} = \sigma_{\text{AoD}}/\sigma_{\text{AS}}$	1.3
Per-path AS at BS	2°
BS per-path AoD $\delta_{m,\text{AoD}}$	$\delta_m^{'} = \mathcal{N}(0, \sigma^2_{\text{AoD}})$ $\lvert\delta^{'}_{(1)}\rvert \leq \ldots \leq \lvert\delta^{'}_{(M)}\rvert$ $\delta_{m,\text{AoD}} = \delta^{'}_{(m)}$ $\sigma_{\text{AoD}} = r_{\text{AS}}\sigma_{\text{AS}}$
Offset of the s–th subpath $\Delta_{m,s,\text{AoD}}$	Fixed (Table A.2)
AoD of the s–th subpath	$\theta_{m,s,\text{AoD}} = \theta_{\text{BS}} + \delta_{m,\text{AoD}} + \Delta_{m,s,\text{AoD}}$
Angular spread (AS) per path at MS	35°
Delay spread $\sigma_{\text{DS}} = 10^{(\epsilon_{\text{DS}} x + \mu_{\text{DS}})}$ $x \sim \mathcal{N}(0,1)$	$\text{E}[\sigma_{\text{DS}}] = 0.65\mu s$ $\mu_{\text{DS}} = -6.18$ $\epsilon_{\text{DS}} = 0.18$
r_{DS}	1.7
Distribution for path delays	$\tau_m^{'} = -r_{\text{DS}}\sigma_{\text{DS}}\ln z_m$ $z_m \sim \mathcal{U}(0,1)$
Power of the m– path	$P_m^{'} = \text{e}^{(1-r_{\text{DS}})(\tau^{'}_{(m)} - \tau^{'}_{(1)})/r_{\text{DS}}\sigma_{\text{DS}}} \cdot 10^{-\xi_m/10}$ $\xi_m \sim \mathcal{N}(0, \sigma_{\text{RND}})$, σ_{RND} =3 dB $P_m = \frac{P_m}{\sum_{j=1}^{6} P_j}$
AoA for the m–th path	$\delta_{m,\text{AoA}} \sim \mathcal{N}(0, \sigma^2_{m,\text{AoA}})$ $\sigma_{m,\text{AoA}} = 104.12(1 - \exp(-0.2175\lvert P_m(\text{dB})\rvert))$
Offset of the s–th subpath $\Delta_{m,s,\text{AoA}}$	Fixed (Table A.2)
AoA for the s–th subpath	$\theta_{m,s,\text{AoA}} = \theta_{\text{MS}} + \delta_{m,\text{AoA}} + \Delta_{m,s,\text{AoA}}$
Lognormal shadowing deviation σ_{SF}	8 dB

Table A.4: Environment Parameters. SCM 2: Urban Macrocell.

Channel Scenario	**Urban micro**
Paths	M=6
Subpaths per path	S=20
Angular spread at BS $\sigma_{\text{AS}} = 10^{(\epsilon_{\text{AS}}x+\mu_{\text{AS}})}$ $x \sim \mathcal{N}(0,1)$	N/A
$r_{\text{AS}} = \sigma_{\text{AoD}}/\sigma_{\text{AS}}$	N/A
Per-path AS at BS	2°
BS per-path AoD $\delta_{m,\text{AoD}}$	$\mathcal{U}(-40°, +40°)$
Offset of the s–th subpath $\Delta_{m,s,\text{AoD}}$	Fixed (Table A.2)
AoD of the m–th subpath	$\theta_{m,s,\text{AoD}} = \theta_{\text{BS}} + \delta_{m,\text{AoD}} + \Delta_{m,s,\text{AoD}}$
Angular spread (AS) per path at MS	35°
Delay spread $\sigma_{\text{DS}} = 10^{(\epsilon_{\text{DS}}x+\mu_{\text{DS}})}$ $x \sim \mathcal{N}(0,1)$	$\text{E}[\sigma_{\text{DS}}] = 0.251\mu s$ N/A
r_{DS}	N/A
Distribution for path delays	$\tau_m \sim \mathcal{U}(0, 1.2\mu s)$
Power of the m– path	$P_m^{'} = 10^{-(\tau_m + z_m/10)}$ z_m Gaussian zero-mean with deviation of 3 dB $P_m = \frac{P_m}{\sum_{j=1}^{6} P_j}$
AoA for the m–th path	$\delta_{m,\text{AoA}} \sim \mathcal{N}(0, \sigma^2_{m,\text{AoA}})$ $\sigma_{m,\text{AoA}} = 104.12(1 - \exp(-0.265\lvert P_m(\text{dB})\rvert))$
Offset of the s–th subpath $\Delta_{m,s,\text{AoA}}$	Fixed (Table A.2)
AoA for the s–th subpath	$\theta_{m,s,\text{AoA}} = \theta_{\text{MS}} + \delta_{m,\text{AoA}} + \Delta_{m,s,\text{AoA}}$
Lognormal shadowing deviation σ_{SF}	NLOS: 10 dB LOS: 4 dB

Table A.5: Environment Parameters. SCM 3: Urban Microcell.

Appendix B

Useful Matrix Properties

B.1 Matrix Inversion Lemma

Consider the partitioned square matrix [107]

$$M = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$$

where it is assumed that $\boldsymbol{A}$ and $\boldsymbol{D}$ are square matrices and the inverses $\boldsymbol{A}^{-1}$, $\boldsymbol{D}^{-1}$, and $\boldsymbol{M}^{-1}$ exist. Under these assumptions, the matrix $\boldsymbol{M}$ can be factorized as follows:

$$M = \begin{bmatrix} \boldsymbol{A} & \mathbf{0} \\ \boldsymbol{C} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \mathbf{0} & \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \end{bmatrix}.$$

Notice that $\boldsymbol{M}$ has been factorized as the product of two matrices that have one diagonal block submatrix equal to the identity matrix and one off-diagonal submatrix equal to zero. Exploiting this product, it is straightforward to obtain the inverse of each of these factors as follows

$$\begin{bmatrix} \boldsymbol{A} & \mathbf{0} \\ \boldsymbol{C} & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{A}^{-1} & \mathbf{0} \\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \mathbf{I} \end{bmatrix}$$
$$\begin{bmatrix} \mathbf{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \mathbf{0} & \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I} & -\boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1} \\ \mathbf{0} & \left(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1} \end{bmatrix}.$$

Thus, we get for the inverse of the matrix $\boldsymbol{M}$

$$\begin{aligned}\boldsymbol{M}^{-1} &= \begin{bmatrix}\mathbf{I} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \mathbf{0} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{A} & \mathbf{0} \\ \boldsymbol{C} & \mathbf{I}\end{bmatrix}^{-1} \\ &= \begin{bmatrix}\mathbf{I} & -\boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1} \\ \mathbf{0} & \left(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}^{-1} & \mathbf{0} \\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \mathbf{I}\end{bmatrix} \\ &= \begin{bmatrix}\boldsymbol{A}^{-1}+\boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1} \\ -\left(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} & \left(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\end{bmatrix}.\end{aligned} \tag{B.1}$$

This result is called the *matrix inversion lemma for partitioned matrices*.

We can obtain the alternative form factorizing $\boldsymbol{M}$ as follows

$$\boldsymbol{M} = \begin{bmatrix}\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{B}\boldsymbol{D}^{-1} \\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0} \\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}.$$

Exploiting again the fact that each factor has one diagonal block submatrix equal to the identity matrix and an off-diagonal submatrix that is zero, the inverse of $\boldsymbol{M}$ can be rewritten in an alternative way

$$\boldsymbol{M}^{-1} = \begin{bmatrix}\left(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}\right)^{-1} & -\left(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}\right)^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{C}\left(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}\right)^{-1} & \boldsymbol{D}^{-1}+\boldsymbol{D}^{-1}\boldsymbol{C}\left(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}\right)^{-1}\boldsymbol{B}\boldsymbol{D}^{-1}\end{bmatrix}. \tag{B.2}$$

By comparing the upper left elements of Eqs. (B.1) and (B.2), we obtain the following relationship

$$\left(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}\right)^{-1} = \boldsymbol{A}^{-1}+\boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} \tag{B.3}$$

which is the *matrix inversion lemma* [107].

B.2 Properties of the Trace Operator

The trace of a square matrix $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ is the sum of its diagonal elements:

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{i,i}$$

where $a_{i,i}$ denotes the element corresponding to the i-th diagonal entry of the matrix $\boldsymbol{A}$. Obviously, the trace is invariant to the transposition of the argument:

$$\operatorname{tr}\left(\boldsymbol{A}^{\mathrm{T}}\right) = \sum_{i=1}^{n} a_{i,i} = \operatorname{tr}(\boldsymbol{A}). \tag{B.4}$$

On the other hand, the trace is not invariant to the conjugate of its argument. In that case, the trace holds

$$\mathrm{tr}\left(\boldsymbol{A}^{*}\right)=\sum_{i=1}^{n}a_{i,i}^{*}=\left(\sum_{i=1}^{n}a_{i,i}\right)^{*}=\mathrm{tr}^{*}\left(\boldsymbol{A}\right). \tag{B.5}$$

Combining Eqs. (B.4) and (B.5), we obtain that $\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}})=\mathrm{tr}^{*}(\boldsymbol{A})$.

Directly from the definition of the trace of a square matrix, it can be seen that

$$\mathrm{tr}\left(\boldsymbol{BC}\right)=\sum_{i=1}^{m}\sum_{j=1}^{n}b_{i,j}c_{j,i}=\sum_{j=1}^{n}\sum_{i=1}^{m}c_{j,i}b_{i,j}=\mathrm{tr}\left(\boldsymbol{CB}\right) \tag{B.6}$$

where $\boldsymbol{B}\in\mathbb{C}^{m\times n}$ and $\boldsymbol{C}\in\mathbb{C}^{n\times m}$. From the above result, we can conclude that the trace is equal to the sum of eigenvalues

$$\mathrm{tr}\left(\boldsymbol{A}\right)=\mathrm{tr}\left(\boldsymbol{U\Delta U}^{\mathrm{H}}\right)=\mathrm{tr}\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U\Delta}\right)=\mathrm{tr}\left(\boldsymbol{\Delta}\right)=\sum_{i=1}^{n}\lambda_{i}. \tag{B.7}$$

Here, $\boldsymbol{U}\in\mathbb{C}^{n\times n}$ is the modal matrix of $\boldsymbol{A}$ containing its normalized eigenvectors. It is well known that $\boldsymbol{U}$ is a unitary matrix, i.e., $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}=\mathbf{I}$. The diagonal matrix $\boldsymbol{\Delta}$ comprises the eigenvalues of $\boldsymbol{A}$, i.e. $\boldsymbol{\Delta}=\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$.

B.3 Derivatives of Vector and Matrix Functions

The derivatives with respect to vectors or matrices of scalar functions that involve the trace operator are widely employed throughout this work. This appendix collects the definitions of these derivatives, as well as some results that have been used in previous chapters.

B.3.1 Real Derivatives

Let $f:\mathbb{R}^m\rightarrow\mathbb{R},\boldsymbol{x}\mapsto f(\boldsymbol{x})$ with $\boldsymbol{x}\in\mathbb{R}^m$. The derivative of $f(\boldsymbol{x})$ with respect to $\boldsymbol{x}$ is the column vector

$$\frac{\partial f(\boldsymbol{x})}{\partial\boldsymbol{x}}=\begin{bmatrix}\frac{\partial f(\boldsymbol{x})}{\partial x_1}\\ \vdots\\ \frac{\partial f(\boldsymbol{x})}{\partial x_m}\end{bmatrix}\in\mathbb{R}^m$$

and supposing now that $\boldsymbol{X}\in\mathbb{R}^{m\times n}$ is transformed by the function $g:\mathbb{R}^{m\times n}\rightarrow\mathbb{R},\boldsymbol{X}\mapsto g(\boldsymbol{X})$, the derivative of $g(\boldsymbol{X})$ with respect to $\boldsymbol{X}$ is defined as

$$\frac{\partial g(\boldsymbol{X})}{\partial\boldsymbol{X}}=\begin{vmatrix}\frac{\partial g(\boldsymbol{X})}{\partial x_{1,1}} & \cdots & \frac{\partial g(\boldsymbol{X})}{\partial x_{1,n}}\\ \vdots & \ddots & \vdots\\ \frac{\partial g(\boldsymbol{X})}{\partial x_{m,1}} & \cdots & \frac{\partial g(\boldsymbol{X})}{\partial x_{m,n}}\end{vmatrix}\in\mathbb{R}^{m\times n}$$

where $x_{i,j}$ denotes the i–th element in the j–th column of $\boldsymbol{X}$.

B.3.2 Complex Derivatives

Suppose that $\boldsymbol{z} = [z_1, ..., z_m]^T = [x_1 + jy_1, ..., x_m + jy_m]^T \in \mathbb{C}^m$ and $\boldsymbol{Z} \in \mathbb{C}^{m \times n}$ are transformed by the functions $f : \mathbb{C}^m \to \mathbb{C}, \boldsymbol{z} \mapsto f(\boldsymbol{z})$ and $g : \mathbb{C}^{m \times n} \to \mathbb{C}, \boldsymbol{Z} \mapsto g(\boldsymbol{Z})$, respectively. The derivative of $f(\boldsymbol{z})$ with respect to $\boldsymbol{z}$ is defined as the column vector

$$\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}} = \begin{bmatrix} \frac{\partial f(\boldsymbol{z})}{\partial z_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{z})}{\partial z_m} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{\partial f(\boldsymbol{z})}{\partial x_1} - \mathrm{j}\, \frac{\partial f(\boldsymbol{z})}{\partial y_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{z})}{\partial x_m} - \mathrm{j}\, \frac{\partial f(\boldsymbol{z})}{\partial y_m} \end{bmatrix} \in \mathbb{C}^m.$$

The derivative with respect to the complex conjugate of $\boldsymbol{z}$ is obtained as follows

$$\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}^*} = \begin{bmatrix} \frac{\partial f(\boldsymbol{z})}{\partial z_1^*} \\ \vdots \\ \frac{\partial f(\boldsymbol{z})}{\partial z_m^*} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{\partial f(\boldsymbol{z})}{\partial x_1} + \mathrm{j}\, \frac{\partial f(\boldsymbol{z})}{\partial y_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{z})}{\partial x_m} + \mathrm{j}\, \frac{\partial f(\boldsymbol{z})}{\partial y_m} \end{bmatrix} \in \mathbb{C}^m.$$

On the other hand, the derivative of $g(\boldsymbol{Z})$ with respect to the matrix $\boldsymbol{Z}$ is the $m \times n$ matrix given by

$$\begin{aligned} \frac{\partial g(\boldsymbol{Z})}{\partial \boldsymbol{Z}} &= \begin{bmatrix} \frac{\partial g(\boldsymbol{Z})}{\partial z_{1,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial z_{1,n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g(\boldsymbol{Z})}{\partial z_{m,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial z_{m,n}} \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} \frac{\partial g(\boldsymbol{Z})}{\partial x_{1,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial x_{1,n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g(\boldsymbol{Z})}{\partial x_{m,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial x_{m,n}} \end{bmatrix} - \frac{\mathrm{j}}{2} \begin{bmatrix} \frac{\partial g(\boldsymbol{Z})}{\partial y_{1,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial y_{1,n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g(\boldsymbol{Z})}{\partial y_{m,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial y_{m,n}} \end{bmatrix} \in \mathbb{C}^{m \times n} \end{aligned}$$

and, similarly, with respect to $\boldsymbol{Z}^*$

$$\frac{\partial g(\boldsymbol{Z})}{\partial \boldsymbol{Z}^*} = \frac{1}{2} \begin{bmatrix} \frac{\partial g(\boldsymbol{Z})}{\partial x_{1,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial x_{1,n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g(\boldsymbol{Z})}{\partial x_{m,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial x_{m,n}} \end{bmatrix} + \frac{\mathrm{j}}{2} \begin{bmatrix} \frac{\partial g(\boldsymbol{Z})}{\partial y_{1,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial y_{1,n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g(\boldsymbol{Z})}{\partial y_{m,1}} & \cdots & \frac{\partial g(\boldsymbol{Z})}{\partial y_{m,n}} \end{bmatrix} \in \mathbb{C}^{m \times n}.$$

Examples

Let $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ and $\boldsymbol{w} \in \mathbb{C}^m$. During the mathematical derivations throughout this work, the following relationships are used

- $\frac{\partial \|\boldsymbol{z}\|_2^2}{\partial \boldsymbol{z}} = \boldsymbol{z}^*$
- $\frac{\partial \left(\mathrm{tr}\left(\boldsymbol{z}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{z}\right)\right)}{\partial \boldsymbol{z}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}^*$
- $\frac{\Re\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{z}\right)}{\partial \boldsymbol{z}} = \frac{1}{2}\boldsymbol{w}$
- $\frac{\partial \,\mathrm{tr}(\boldsymbol{A}\boldsymbol{Z})}{\partial \boldsymbol{Z}} = \boldsymbol{A}^{\mathrm{T}}$
- $\frac{\partial \,\mathrm{tr}(\boldsymbol{A}\boldsymbol{Z}^*)}{\partial \boldsymbol{Z}} = \boldsymbol{0}$
- $\frac{\partial \,\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{Z}^{\mathrm{H}}\right)}{\partial \boldsymbol{Z}^*} = \boldsymbol{A}$
- $\frac{\partial \,\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{Z}^{\mathrm{H}}\right)}{\partial \boldsymbol{Z}} = \boldsymbol{0}$
- $\frac{\partial \Re(\mathrm{tr}(\boldsymbol{A}\boldsymbol{Z}))}{\partial \boldsymbol{Z}} = \frac{1}{2}\frac{\partial \,\mathrm{tr}(\boldsymbol{A}\boldsymbol{Z})}{\partial \boldsymbol{Z}} + \frac{1}{2}\frac{\partial \,\mathrm{tr}(\boldsymbol{A}^*\boldsymbol{Z}^*)}{\partial \boldsymbol{Z}} = \frac{1}{2}\boldsymbol{A}^{\mathrm{T}}$
- $\frac{\partial \left(\mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Z}\right)\right)}{\partial \boldsymbol{Z}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{Z}^*$
- $\frac{\partial \left(\mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Z}\right)\right)}{\partial \boldsymbol{Z}^*} = \boldsymbol{A}\boldsymbol{Z}.$

B.4 Kronecker Product

The Kronecker product between $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{C}^{p\times q}$ is a matrix operation defined as

$$\boldsymbol{A}\otimes\boldsymbol{B} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} \otimes \boldsymbol{B} = \begin{bmatrix} a_{1,1}\boldsymbol{B} & \cdots & a_{1,n}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{m,1}\boldsymbol{B} & \cdots & a_{m,n}\boldsymbol{B} \end{bmatrix} \in \mathbb{C}^{mp\times nq}. \tag{B.8}$$

The following properties hold for the Kronecker product of Eq. (B.8) (see [123]):

$$\begin{aligned} (\boldsymbol{A}\otimes\boldsymbol{B})^{\mathrm{T}} &= \boldsymbol{A}^{\mathrm{T}}\otimes\boldsymbol{B}^{\mathrm{T}} \in \mathbb{C}^{nq\times mp} \\ \boldsymbol{A}\otimes\alpha &= \alpha\otimes\boldsymbol{A} \in \mathbb{C}^{m\times n} \\ \boldsymbol{a}^{\mathrm{T}}\otimes\boldsymbol{b} &= \boldsymbol{b}\otimes\boldsymbol{a}^{\mathrm{T}} = \boldsymbol{b}\boldsymbol{a}^{\mathrm{T}} \in \mathbb{C}^{p\times m} \\ (\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D}) &= (\boldsymbol{A}\boldsymbol{C})\otimes(\boldsymbol{B}\boldsymbol{D}) \in \mathbb{C}^{mp\times rs} \\ (\boldsymbol{b}\otimes\boldsymbol{A})\boldsymbol{C} &= \boldsymbol{b}\otimes(\boldsymbol{A}\boldsymbol{C}) \in \mathbb{C}^{pm\times r} \\ (\boldsymbol{A}\otimes\boldsymbol{b})\boldsymbol{C} &= (\boldsymbol{A}\boldsymbol{C})\otimes\boldsymbol{b} \in \mathbb{C}^{pm\times r} \\ (\boldsymbol{E}\otimes\boldsymbol{F})^{-1} &= \boldsymbol{E}^{-1}\otimes\boldsymbol{F}^{-1} \in \mathbb{C}^{mn\times mn} \\ \mathrm{tr}\,(\boldsymbol{E}\otimes\boldsymbol{F}) &= \mathrm{tr}(\boldsymbol{E})\,\mathrm{tr}(\boldsymbol{F}) \in \mathbb{C}. \end{aligned} \tag{B.9}$$

Here, $\alpha \in \mathbb{C}$, $\boldsymbol{a} \in \mathbb{C}^m$, $\boldsymbol{b} \in \mathbb{C}^p$, $\boldsymbol{C} \in \mathbb{C}^{n\times r}$, $\boldsymbol{D} \in \mathbb{C}^{q\times s}$, $\boldsymbol{E} \in \mathbb{C}^{m\times m}$, and $\boldsymbol{F} \in \mathbb{C}^{n\times n}$.

B.5 Real-Valued Notation

Systems can be transformed from the complex-valued notation to the real-valued one according to the following rule

$$\begin{aligned}\boldsymbol{A}_{\text{RV}} &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \otimes \Re\left(\boldsymbol{A}\right) + \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \otimes \Im\left(\boldsymbol{A}\right)\\ &= \begin{pmatrix}\Re\left(\boldsymbol{A}\right) & -\Im\left(\boldsymbol{A}\right)\\ \Im\left(\boldsymbol{A}\right) & \Re\left(\boldsymbol{A}\right)\end{pmatrix} \in \mathbb{R}^{2m\times 2n}\end{aligned} \tag{B.10}$$

where $\otimes$ denotes the Kronecker product (see Section B.4), the prefix $\Re$ and $\Im$ denote the real and imaginary part of its argument, respectively, and $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. It is important to note that the transformation given by Eq. (B.10) preserves all the properties of complex numbers.

The transformation rule for signals is

$$\boldsymbol{b}_{\text{RV}} = \begin{pmatrix}1\\ 0\end{pmatrix} \otimes \Re\left(\boldsymbol{b}\right) + \begin{pmatrix}0\\ 1\end{pmatrix} \otimes \Im\left(\boldsymbol{b}\right) = \begin{pmatrix}\Re\left(\boldsymbol{b}\right)\\ \Im\left(\boldsymbol{b}\right)\end{pmatrix} \in \mathbb{R}^{2n} \tag{B.11}$$

where the signal vector $\boldsymbol{b}$ is a complex column vector of dimensionality n.

Considering the transmission over MIMO channels, if the K-dimensional received signal $\boldsymbol{y}$ is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\eta}$$

where $\boldsymbol{H} \in \mathbb{C}^{K\times N}$ represents the flat fading channel, $\boldsymbol{x} \in \mathbb{C}^{N}$ represents the transmit signal, and $\boldsymbol{\eta} \in \mathbb{C}^{K}$ is the additive white Gaussian noise, the equivalent $2K$-dimensional real valued transmission model is

$$\begin{pmatrix}\Re\left(\boldsymbol{y}\right)\\ \Im\left(\boldsymbol{y}\right)\end{pmatrix} = \begin{pmatrix}\Re\left(\boldsymbol{H}\right) & -\Im\left(\boldsymbol{H}\right)\\ \Im\left(\boldsymbol{H}\right) & \Re\left(\boldsymbol{H}\right)\end{pmatrix} \begin{pmatrix}\Re\left(\boldsymbol{x}\right)\\ \Im\left(\boldsymbol{x}\right)\end{pmatrix} + \begin{pmatrix}\Re\left(\boldsymbol{\eta}\right)\\ \Im\left(\boldsymbol{\eta}\right)\end{pmatrix} \tag{B.12}$$

which can be written as

$$\boldsymbol{y}_{\text{RV}} = \boldsymbol{H}_{\text{RV}}\boldsymbol{x}_{\text{RV}} + \boldsymbol{\eta}_{\text{RV}}. \tag{B.13}$$

As shown in [124], real-valued processing can provide a gain in performance for certain applications.

Appendix C

Karush-Kuhn-Tucker Conditions

Let us consider the following possibly nonlinear optimization problem with equality and inequality constraints:

$$\begin{aligned} \boldsymbol{X}_{\text{opt}} &= \underset{\boldsymbol{X}}{\operatorname{argmin}}\, f(\boldsymbol{X}) \\ &\text{subject to: } g_i(\boldsymbol{X}) \leq 0 \text{ and } h_j(\boldsymbol{X}) = 0 \; \forall i \in \{1,\ldots,l\}, j \in \{1,\ldots,p\} \end{aligned} \tag{C.1}$$

where $\boldsymbol{X}$ and $\boldsymbol{X}_{\text{opt}} \in \mathbb{C}^{m\times n}$. The functions $f(\boldsymbol{X})$, $g_i(\boldsymbol{X})$, $i = 1,\ldots,l$, and $h_j(\boldsymbol{X})$, $j = 1,\ldots,p$, are real-valued with complex-valued arguments, i.e.

$$\begin{aligned} f &: \mathbb{C}^{m\times n} \to \mathbb{R} \\ g_i &: \mathbb{C}^{m\times n} \to \mathbb{R}, \quad i = 1,\ldots,l \\ h_j &: \mathbb{C}^{m\times n} \to \mathbb{R}, \quad j = 1,\ldots,p. \end{aligned}$$

The function to be minimized is $f(\boldsymbol{X})$; $g_i(\boldsymbol{X})$ is the i–th inequality constraint; and $h_j(\boldsymbol{X})$ is the j–th equality constraint, with l and p being the number of inequality and equality constraints, respectively.

Necessary optimality conditions of the optimization in Eq. (C.1) can be found with the Lagrangian function

$$L\left(\boldsymbol{X}, \lambda_{1,1},\ldots,\lambda_{1,l},\lambda_{2,1},\ldots,\lambda_{2,p}\right) = f(\boldsymbol{X}) + \sum_{i=1}^{l} \lambda_{1,i} g_i(\boldsymbol{X}) + \sum_{j=1}^{p} \lambda_{2,j} h_j(\boldsymbol{X})$$

with $\lambda_{1,i} \in \mathbb{R}^{0,+}$, for $i = 1,\ldots,l$, and $\lambda_{2,j} \in \mathbb{R}$, for $j = 1,\ldots,p$.

The *Karush-Kuhn-Tucker* conditions (also known as *KKT* conditions) are necessary for any solution of an optimization problem [68–71]. It is a generalization of the method of Lagrangian multipliers to inequality constraints. These necessary first–order conditions

for the optimization problem of Eq. (C.1) are given by

$$\begin{aligned} \frac{\partial L\left(\boldsymbol{X}, \lambda_{1,1}, \ldots, \lambda_{1,l}, \lambda_{2,1}, \ldots, \lambda_{2,p}\right)}{\partial \boldsymbol{X}} &= \mathbf{0} \\ g_i(\boldsymbol{X}) &\leq 0 \quad i = 1, \ldots, l \\ \lambda_{1,i} g_i(\boldsymbol{X}) &= 0 \quad i = 1, \ldots, l \quad \text{and} \\ \lambda_{1,i} &\geq 0 \quad i = 1, \ldots, l \\ h_j(\boldsymbol{X}) &= 0 \quad j = 1, \ldots, p. \end{aligned}$$

Any optimizer for Eq. (C.1) must fulfill these KKT conditions. However, since the KKT conditions are not sufficient in general, not all candidates obtained from the KKT conditions are optimal.

A function $f(\boldsymbol{X})$ (or $g_i(\boldsymbol{X})$) is convex if its domain $\mathbf{dom} f$ is a convex set, i.e. if $\mathbf{dom} f$ contains the line segment joining any two of its points, and if for all $\boldsymbol{X}, \boldsymbol{Y} \in \mathbf{dom} f$, $\theta \in [0, 1]$

$$f\left(\theta \boldsymbol{X} + (1-\theta)\boldsymbol{Y}\right) \leq \theta f(\boldsymbol{X}) + (1-\theta) f(\boldsymbol{Y}). \tag{C.2}$$

The convexity of a twice differentiable function is also characterized by its second derivative $\nabla^2 f(\boldsymbol{X})$, i.e. a twice differentiable function is convex if and only if its Hessian is positive semidefinite on its domain, and is strictly convex when the Hessian is positive definite.

If the objective and inequality constraint functions are convex and the equality constraint functions are linear (or, more generally, affine), the problem is then a *convex optimization problem* (or convex program) [125]. In the case of a convex optimization problem, the KKT conditions are not only necessary but also sufficient, i.e. any candidate is globally optimal. If the cost function in Eq. (C.1) is strictly convex, then this global solution is unique. In this case, the KKT conditions are also sufficient to solve the given convex programming problem.

Appendix D

Multivariate Normal Distribution

D.1 Mean Vector, Covariance Matrix, and PDF of a Multivariate Normal Distribution

The multivariate normal distribution is the most important distribution in science and engineering. Let $\boldsymbol{x} = [x_1, \ldots, x_m]^{\mathrm{T}} \in \mathbb{C}^m$. The mean value of $\boldsymbol{x}$ is

$$\boldsymbol{\mu}_{\boldsymbol{x}} = \mathrm{E}[\boldsymbol{x}] = [\mu_1, \ldots, \mu_m]^{\mathrm{T}} \in \mathbb{C}^m \tag{D.1}$$

where $\mu_i = \mathrm{E}[x_i]$. Therefore, the mean vector $\boldsymbol{\mu}_{\boldsymbol{x}}$ is a vector of means.

The covariance matrix of $\boldsymbol{x}$ is

$$\boldsymbol{C}_{\boldsymbol{x}} = \mathrm{E}[(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}})(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}})^{\mathrm{H}}] = \{c_{ij}\} \in \mathbb{C}^{m \times m} \tag{D.2}$$

where $\{c_{ij}\}$ denotes the covariance matrix whose elements c_{ij} are given by $\mathrm{E}[(x_i - \mu_i)(x_j - \mu_j)^*]$, i.e. the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ is a matrix of covariances c_{ij}.

The random vector $\boldsymbol{x}$ is said to be multivariate normal (so–called *Gaussian*) if its *Probability Density Function* (PDF) is given by

$$f_{\mathrm{G}}\left(\boldsymbol{x}, \boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{C}_{\boldsymbol{x}}\right) = \frac{\exp\left(-\left(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}}\right)^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \left(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}}\right)\right)}{\pi^m \det(\boldsymbol{C}_{\boldsymbol{x}})}. \tag{D.3}$$

where the notation $\det(\boldsymbol{C}_{\boldsymbol{x}})$ is used for the determinant of $\boldsymbol{C}_{\boldsymbol{x}}$.

D.2 Invariance of Uncorrelated Complex Gaussian Distribution to Unitary Rotations

Let $\boldsymbol{W} \in \mathbb{C}^{m \times n}$ be random, whose elements are i.i.d., zero-mean circularly symmetric complex Gaussian distributed, i.e. $\boldsymbol{w} = \mathrm{vec}(\boldsymbol{W}) \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}_{mn}, \sigma^2 \boldsymbol{I}_{mn})$. Equivalently,

when denoting the i–th column of $\boldsymbol{W}$ as $\boldsymbol{w}_i \in \mathbb{C}^m$, we have that $\boldsymbol{w}_i \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}_m, \sigma^2 \boldsymbol{I}_m)$, $\forall i$, and $\mathrm{E}[\boldsymbol{w}_i \boldsymbol{w}_j^{\mathrm{H}}] = \boldsymbol{0}_{m \times m}$ for $i \neq j$. Suppose a unitary matrix $\boldsymbol{U} \in \mathbb{C}^{m \times m}$ is applied to $\boldsymbol{W}$ from the left leading to

$$\boldsymbol{V} = \boldsymbol{U}\boldsymbol{W} \in \mathbb{C}^{m \times n}.$$

Clearly, $\boldsymbol{V}$ is zero-mean circularly symmetric complex Gaussian distributed as $\boldsymbol{W}$, since a linear transformation applied to $\boldsymbol{W}$ gives $\boldsymbol{V}$. Clearly, the columns of $\boldsymbol{V}$ are uncorrelated just as the columns of $\boldsymbol{W}$ are. As the columns are zero-mean complex Gaussian distributed, we only have to investigate their covariance matrices

$$\mathrm{E}[\boldsymbol{v}_i \boldsymbol{v}_i^{\mathrm{H}}] = \mathrm{E}[\boldsymbol{U}\boldsymbol{w}_i \boldsymbol{w}_i^{\mathrm{H}} \boldsymbol{U}^{\mathrm{H}}] = \boldsymbol{U}\,\mathrm{E}[\boldsymbol{w}_i \boldsymbol{w}_i^{\mathrm{H}}] \boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}\sigma^2 \boldsymbol{I}_m \boldsymbol{U}^{\mathrm{H}} = \sigma^2 \boldsymbol{I}_m$$

where $\boldsymbol{v}_i \in \mathbb{C}^m$ denotes the i–th column of $\boldsymbol{V}$. We see that the entries of the columns are independent, that is, the entries of $\boldsymbol{V}$ are i.i.d. zero-mean circularly symmetric complex Gaussian with variance σ^2. Thus, we have that $\boldsymbol{v} = \mathrm{vec}(\boldsymbol{V}) \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}_{mn}, \sigma^2 \boldsymbol{I}_{mn})$, i.e. the distribution of $\boldsymbol{V}$ is the same as the distribution of $\boldsymbol{W}$. Therefore, the distribution of $\boldsymbol{W}$ is invariant to a unitary rotation from the left. It is also easy to see that the distribution of $\boldsymbol{W}$ is invariant to a unitary rotation from the right (just consider the Hermitian of $\boldsymbol{W}$).

Appendix E

Error Covariance Matrix for Random Vector Quantization

In this appendix, we show how to solve $\mathrm{E}[\min_i ||\hat{\boldsymbol{h}}_{\mathrm{es},k}[\nu] - \boldsymbol{y}_{k,i}||_2^2]$ of Eq. (6.48) to find a closed-form solution for modeling the error matrix for robust precoder designs when random vector quantization is performed instead of scalar quantization.

We have that $||\boldsymbol{\epsilon}_i||_2^2$ with $\boldsymbol{\epsilon}_i \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\hat{\boldsymbol{h}},k} + \boldsymbol{C}_{\boldsymbol{h},k})$ has the same distribution as $\beta_i = ||\boldsymbol{z}_i||_2^2$ with $\boldsymbol{z}_i \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{\Lambda})$, where $\boldsymbol{\Lambda}$ is the diagonal matrix containing the eigenvalues of $\boldsymbol{C}_{\hat{\boldsymbol{h}},k} + \boldsymbol{C}_{\boldsymbol{h},k}$. In the following steps, we demonstrate how to obtain the *Probability Density Function* (PDF) of $\beta_i \in \mathbb{R}^{0,+}$ for $N = 4$ transmit antennas and also the mean of $\min_i \beta_i$. Clearly, we can exploit the independence of the entries of $\boldsymbol{z}_i$, since $\boldsymbol{z}_i$ is complex Gaussian. It is known that if $x \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_x^2)$, $|x|^2$ is exponentially distributed, i.e.

$$f_{|x|^2}(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{\sigma_x^2} \exp\left(-\frac{y}{\sigma_x^2}\right) & \text{otherwise.} \end{cases}$$

In other words, the squares of the entries of $\boldsymbol{z}_i$ are χ^2-distributed with two degrees of freedom and the variance of the j–th entry $z_{i,j}$ of $\boldsymbol{z}_i$ is λ_j. Therefore, the sum of the squares of the first two entries of $\boldsymbol{z}_i$ has the PDF

$$\begin{aligned} f_{|z_{i,1}|^2+|z_{i,2}|^2}(y) &= \int_{-\infty}^{\infty} f_{|z_{i,1}|^2}(y-x) f_{|z_{i,2}|^2}(x) \mathrm{d}x \\ &= \int_0^y \frac{1}{\lambda_1 \lambda_2} \exp\left(-\frac{y-x}{\lambda_1}\right) \exp\left(-\frac{x}{\lambda_2}\right) \mathrm{d}x \\ &= \frac{1}{\lambda_1 \lambda_2} \exp\left(-\frac{y}{\lambda_1}\right) \int_0^y \exp\left(\frac{\lambda_2 - \lambda_1}{\lambda_1 \lambda_2} x\right) \mathrm{d}x \\ &= \frac{1}{\lambda_1 - \lambda_2} \left(\exp\left(-\frac{y}{\lambda_1}\right) - \exp\left(-\frac{y}{\lambda_2}\right)\right) \end{aligned}$$

for $y \geq 0$ and is zero otherwise. Similarly, the PDF for the third and the fourth entry of $\boldsymbol{z}_i$ can be written as

$$f_{|z_{i,3}|^2+|z_{i,4}|^2}(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{\lambda_3-\lambda_4}\left(\exp\left(-\frac{y}{\lambda_3}\right) - \exp\left(-\frac{y}{\lambda_4}\right)\right) & \text{otherwise.} \end{cases}$$

These results help to find the PDF of $\beta_i = \sum_{j=1}^{4} |z_{i,j}|^2$, because

$$\begin{aligned} f_{\beta_i}(y) &= \int_0^y f_{|z_{i,1}|^2+|z_{i,2}|^2}(y-x) f_{|z_{i,3}|^2+|z_{i,4}|^2}(x)\mathrm{d}x = \frac{1}{\lambda_1-\lambda_2}\frac{1}{\lambda_3-\lambda_4} \\ &\quad \times \left(\exp\left(-\frac{y}{\lambda_1}\right) \int_0^y \exp\left(\frac{\lambda_3-\lambda_1}{\lambda_1\lambda_3}x\right)\mathrm{d}x - \exp\left(-\frac{y}{\lambda_1}\right)\int_0^y \exp\left(\frac{\lambda_4-\lambda_1}{\lambda_1\lambda_4}x\right)\mathrm{d}x \right. \\ &\quad \left. + \exp\left(-\frac{y}{\lambda_2}\right) \int_0^y \exp\left(\frac{\lambda_4-\lambda_2}{\lambda_2\lambda_4}x\right)\mathrm{d}x - \exp\left(-\frac{y}{\lambda_2}\right)\int_0^y \exp\left(\frac{\lambda_3-\lambda_2}{\lambda_2\lambda_3}x\right)\mathrm{d}x \right) \\ &= \sum_{k=1}^{4} \frac{\lambda_k^2}{\prod_{j=1,j\neq k}^{4}(\lambda_k-\lambda_j)} \exp\left(-\frac{y}{\lambda_k}\right). \end{aligned}$$

The *Cumulative Distribution Function* (CDF) of β_i is found by integration

$$\begin{aligned} F_{\beta_i}(\beta) = \Pr[\beta_i \leq \beta] &= \int_0^\beta f_{\beta_i}(y)\mathrm{d}y = \sum_{k=1}^{4} \frac{\lambda_k^2}{\prod_{j=1,j\neq k}^{4}(\lambda_k-\lambda_j)} \int_0^\beta \exp\left(-\frac{y}{\lambda_k}\right)\mathrm{d}y \\ &= \sum_{k=1}^{4} \frac{\lambda_k^3}{\prod_{j=1,j\neq k}^{4}(\lambda_k-\lambda_j)} \left(1-\exp\left(-\frac{\beta}{\lambda_k}\right)\right). \end{aligned}$$

Note that $f_{\beta_i}(y)$ is independent of the index i. Since we take the minimum of M square errors β_i and the errors are independent (remember that the codebook entries are independent), the complementary cumulative distribution of $\beta_{\text{min}} = \min_i \beta_i$ can be expressed as

$$\begin{aligned} 1 - F_{\beta_{\text{min}}}(\beta) &= \prod_{i=1}^{M}(1-F_{\beta_i}(\beta)) \\ &= \left(1 - \sum_{k=1}^{4} \frac{\lambda_k^3}{\prod_{j=1,j\neq k}^{4}(\lambda_k-\lambda_j)} \left(1-\exp\left(-\frac{\beta}{\lambda_k}\right)\right)\right)^M. \end{aligned} \tag{E.1}$$

Therefore, the mean of β_{min} can be found as (see [126], (5-27))

$$\mathrm{E}[\beta_{\text{min}}] = \mathrm{E}[\min_i ||\hat{\boldsymbol{h}}_{\text{es},k}[\nu] - \boldsymbol{y}_i||_2^2] = \int_0^\infty (1 - F_{\beta_{\text{min}}}(\beta))\mathrm{d}\beta. \tag{E.2}$$

An analytical integration is in principle possible, but the resulting number of terms even for moderate M ($M = 2^{10} = 1024$) is huge. Therefore, we cannot obtain a closed-form solution for this integral and we have to solve it by means of a numerical integration.

Appendix F

Rectangular Multivariate Gaussian Integrals

F.1 Rectangular Multivariate Gaussian Probability

With $\boldsymbol{\omega}_k = \Re(\breve{\boldsymbol{h}}_k)$ and $\boldsymbol{\chi}_k = \Im(\breve{\boldsymbol{h}}_k)$, the PDF of $\breve{\boldsymbol{h}}_k$ can be decomposed as follows

$$\kappa_k = f_{\breve{\boldsymbol{h}}_k}\left(\breve{\boldsymbol{h}}_k = \boldsymbol{\omega}_k + \mathrm{j}\,\boldsymbol{\chi}_k\right) = f_{\boldsymbol{\omega}_k}(\boldsymbol{\omega}_k) f_{\boldsymbol{\omega}_k}(\boldsymbol{\chi}_k).$$

Let the coefficients defining $\mathbb{S}_{\ell_k}$ [see Eq. (6.85)] be renamed as α_i^{Re}, β_i^{Re}, α_i^{Im}, and β_i^{Im} (the index k and the argument ℓ_k are dropped). Then, we have that

$$\begin{aligned}
\kappa_k &= \int_{\alpha_1^{\mathrm{Re}}}^{\beta_1^{\mathrm{Re}}} \mathrm{d}\omega_{k,1} \cdots \int_{\alpha_d^{\mathrm{Re}}}^{\beta_d^{\mathrm{Re}}} \mathrm{d}\omega_{k,d} \frac{1}{(\sqrt{\pi})^d} \exp\left(-\sum_{j=1}^{d} \omega_{k,j}^2\right) \\
&\quad \times \int_{\alpha_1^{\mathrm{Im}}}^{\beta_1^{\mathrm{Im}}} \mathrm{d}\chi_{k,1} \cdots \int_{\alpha_d^{\mathrm{Im}}}^{\beta_d^{\mathrm{Im}}} \mathrm{d}\chi_{k,d} \frac{1}{(\sqrt{\pi})^d} \exp\left(-\sum_{j=1}^{d} \chi_{k,j}^2\right) \\
&= \prod_{j=1}^{d} \int_{\alpha_j^{\mathrm{Re}}}^{\beta_j^{\mathrm{Re}}} \frac{1}{\sqrt{\pi}} \exp\left(-\omega_{k,j}^2\right) \mathrm{d}\omega_{k,j} \int_{\alpha_j^{\mathrm{Im}}}^{\beta_j^{\mathrm{Im}}} \frac{1}{\sqrt{\pi}} \exp\left(-\chi_{k,j}^2\right) \mathrm{d}\chi_{k,j} \\
&= \prod_{j=1}^{d} \left(\Phi\left(\sqrt{2}\alpha_j^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_j^{\mathrm{Re}}\right)\right)\left(\Phi\left(\sqrt{2}\alpha_j^{\mathrm{Im}}\right) - \Phi\left(\sqrt{2}\beta_j^{\mathrm{Im}}\right)\right) \qquad \text{(F.1)}
\end{aligned}$$

where we use $\int_a^b \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\,\mathrm{d}t = \Phi(a) - \Phi(b)$ for the last equality.

F.2 Rectangular Multivariate Gaussian Centroid

Due to the symmetry of the real and imaginary part of $\boldsymbol{m}_k$, it suffices to find the real part of $\boldsymbol{m}_k$ to prove Eq. (6.86). Let us split up $\breve{\boldsymbol{h}}_k$ into its real and imaginary part, i.e. $\boldsymbol{\omega}_k = \Re(\breve{\boldsymbol{h}}_k)$ and $\boldsymbol{\chi}_k = \Im(\breve{\boldsymbol{h}}_k)$, as we did in Section F.1. Thus, we have that

$$\boldsymbol{\mu}_k = \frac{1}{\kappa_k} \int_{\mathbb{S}_{\ell_k}} \frac{1}{\pi^d} \left(\boldsymbol{\omega}_k + \mathrm{j}\, \boldsymbol{\chi}_k\right) \exp\left(- \sum_{j=1}^{d} \left(\omega_{k,j}^2 + \chi_{k,j}^2\right) \right) \mathrm{d}\boldsymbol{\omega}_k \, \mathrm{d}\boldsymbol{\chi}_k$$

where κ_k is the same integral as the one considered in Section F.1. From Eq. (6.86), we have that $\boldsymbol{m}_k = \boldsymbol{\mu}_k^{\mathrm{Re}} + \mathrm{j}\, \boldsymbol{\mu}_k^{\mathrm{Im}}$. Let $\mu_{k,i}^{\mathrm{Re}}$ be the i–th entry of $\boldsymbol{\mu}_k^{\mathrm{Re}}$. Taking into account that each entry of $\boldsymbol{\omega}_k$ is a standard Gaussian with variance $1/2$, $\mu_{k,i}^{\mathrm{Re}}$ is given by

$$\begin{aligned}
\mu_{k,i}^{\mathrm{Re}} &= \prod_{j=1}^{d} \left(\Phi\left(\sqrt{2}\alpha_j^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_j^{\mathrm{Re}}\right) \right)^{-1} \\
&\quad \times \int_{\alpha_1^{\mathrm{Re}}}^{\beta_1^{\mathrm{Re}}} \mathrm{d}\omega_{k,1} \cdots \int_{\alpha_d^{\mathrm{Re}}}^{\beta_d^{\mathrm{Re}}} \mathrm{d}\omega_{k,d} \frac{\omega_{k,i}}{\sqrt{2}\left(\sqrt{\pi}\right)^d} \exp\left(- \sum_{j=1}^{d} \omega_{k,j}^2 \right) \\
&= \frac{\prod_{j=1, j\neq i}^{d} \left(\Phi\left(\sqrt{2}\alpha_j^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_j^{\mathrm{Re}}\right) \right)}{\prod_{j=1}^{d} \left(\Phi\left(\sqrt{2}\alpha_j^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_j^{\mathrm{Re}}\right) \right)} \int_{\alpha_i^{\mathrm{Re}}}^{\beta_i^{\mathrm{Re}}} \frac{1}{\sqrt{2\pi}} \omega_{k,i} \exp\left(-\omega_{k,i}^2\right) \mathrm{d}\omega_{k,i} \\
&= \left(\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Re}}\right) \right)^{-1} \int_{\sqrt{2}\alpha_i^{\mathrm{Re}}}^{\sqrt{2}\beta_i^{\mathrm{Re}}} \frac{1}{2\sqrt{\pi}} t \exp\left(-t^2/2\right) \mathrm{d}t \\
&= \frac{1}{2\sqrt{\pi}} \frac{\exp\left(-\alpha_i^{\mathrm{Re},2}\right) - \exp\left(-\beta_i^{\mathrm{Re},2}\right)}{\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Re}}\right)}.
\end{aligned} \tag{F.2}$$

Following similar steps for $\mu_{k,i}^{\mathrm{Im}}$, we obtained that

$$\mu_{k,i}^{\mathrm{Im}} = \frac{1}{2\sqrt{\pi}} \frac{\exp\left(-\alpha_i^{\mathrm{Im},2}\right) - \exp\left(-\beta_i^{\mathrm{Im},2}\right)}{\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Im}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Im}}\right)} \tag{F.3}$$

and thus, we have obtained a closed-form solution to Eq. (6.86).

F.3 Rectangular Multivariate Gaussian Covariance

That Eq. (6.87) holds for the off-diagonal elements can be easily shown with similar steps as in Section F.2. So, we only have to obtain the expression for $\sigma_{k,i}$ that can be found in Eq. (6.90). Due to Eq. (6.87), we have that

$$\sigma_{k,i} = [\boldsymbol{M}_k]_{i,i} - |m_{k,i}|^2$$

where $[\boldsymbol{M}_k]_{i,i}$ denotes the i-th diagonal element of $\boldsymbol{M}_k$. With $\boldsymbol{\omega}_k = \Re(\breve{\boldsymbol{h}}_k)$, $\boldsymbol{\chi}_k = \Im(\breve{\boldsymbol{h}}_k)$, and $\lambda_{k,i} = [\boldsymbol{M}_k]_{i,i}$, we have that

$$\lambda_{k,i} = \frac{1}{\kappa_k} \int_{\mathbb{S}_{\ell_k}} \frac{1}{\pi^d} \left(\omega_{k,i} + \mathrm{j}\chi_{k,i}\right) \left(\omega_{k,i} - \mathrm{j}\chi_{k,i}\right) \exp\left(- \sum_{j=1}^{d} \left(\omega_{k,j}^2 + \chi_{k,j}^2\right) \right) \mathrm{d}\boldsymbol{\omega}_k \, \mathrm{d}\boldsymbol{\chi}_k.$$

As shown in Appendix F.1, $\kappa_k = \prod_{j=1}^{d} (\Phi(\sqrt{2}\alpha_j^{\mathrm{Re}}) - \Phi(\sqrt{2}\beta_j^{\mathrm{Re}}))(\Phi(\sqrt{2}\alpha_j^{\mathrm{Im}}) - \Phi(\sqrt{2}\beta_j^{\mathrm{Im}}))$. Taking into account that each entry of $\boldsymbol{\omega}_k$ or $\boldsymbol{\chi}_k$ follows a zero–mean Gaussian distribution with variance $1/2$, we have that

$$\begin{aligned}
\lambda_{k,i} &= \left(\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Re}}\right)\right)^{-1} \int_{\sqrt{2}\alpha_i^{\mathrm{Re}}}^{\sqrt{2}\beta_i^{\mathrm{Re}}} \frac{1}{\sqrt{2\pi}} \frac{\omega_{k,i}^2}{2} \exp\left(-\omega_{k,i}^2/2\right) \mathrm{d}\omega_{k,i} \\
&\quad + \left(\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Im}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Im}}\right)\right)^{-1} \int_{\sqrt{2}\alpha_i^{\mathrm{Im}}}^{\sqrt{2}\beta_i^{\mathrm{Im}}} \frac{1}{\sqrt{2\pi}} \frac{\chi_{k,i}^2}{2} \exp\left(-\chi_{k,i}^2/2\right) \mathrm{d}\chi_{k,i} \\
&= 1 + \frac{1}{2\sqrt{\pi}} \frac{\alpha_i^{\mathrm{Re}} \exp\left(-\alpha_i^{\mathrm{Re},2}\right) - \beta_i^{\mathrm{Re}} \exp\left(-\beta_i^{\mathrm{Re},2}\right)}{\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Re}}\right)} \\
&\quad + \frac{1}{2\sqrt{\pi}} \frac{\alpha_i^{\mathrm{Im}} \exp\left(-\alpha_i^{\mathrm{Im},2}\right) - \beta_i^{\mathrm{Im}} \exp\left(-\beta_i^{\mathrm{Im},2}\right)}{\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Im}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Im}}\right)}
\end{aligned} \tag{F.4}$$

where $\int \frac{1}{\sqrt{2\pi}} x^2 \exp(-x^2/2)\,\mathrm{d}x = \int \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\,\mathrm{d}x - \frac{1}{\sqrt{2\pi}} x \exp(-x^2/2)$ is applied. For brevity, the last equality of Eq. (F.4) will be denoted as $\lambda_{k,i}^{\mathrm{Re}} + \lambda_{k,i}^{\mathrm{Im}}$, where $\lambda_{k,i}^{\mathrm{Re}}$ and $\lambda_{k,i}^{\mathrm{Im}}$ are then given by

$$\begin{aligned}
\lambda_{k,i}^{\mathrm{Re}} &= \frac{1}{2} + \frac{1}{2\sqrt{\pi}} \frac{\alpha_i^{\mathrm{Re}} \exp\left(-\alpha_i^{\mathrm{Re},2}\right) - \beta_i^{\mathrm{Re}} \exp\left(-\beta_i^{\mathrm{Re},2}\right)}{\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Re}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Re}}\right)} \\
\lambda_{k,i}^{\mathrm{Im}} &= \frac{1}{2} + \frac{1}{2\sqrt{\pi}} \frac{\alpha_i^{\mathrm{Im}} \exp\left(-\alpha_i^{\mathrm{Im},2}\right) - \beta_i^{\mathrm{Im}} \exp\left(-\beta_i^{\mathrm{Im},2}\right)}{\Phi\left(\sqrt{2}\alpha_i^{\mathrm{Im}}\right) - \Phi\left(\sqrt{2}\beta_i^{\mathrm{Im}}\right)}
\end{aligned} \tag{F.5}$$

respectively.

From Eq. (6.86), $|m_{k,i}|^2 = \mu_{k,i}^{\mathrm{Re},2} + \mu_{k,i}^{\mathrm{Im},2}$ and thus [cf. Eq. (6.90)]

$$\sigma_{k,i} = \tau_{k,i}^{\mathrm{Re}} + \tau_{k,i}^{\mathrm{Im}}$$

where $\tau_{k,i}^{\mathrm{Re}} = \lambda_{k,i}^{\mathrm{Re}} - \mu_{k,i}^{\mathrm{Re},2}$ and $\tau_{k,i}^{\mathrm{Im}} = \lambda_{k,i}^{\mathrm{Im}} - \mu_{k,i}^{\mathrm{Im},2}$.

Appendix G

List of Acronyms

AoA Angle of Arrival

AoD Angle of Departure

AS Angular Spread

AWGN Additive White Gaussian Noise

BC Broadcast Channel

BER Bit Error Rate

BS Base Station

CDF Cumulative Distribution Function

CDMA Code Division Multiple Access

CSI Channel State Information

CSIR Receiver Channel State Information

CSIT Transmitter Channel State Information

dB Decibels

DFE Decision Feedback Equalization

DPC Dirty Paper Coding

DS Delay Spread

EASI Equivariant Adaptive Separation via Independence

EXIT EXtrinsic Information Transfer

FDD Frequency-Division Duplex

FDMA Frequency-Division Multiple Access

3GPP Third Generation Partnership Project

iid independent and identically distributed

ISI Intersymbol Interference

KKT Karush-Kuhn-Tucker

LLL Lenstra-Lenstra-Lovász

LOS Line of Sight

LP Linear Precoding

LS Least Squares

LTE Long-Term Evolution

MAN Metropolitan Area Networks

MF Matched Filter

MIMO Multiple-Input Multiple-Output

MISO Multiple-Input Single-Output

ML Maximum Likelihood

MMSE Minimum Mean Square Error

MS Mobile Station

MSE Mean Square Error

MU Multi-User

NLOS Non Line of Sight

OFDM Orthogonal Frequency Division Multiplexing

PDF Probability Density Function

QAM Quadrature Amplitude Modulation

QPSK Quadrature Phase Shift Keying

RF Radio Frequency

RF-FE Radio Frequency Front-End

RMS Root Mean Square

RVQ Random Vector Quantization

RxMF Receive Matched Filter

RxWF Receive Wiener Filter

RxZF Receive Zero-Forcing Filter

SCM Spatial Channel Model

SINR Signal-to-Interference-plus-Noise-Ratio

SISO Single-Input Single-Output

SNR Signal-to-Noise Ratio

ST Space Time

SU Single-User

SVD Singular-Value Decomposition

TDD Time-Division Duplex

TDMA Time-Division Multiple Access

THP Tomlinson-Harashima Precoding

TxMF Transmit Matched Filter

TxWF Transmit Wiener Filter

TxZF Transmit Zero-Forcing Filter

US Uncorrelated Scattering

VP Vector Precoding

VQ Vector Quantizer

WF Wiener Filter

WiMAX Worldwide Interoperability for Microwave Access

WSS Wide Sense Stationary

ZF Zero–Forcing

References

[1] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient Use of Side Information in Multiple-Antenna Data Transmission over Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, October 1998.

[2] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback Gain in Multiple Antenna Systems," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 785–798, May 2002.

[3] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian Beamforming for Multiple-Input Multiple-Output Wireless Systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2735–2747, October 2003.

[4] V. Lau, Y. Liu, and T. Chen, "On the Design of MIMO Block-Fading Channels With Feedback-Link Capacity Constraint," *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 62–70, January 2004.

[5] J. C. Roh and B. D. Rao, "Transmit Beamforming in Multiple-Antenna Systems With Finite Rate Feedback: A VQ-Based Approach," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1101–1112, March 2006.

[6] M. Costa, "Writing on Dirty Paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[7] P. Viswanath and D. N. C. Tse, "Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, August 2003.

[8] M. Schubert and H. Boche, "Iterative Multiuser Uplink and Downlink Beamforming Under SINR Constraints," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2324–2334, July 2005.

[9] G. Ginis and J. M. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1085–1104, June 2002.

[10] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in Multiantenna and Multiuser Communications," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1305–1316, July 2004.

[11] K. Kusume, M. Joham, W. Utschick, and G. Bauch, "Cholesky Factorization with Symmetric Permutation Applied to Detecting and Precoding Spatially Multiplexed Data Streams," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 3089–3103, June 2007.

[12] R. El Assir, F. A. Dietrich, M. Joham, and W. Utschick, "Min-Max MSE Precoding for Broadcast Channels based on Statistical Channel State Information," in *Proc. SPAWC 2006*, July 2006.

[13] B. Chun, "A Downlink Beamforming in Consideration of Common Pilot and Phase Mismatch," in *Proc. European Conference on Circuit Theory and Design*, September 2005.

[14] M. Schubert and S. Shi, "MMSE Transmit Optimization with Interference Pre-Compensation," in *Proc. VTC 2005 Spring*, vol. 2, May 2005, pp. 845–849.

[15] F. Rey, M. Lamarca, and G. Vázquez, "Optimal Power Allocation with Partial Channel Knowledge for MIMO Multicarrier Systems," in *Proc. VTC 2002-Fall*, vol. 4, September 2002, pp. 2121–2125.

[16] A. P. Liavas, "Tomlinson-Harashima Precoding With Partial Channel Knowledge," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 5–9, January 2005.

[17] F. A. Dietrich, W. Utschick, and P. Breun, "Linear Precoding Based on a Stochastic MSE Criterion," in *Proc. of the 13th European Signal Processing Conference*, September 2005.

[18] F. A. Dietrich, P. Breun, and W. Utschick, "Robust Tomlinson-Harashima Precoding for the Wireless Broadcast Channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 631–644, February 2007.

[19] R. L. Choi and R. D. Murch, "New Transmit Schemes and Simplified Receiver for MIMO Wireless Communication Systems," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1217–1230, November 2003.

[20] M. Joham, W. Utschick, and J. A. Nossek, "Linear Transmit Processing in MIMO Communications Systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, August 2005.

[21] M. Joham and W. Utschick, "Ordered Spatial Tomlinson Harashima Precoding," in *Smart Antennas — State-of-the-Art*, ser. EURASIP Book Series on Signal Processing and Communications, T. Kaiser, A. Bourdoux, H. Boche, J. R. Fonollosa, J. B. Andersen, and W. Utschick, Eds. EURASIP, Hindawi Publishing Corporation, 2006, vol. 3, ch. III. Transmitter, pp. 401–422.

[22] D. Schmidt, M. Joham, and W. Utschick, "Minimum Mean Square Error Vector Precoding," in *Proc. PIMRC'05*, vol. 1, September 2005, pp. 107–111.

[23] D. A. Schmidt, M. Joham, and W. Utschick, "Minimum Mean Square Error Vector Precoding," *European Transactions on Telecommunications*, vol. 19, no. 3, pp. 219–231, March/April 2008.

[24] W. Keusgen, C. M. Walke, and B. Rembold, "A System Model Considering the Influence of Front-End Imperfections on the Reciprocity of Up- and Downlink System Impulse Responses," in *Proc. ASST 2001*, September 2001, pp. 243–248.

[25] D. Love, R. Heath, W. Santipach, and M. L. Honig, "What is the Value of Limited Feedback for MIMO Channels?" *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, October 2004.

[26] S. M. Kay, *Fundamentals of Statistical Signal Processing – Estimation Theory*. Prentice Hall, 1993.

[27] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[28] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[29] T. S. Rappaport, *Wireless Communications-Principles and Practice*. Prentice-Hall, Englewood Cliffs, NJ, 2001.

[30] J. S. Seybold, *Introduction to RF Propagation*. John Wiley & Sons, 2005.

[31] J. G. Proakis and J. Salehi, *Digital Communications*. McGraw Hill, 1983.

[32] T. M. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge University Press, 2007.

[33] W. Jakes, *Microwave Mobile Communications*. John Wiley & Sons, 1974.

[34] P. Dent, G. E. Bottomley, and T. Croft, "Jakes Fading Model Revisited," *Electronics Letters*, vol. 29, no. 13, pp. 1162–1163, June 1993.

[35] T. Zemen and C. F. Mecklenbräuker, "Time-Variant Channel Estimation Using Discrete Prolate Spheroidal Sequences," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3597–3607, September 2005.

[36] 3rd Generation Partnership Project, Technical Specification Group Radio Access Network, "Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations (Release 6)," September 2003, 3GPP TR 25.996, v6.1.0.

[37] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of Spatial Channel Models for Antenna Array Communication Systems," *IEEE Personal Communications*, vol. 5, pp. 10–22, February 1998.

[38] R. Esmailzadeh and M. Nakagawa, *TDD-CDMA for Wireless Communications*. Artech House, 2002.

[39] ——, "TDD-CDMA for the 4th Generation of Wireless Communications," *IEEE Communications Magazine*, vol. 41, no. 8, pp. 8–15, August 2003.

[40] Philips, *Comparison Between MU-MIMO Codebook-based Channel Reporting Techniques for LTE Downlink*. 3GPP TSG RAN WG1, Tech. Rep. R1-062483, October 2006.

[41] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*. John Wiley & Sons, 2009.

[42] J. G. Andrews, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*. Prentice Hall, 2007.

[43] F. Dietrich, R. Hunger, M. Joham, and W. Utschick, "Robust Transmit Wiener Filter for Time Division Duplex Systems," in *Proc. 3rd International Symposium on Signal Processing and Information Technology*, December 2003, pp. 415–418, Darmstadt, Germany.

[44] F. A. Dietrich and W. Utschick, "Robust Tomlinson-Harashima Precoding," in *Proc. PIMRC 2005*, vol. 1, September 2005, pp. 136–140.

[45] J. P. Costas, "Coding with Linear Systems," *Proceedings of the I.R.E.*, vol. 40, pp. 1101–1103, September 1952.

[46] A. Scaglioni, G. B. Giannakis, and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers Part I: Unification and Optimal Designs," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1988–2006, July 1999.

[47] A. Scaglioni, P. Stoica, S. Barbarossa, and G. B. Giannakis, "Optimal Designs for Space–Time Linear Precoders and Decoders," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1051–1064, May 2002.

[48] M. Joham, *Optimization of Linear and Nonlinear Transmit Signal Processing. PhD dissertation*. Munich University of Technology, 2004.

[49] S. Verdú, *Multiuser Detection*. Cambridge University Press, 1998.

[50] R. N. McDonough and A. D. Whalen, *Detection of Signals in Noise*. Academic Press, 1995.

[51] R. Price and P. E. Green, "A Communication Technique for Multipath Channels," in *Proc. IRE*, vol. 46, March 1958, pp. 555–570.

[52] G. K. Kaleh, "Channel Equalization for Block Transmission Systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 110–121, January 1995.

[53] W. van Etten, "An Optimum Linear Receiver for Multiple Channel Digital Transmission Systems," *IEEE Transactions on Communication Technology*, vol. COM-18, pp. 520–526, October 1970.

[54] D. A. Shnidman, "A Generalized Nyquist Criterion and an Optimum Linear Receiver for a Pulse Modulation System," *The Bell System Technical Journal*, vol. 46, pp. 2163–2177, November 1967.

[55] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. MIT Press, 1949.

[56] A. R. Kaye and D. A. George, "Transmission of Multiplexed PAM Signals over Multiple Channel and Diversity Schemes," *IEEE Transactions on Communication Technology*, vol. COM-18, pp. 520–526, October 1970.

[57] J. H. Winters, "Optimum Combining for Indoor Radio Systems with Multiple Users," *IEEE Transactions on Communications*, vol. COM-35, pp. 1222–1230, November 1987.

[58] R. Esmailzadeh and M. Nakagawa, "Pre-RAKE Diversity Combination for Direct Sequence Spread Spectrum Mobile Communications Systems," *IEEE Transactions on Communications*, vol. E76-B, no. 8, pp. 1008–1015, August 1993.

[59] M. Joham, W. Utschick, and J. A. Nossek, "On the Equivalence of Prerake and Transmit Matched Filter," in *Proc. ASST 2001*, September 2001, pp. 313–318.

[60] J. A. Nossek, M. Joham, and W. Utschick, "Transmit Processing in MIMO Wireless Systems," in *Proc. of the 6th IEEE Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication*, May/June 2004, pp. I–18 – I–23, Shanghai, China.

[61] B. R. Vojčić and W. M. Jang, "Transmitter Precoding in Synchronous Multiuser Communications," *IEEE Transactions on Communications*, vol. 46, pp. 1346–1355, October 1998.

[62] H. R. Karimi, M. Sandell, and J. Salz, "Comparison between Transmitter and Receiver Array Processing to Achieve Interference Nulling and Diversity," in *Proc. PIMRC*, vol. 3, September 1999, pp. 997–1001.

[63] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener Filter for the Downlink of TDD DS-CDMA Systems," in *Proc. ISSSTA*, vol. 1, September 2002, pp. 9–13.

[64] M. E. Austin, "Decision Feedback Equalization for Digital Communication over Dispersive Channels," *Technical report 437. M.I.T. Lincoln Laboratory*, August 1967.

[65] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-Antenna Multi-User Communication—Part II: Perturbation," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 195–202, March 2005.

[66] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, no. 5/6, pp. 138–139, March 1971.

[67] H. Harashima and H. Miyakawa, "Matched-Transmission Technique for Channels With Intersymbol Interference," *IEEE Transactions on Communications*, vol. 20, no. 4, pp. 774–780, August 1972.

[68] W. Karush, *Minima of Functions of Several Variables with Inequalities as Side Conditions. M.S. Thesis*. The University of Chicago, 1939.

[69] R. Fletcher, *Practical Methods of Optimization*. John Wiley & Sons, 1967.

[70] D. G. Luenberger, *Linear and Nonlinear Programming*. Addison-Wesley, 1989.

[71] H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," in *Proc. Second Berkeley Symposium on Mathematical Statistics and Probability*. J. Neyman, University of California Press, 1951, pp. 481–492.

[72] R. Hunger, M. Joham, and W. Utschick, "Extension of linear and nonlinear transmit filters for decentralized receivers," in *European Wireless 2005*, April 2005, pp. 40–46, vol. 1.

[73] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum–likelihood detection and the search for the closest lattice point," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2389–2402, October 2003.

[74] E. Viterbo and E. Biglieri, "A universal decoding algorithm for lattice codes," in *Quatorzieme colloque GRETSI*, September 1993, pp. 611–614, Juan–les–Pins.

[75] B. Hassibi and H. Vikalo, "On the Sphere-decoding Algorithm I. Expected Complexity," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2806–2818, August 2005.

[76] H. Vikalo and B. Hassibi, "On the Sphere-decoding Algorithm II. Generalization, Second-order Statistics, and Applications to Communications," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2819–2834, August 2005.

[77] J. Jaldén and B. Ottersten, "On the Complexity of Sphere Decoding in Digital Communications," in *Proc. IEEE Transactions on Signal Processing*, vol. 53, no. 4, April 2005, pp. 1474–1484.

[78] R. F. H. Fischer, "Lattice-Reduction-Aided Equalization and Generalized Partial Response Signaling for Point-to-Point Transmission over Flat-Fading MIMO Channels," in *Proc. of 4th International Symposium on Turbo Codes in Connection with the 6th International ITG-Conference on Source and Channel Coding*, April 2006, pp. 7–1–7–6.

[79] G. K. Psaltopoulos, M. Joham, and W. Utschick, "Generalized MMSE Detection Techniques for Multipoint–to–Point Systems," in *Proc. Global Telecommunications Conference (GLOBECOM)*, November/December 2006.

[80] C. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment Using Multiple Antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[81] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. Wolniansky, "Detection algorithm and initial laboratory results using V-BlAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14–15, January 1999.

[82] G. Ginis and J. M. Cioffi, "On the Relation between V-BLAST and the GDFE," *IEEE Communications Letters*, vol. 5, no. 9, pp. 364–366, September 2001.

[83] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. John Wiley & Sons, 2002.

[84] A. Duel-Hallen, "A Family of Multiuser Decision-Feedback Detectors for Asymchronous Code-Division Multiple-Access Channels," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 421–434, February/March/April 1995.

[85] K. Kusume, M. Joham, and W. Utschick, "MMSE Block Decision-Feedback Equalizer for Spatial Multiplexing with Reduced Complexity," in *Proc. IEEE Global Telecommunications Conference*, November/December 2004, pp. 2540–2544, vol. 4, Dallas, Texas, USA.

[86] G. H. Golub and C. F. V. Loan, *Matrix Computations*. The Johns Hopkins University Press, 1996.

[87] C. P. Schnorr and M. Euchnerr, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, pp. 188–191, September 1994.

[88] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest Point Search in Lattices," *IEEE Transactions on Information Theory*, vol. 48, no. 8, pp. 2201–2214, August 2002.

[89] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*. Springer, Berlin, 1988.

[90] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring Polynomials with Rational Coefficients," *Math. Annalen*, vol. 261, pp. 515–534, December 1982.

[91] G. Psaltopoulos, M. Joham, and W. Utschick, "Comparison of Lattice Search Techniques for Nonlinear Precoding," in *Proc. International ITG Workshop on Smart Antennas*, March 2006, pp. 997–1001.

[92] K. Kusume, M. Joham, W. Utschick, and G. Bauch, "Efficient Tomlinson-Harashima Precoding for Spatial Multiplexing on Flat MIMO Channel," in *Proc. International Conference on Communications*, vol. 3, May 2005, pp. 2021–2025, Seoul, Korea.

[93] G. Ginis and J. M. Cioffi, "A Multi-User Precoding Scheme achieving Crosstalk Cancellation with Application to DSL Systems," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, vol. 2, October 2000, pp. 1627–1631.

[94] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. H. Huber, "Space-Time Transmission using Tomlinson-Harashima Precoding," in *Proc. of the 4th ITG Conference on Source and Channel Coding*, January 2002, pp. 139–147.

[95] A. Papoulis, *Signal Analysis*. McGraw Hill, 1977.

[96] M. Joham, J. Brehmer, and W. Utschick, "MMSE Approaches to Multiuser Spatio-Temporal Tomlinson-Harashima Precoding," in *Proc. of the 5th International ITG Conference on Source and Channel Coding*, January 2004, pp. 387–394, Erlangen, Germany.

[97] F. A. Dietrich, T. Ivanov, and W. Utschick, "Estimation of Channel and Noise Correlations for MIMO Channel Estimation," in *Proc. ITG/IEEE Workshop on Smart Antennas*, March 2006.

[98] P. M. Castro and L. Castedo, "Adaptive Tomlinson-Harashima precoding for wireless communication systems," in *Proc. Sixth Baiona Workshop on Signal Processing in Communications*, July 2003, pp. 253–258.

[99] ——, "Precodificación Tomlinson-Harashima adaptativa para sistemas de comunicaciones móviles," in *Proc. XVII Simposium Nacional de la Unión Científica Internacional de Radio (URSI)*, September 2003.

[100] P. M. Castro, L. Castedo, and J. Míguez, "Precoding in Wireless communications systems using Particle Filtering for Blind Channel Prediction," in *Proc. of the 5th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2004.

[101] P. M. Castro and L. Castedo, "Predicción de canal para sistemas MIMO con precodificación Tomlinson-Harashima," in *Proc. XVIII Simposium Nacional de la Unión Científica Internacional de Radio (URSI)*, September 2004.

[102] P. M. Castro, L. Castedo, and J. Míguez, "Adaptive Precoding in MIMO Wireless Communication Systems using Blind Channel Prediction over Frequency Selective Fading Channels," in *Proc. of the 13th IEEE Workshop on Statistical Signal Processing (SSP)*, July 2005.

[103] P. M. Castro and L. Castedo, "Adaptive Vector Quantization for Precoding using Blind Channel Prediction in Frequency Selective MIMO Mobile Channels," in *Proc. International ITG/IEEE Workshop on Smart Antennas (WSA)*, April 2005.

[104] P. M. Castro, H. Pérez-Iglesias, A. Dapena, and L. Castedo, "Utilization of Blind Source Separation Algorithms for MIMO Linear Precoding," in *Proc. Independent Component Analysis and Signal Separation (ICA)*, March 2006, pp. 577–584.

[105] P. M. Castro, A. Dapena, and L. Castedo, "Implementation of Linearly Precoded MIMO Communication Systems with BSS Techniques," in *Proc. of the 12th European Wireless (EW)*, April 2006.

[106] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, 1996.

[107] L. L. Scharf, *Statistical Signal Processing – Detection, Estimation, and Time Series Analysis*. Addison-Wesley, 1991.

[108] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer, 1993.

[109] P. M. Castro, M. Joham, L. Castedo, and W. Utschick, "Robust MMSE Linear Precoding for Multiuser MISO Systems with Limited Feedback and Channel Prediction," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2008.

[110] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. H. Huber, "Tomlinson-Harashima Precoding in Space-Time Transmission for Low-Rate Backward Channel," in *Proc. International Zurich Seminar on Broadband Communications*, February 2002, pp. 7–1–7–6.

[111] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Linear and Nonlinear Preequalization/Equalization for MIMO Systems with Long-Term Channel State Information at the Transmitter," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 373–378, March 2004.

[112] V. Lau, Y. Liu, and T. Chen, "Capacity of Memoryless Channels and Block-Fading Channels With Designable Cardinality-Constrained Channel State Feedback," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 2038–2049, September 2004.

[113] T. Yoo and A. Goldsmith, "On the Optimality of Multiantenna Broadcast Scheduling Using Zero-Forcing Beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, March 2006.

[114] P. M. Castro, M. Joham, W. Utschick, and L. Castedo, "Robust Precoding for Multi-User MISO Systems with Limited-Feedback Channels," in *Proc. ITG/IEEE WSA 2007*, February 2007.

[115] P. M. Castro, M. Joham, L. Castedo, and W. Utschick, "Robust multi-user precoding schemes with limited-feedback channels," in *Proc. of XXII Simposium Nacional de la Unión Científica Internacional de Radio (URSI)*, September 2007.

[116] P. M. Castro, L. Castedo, M. Joham, and W. Utschick, "Precodificación TH MMSE robusta para sistemas MISO multiusuario con predicción de canal," in *Proc. of XXIII Simposium Nacional de la Unión Científica Internacional de Radio (URSI)*, September 2008.

[117] P. M. Castro, M. Joham, W. Utschick, and L. Castedo, "Optimized CSI Feedback for Robust THP Design," in *Proc. 41st Asilomar Conference on Signals, Systems and Computers*, October 2007, pp. 1956–1960.

[118] M. Joham, P. M. Castro, L. Castedo, and W. Utschick, "MMSE Optimal Feedback of Correlated CSI for Multi-User Precoding," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008)*, April 2008, pp. 3129–3132.

[119] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. IT-28, no. 2, pp. 129–137, March 1982.

[120] M. Joham, P. M. Castro, J. Zhen, L. Castedo, and W. Utschick, "Joint Design of Limited Feedback and Multiuser Precoding Based on a Precoding MSE Metric," in *Proc. 42nd Asilomar Conference on Signals, Systems, and Computers*, October 2008, pp. 1305–1309.

[121] S. M. Kay, *Fundamentals of Statistical Signal Processing – Estimation Theory*. Prentice Hall, 1993.

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

[123] J. W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Transactions on Circuits and Systems*, vol. 25, pp. 772–781, September 1978.

[124] R. F. H. Fischer and C. Windpassinger, "Real- vs. Complex-valued Equalization in V-BLAST Systems," *Electronics Letters*, no. 39, pp. 470–471, March 2003.

[125] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[126] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw–Hill, 1965.