MIMO Transceiver Design via Majorization Theory

MIMO Transceiver Design via Majorization Theory

Daniel P. Palomar

Dept. Electronic and Computer Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong
palomar@ust.hk

Yi Jiang

Dept. Electrical and Computer Engineering
University of Colorado
Boulder, Colorado 80309
USA
yjiang@dsp.colorado.edu

the essence of knowledge

Boston – Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is D. P. Palomar and Y. Jiang, MIMO Transceiver Design via Majorization Theory, Foundation and Trends® in Communications and Information Theory, vol 3, nos 4-5, pp 331–551, 2006

Printed on acid-free paper

ISBN: 978-1-60198-030-4

Foundations and Trends® in Communications and Information Theory

Volume 3 Issues 4-5, 2006

Editorial Scope

Foundations and Trends® in Communications and Information Theory will publish survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2006, Volume 3, 6 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328. Also available as a combined paper and online subscription.

Foundations and Trends® in
Communications and Information Theory
Vol. 3, Nos. 4-5 (2006) 331–551

DOI: 10.1561/0100000018

MIMO Transceiver Design via Majorization Theory

Daniel P. Palomar[1] **and Yi Jiang**[2]

[1] *Dept. Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, palomar@ust.hk*
[2] *Dept. Electrical and Computer Engineering, University of Colorado, Boulder, Colorado 80309, USA, yjiang@dsp.colorado.edu*

Abstract

Multiple-input multiple-output (MIMO) channels provide an abstract and unified representation of different physical communication systems, ranging from multi-antenna wireless channels to wireless digital subscriber line systems. They have the key property that several data streams can be simultaneously established.

In general, the design of communication systems for MIMO channels is quite involved (if one can assume the use of sufficiently long and good codes, then the problem formulation simplifies drastically). The first difficulty lies on how to measure the global performance of such systems given the tradeoff on the performance among the different data streams. Once the problem formulation is defined, the resulting mathematical problem is typically too complicated to be optimally solved as it is a matrix-valued nonconvex optimization problem. This design problem has been studied for the past three decades (the first papers dating back to the 1970s) motivated initially by cable systems and more recently by wireless multi-antenna systems. The approach was to

choose a specific global measure of performance and then to design the system accordingly, either optimally or suboptimally, depending on the difficulty of the problem.

This text presents an up-to-date unified mathematical framework for the design of point-to-point MIMO transceivers with channel state information at both sides of the link according to an *arbitrary* cost function as a measure of the system performance. In addition, the framework embraces the design of systems with given individual performance on the data streams.

Majorization theory is the underlying mathematical theory on which the framework hinges. It allows the transformation of the originally complicated matrix-valued nonconvex problem into a simple scalar problem. In particular, the *additive* majorization relation plays a key role in the design of *linear* MIMO transceivers (i.e., a linear precoder at the transmitter and a linear equalizer at the receiver), whereas the *multiplicative* majorization relation is the basis for *nonlinear decision-feedback* MIMO transceivers (i.e., a linear precoder at the transmitter and a decision-feedback equalizer at the receiver).

Contents

1
Introduction

This chapter starts by introducing in a concise way the concept and relevance of multiple-input multiple-output (MIMO) channels and by highlighting some of the successful schemes for MIMO communication systems that have been proposed such as space–time coding and linear precoding. Then, a first glimpse at linear transceivers is presented, starting from the classical receive beamforming schemes in *smart antennas* and gradually building on top in a natural way. Finally, a historical account on MIMO transceivers is outlined.

1.1 MIMO Channels

MIMO channels arise in many different scenarios such as wireline systems or multi-antenna wireless systems, where there are multiple transmit and receive dimensions. A MIMO channel is mathematically denoted by a channel matrix which provides an elegant, compact, and unified way to represent physical channels of completely different nature.

The use of multiple dimensions at both ends of a communication link offers significant improvements in terms of *spectral efficiency* and *link reliability*. The most important characteristic of MIMO channels is the *multiplexing gain*, obtained by exploiting the multiple dimensions to open up several parallel *subchannels* within the MIMO channel, also termed channel *eigenmodes*, which leads to an increase of rate. The multiplexing property allows the transmission of several symbols simultaneously or, in other words, the establishment of several *substreams* for communication.

1.1.1 Basic Signal Model

The transmission over a general MIMO communication channel with n_T transmit and n_R receive dimensions can be described with the baseband signal model

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{1.1}$$

as depicted in Figure 1.1, where $\mathbf{s} \in \mathbb{C}^{n_T \times 1}$ is the transmitted vector, $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ is the channel matrix, $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$ is the received vector, and $\mathbf{n} \in \mathbb{C}^{n_R \times 1}$ denotes the noise.

A multicarrier MIMO channel can be similarly described, either explicitly for the N carriers as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{s}_k + \mathbf{n}_k \quad 1 \le k \le N, \tag{1.2}$$

or implicitly as in (1.1) by defining the block-diagonal equivalent matrix $\mathbf{H} = \mathrm{diag}\left(\{\mathbf{H}_k\}\right)$.

When $n_T = 1$, the MIMO channel reduces to a single-input multiple-output (SIMO) channel (e.g., with multiple antennas only at the receiver). Similarly, when $n_R = 1$, the MIMO channel reduces to a

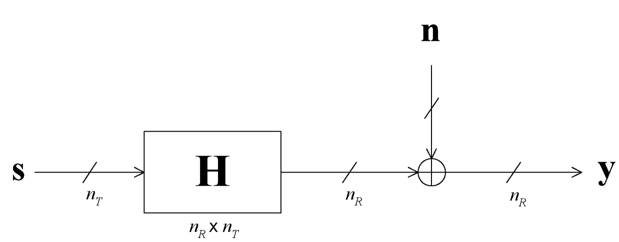

Fig. 1.1 Scheme of a MIMO channel.

multiple-input single-output (MISO) (e.g., with multiple antennas only at the transmitter). When both $n_T = 1$ and $n_R = 1$, the MIMO channel simplifies to a simple scalar or single-input single-output (SISO) channel.

1.1.2 Examples of MIMO Channels

We now briefly illustrate how different physical communication channels can be conveniently modeled as a MIMO channel.

1.1.2.1 Inter-Symbol Interference (ISI) Channel

Consider the discrete-time signal model after symbol-rate sampling

$$y(n) = \sum_{k=0}^{L} h(k)\, s(n-k) + n(n), \tag{1.3}$$

where $h(k)$ are the coefficients of the finite-impulse response (FIR) filter of order L representing the channel.

If the transmitter inserts at least L zeros between blocks of N symbols (termed zero-padding), the MIMO channel model in (1.1) is obtained where the channel matrix $\mathbf{H}$ is a convolutional matrix [122, 131]:

$$\mathbf{H} = \begin{bmatrix} h(0) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ h(L) & & \ddots & 0 \\ 0 & \ddots & & h(0) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h(L) \end{bmatrix}. \tag{1.4}$$

Alternatively, if the transmitter uses a cyclic prefix of at least L symbols between blocks of N symbols, then the linear convolution becomes a circular convolution and the MIMO channel model in (1.1) is obtained where the channel matrix $\mathbf{H}$ is a circulant matrix[1] [122, 131].

[1] In a circulant matrix, the rows are composed of cyclically shifted versions of a sequence [66].

1.1.2.2 Multicarrier Channel

In a multicarrier communication system, the available bandwidth is partitioned into N subbands and then each subband is independently used for transmission [17, 80]. Such an approach not only simplifies the communication process but it is also a capacity-achieving structure for a sufficiently large N [46, 62, 122].

The signal model follows from a block transmission with a cyclic prefix, obtaining a circulant matrix, combined with an inverse/direct discrete Fourier transform (DFT) at the transmitter/receiver. The MIMO channel model in (1.1) is obtained where the channel matrix $\mathbf{H}$ is a diagonal matrix with diagonal elements given by DFT coefficients [54, 86].

1.1.2.3 Multi-Antenna Wireless Channel

The multi-antenna wireless channel with multiple antennas at both sides of the link (see Figure 1.2) is the paradigmatic example of a MIMO channel. In fact, the publication of [43, 146, 148] in the late 1990s about multi-antenna wireless channels boosted the research on MIMO systems. The popularity of this particular scenario is mainly due to the linear increase of capacity with the number of antennas [43, 148] for the same bandwidth.

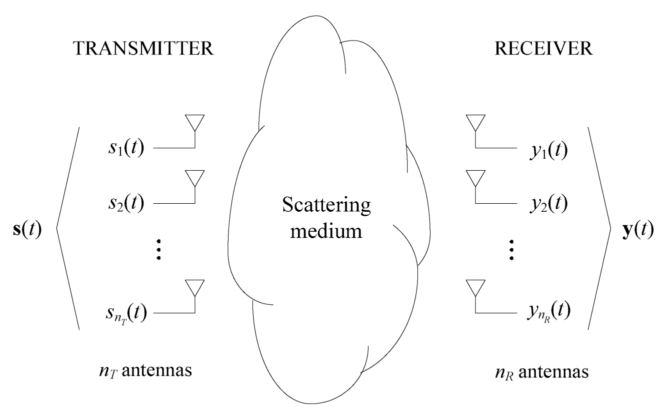

Fig. 1.2 Example of a MIMO channel arising in wireless communications when multiple antennas are used at both the transmitter and the receiver.

If the channel is flat in frequency, then the MIMO channel model in (1.1) follows naturally by defining the (i,j)th element of matrix $\mathbf{H}$ as the channel gain/fading between the jth transmit antenna and the ith receive one. In general, however, the channel will be frequency-selective according to the following matrix convolution:

$$\mathbf{y}(n) = \sum_{k=0}^{L} \mathbf{H}(k)\, \mathbf{s}(n-k) + \mathbf{n}(n) \tag{1.5}$$

where $\mathbf{H}(n)$ are the matrix-coefficients of the FIR matrix filter representing the channel ($[\mathbf{H}(n)]_{ij}$ is the discrete-time channel from the jth transmit antenna to the ith receive one). At this point, the frequency-selective channel in (1.5) can be manipulated as in (1.3) to obtain a block-matrix with each block corresponding to the channel between each transmit–receive pair of antennas; in particular, with zero padding each block will be a convolutional matrix, whereas with cyclic prefix each block will be a circulant matrix [122]. In the case of cyclic prefix, after applying the inverse/direct DFT to each block and a posterior rearrangement of the elements, the multicarrier MIMO signal model in (1.2) is obtained [122], i.e., one basic MIMO channel per carrier.

1.1.2.4 Wireline DSL Channel

Digital Subscriber Line technology has gained popularity as a broadband access technology capable of reliably delivering high data rates over telephone subscriber lines [144]. Modeling a DSL system as a MIMO channel presents many advantages with respect to treating each twisted pair independently [47, 63]. If fact, modeling a wireline channel as a MIMO channel was done three decades ago [90, 129].

The dominant impairment in DSL systems is crosstalk arising from electromagnetic coupling between neighboring twisted-pairs. Near-end crosstalk (NEXT) comprises the signals originated in the same side of the received signal (due to the existence of downstream and upstream transmission) and far-end crosstalk (FEXT) includes the signal originated in the opposite side of the received signal. The impact of NEXT is generally suppressed by employing frequency division duplex (FDD) to separate downstream and upstream transmission.

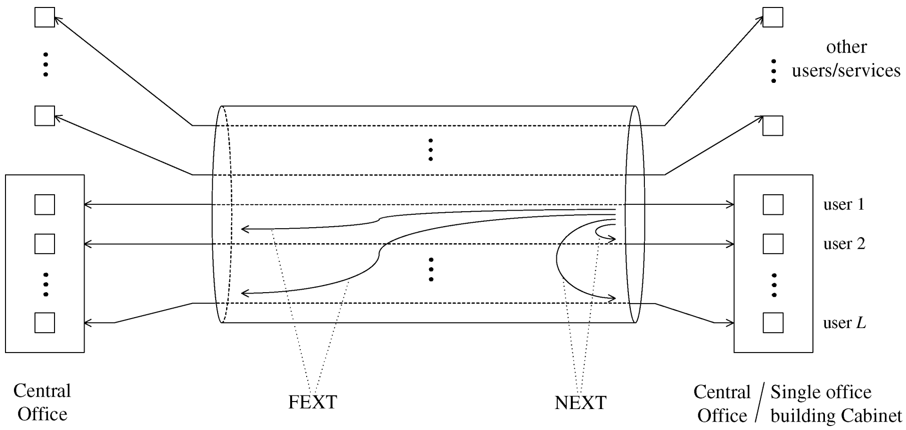

Fig. 1.3 Scheme of a bundle of twisted pairs of a DSL system.

The general case under analysis consists of a binder group composed of L users in the same physical location plus some other users that possibly belong to a different service provider and use different types of DSL systems (see Figure 1.3). The MIMO channel represents the communication of the L intended users while the others are treated as interference.

DSL channels are highly frequency-selective with a signal model as in (1.5); as a consequence, practical communication systems are based on the multicarrier MIMO signal model in (1.2).

1.1.2.5 CDMA Channel

Excess-bandwidth systems (the majority of practical systems) utilize a transmit bandwidth larger than the minimum (Nyquist) bandwidth. Examples are systems using spreading codes and systems using a root-raised cosine transmit shaping pulse (with a nonzero rolloff factor) [120]. For these systems, fractional-rate sampling (sampling at a rate higher than the symbol rate) has significant practical advantages compared to symbol-rate sampling such as the insensitivity with respect to the sampling phase and the possibility to implement in discrete time many of the operations performed at the receiver such as the matched-filtering operation (cf. [121]). Fractionally sampled systems can be modeled as a multirate convolution which can be easily converted into a more convenient vector convolution as in (1.5).

One relevant example of excess-bandwidth system is code division multiple access (CDMA) systems, where multiple users transmit overlapping in time and frequency but using different signature waveforms or spreading codes (which are excess-bandwidth shaping pulses). The discrete-time model for such systems is commonly obtained following a matched filtering approach by sampling at the symbol rate the output of a bank of filters where each filter is matched to one of the signature waveforms [157]. An alternative derivation of the discrete-time model for CDMA systems is based on a fractionally sampled scheme by sampling at the chip rate. Adding up the effect of U users, the final discrete-time (noiseless) signal model is

$$\mathbf{y}(n) = \sum_{u=1}^{U} \sum_{l=0}^{L} \mathbf{h}_u(l)\, s_u(n-l), \tag{1.6}$$

where $\mathbf{h}_u(n)$ is the equivalent chip-rate sampled channel of the uth user defined as $\mathbf{h}_u(n) \triangleq [h_u(nP), \ldots, h_u(nP+(P-1))]^T$, $h_u(n)$ corresponds to the continuous impulse response $h_u(t)$ sampled at time $t = nT/P$, P denotes the oversampling factor or spreading factor, and L is the support of the channel $\mathbf{h}_u(n)$.

1.2 MIMO Communication Systems

A plethora of communication techniques exists for transmission over MIMO channels which essentially depend on the degree of channel state information (CSI) available at the transmitter and at the receiver. Clearly, the more channel information, the better the performance of the system. The reader interested in space–time wireless communication systems is referred to the two 2003 textbooks [87, 115] and to the more extensive 2005 textbooks [16, 50, 151].

CSI at the receiver (CSIR) is traditionally acquired via the transmission of a training sequence (pilot symbols) that allows the estimation of the channel. It is also possible to use blind methods that do not require any training symbols but exploit knowledge of the structure of the transmitted signal or of the channel. CSI at the transmitter (CSIT) is typically obtained either via a feedback channel from the receiver (this technique requires the channel to be sufficiently slowly varying and has a loss in spectral efficiency due to the utilization of part of the bandwidth to transmit the channel state) or by exploiting (whenever possible) the channel reciprocity that allows to infer the channel from previous receive measurements (cf. [10]).

It is generally assumed that perfect CSIR is available. Regarding CSIT, there are two main families of transmission methods that consider either no CSIT or perfect CSIT. In practice, however, it is more realistic to consider imperfect or partial CSIT.

1.2.1 Schemes with No CSIT

Space–time coding generalizes the classical concept of coding on the temporal domain [22] to coding on both spatial and temporal dimensions [1, 146]. The idea is to introduce redundancy in the transmitted signal, both over space and time, to allow the receiver to recover the signal even in difficult propagation situations. The conventional space–time coding trades off spectral efficiency for improved communication reliability. Since the initial papers in 1998 [1, 146], an extraordinary number of publications has flourished in the literature (cf. [37, 38, 87, 105]). The recent space–time block codes proposed in [38] and [105] can achieve the optimum tradeoff between spectral efficiency and transmission reliability, or the diversity–multiplexing gain tradeoff as charted in [178]. The better performance of the advanced space–time codes come with high decoding complexity.

Layered architectures (also termed BLAST[2]) refer to a particular case of a space–time coding when a separate coding scheme is used for each spatial branch, i.e., they are constructed by assembling one-dimensional constituent codes. The diagonal BLAST originally proposed by Foschini in 1996 [45] can in principle achieve the optimal diversity–multiplexing gain tradeoff [178]. However it requires short and powerful coding to eliminate error propagation, which makes it difficult to implement. The simpler vertical BLAST proposed in [44] admits independent coding and decoding for each spatially multiplexed

[2] BLAST stands for Bell-labs LAyered Space–Time architecture [44, 45].

substream, but the simplicity in the equalization and decoding aspects comes with low reliability since vertical BLAST does not collect the diversity across different layers. Hybrid schemes combining layered architectures with constituent space–time codes have been proposed as a reasonable tradeoff between performance and complexity, e.g., [5].

1.2.2 Schemes with Perfect CSIT

When perfect CSIT is available, the transmission can be adapted to each channel realization using signal processing techniques. Historically speaking, there are two main scenarios that have motivated the development of communication methods for MIMO channels with CSIT: wireline channels, and wireless channels.

The initial motivation to design techniques for communication over MIMO channels can be found in wireline systems by treating all the links within a bundle of cables as a whole, e.g., [63, 90, 129, 170, 171]. Another more recent source of motivation to design methods for communication over MIMO channels follows from multi-antenna wireless systems e.g., [3, 122, 131]. A historical perspective on signal processing methods for MIMO systems is given in Section 1.4.

1.2.3 Schemes with Imperfect/Partial CSIT

In real scenarios, it is seldom the case that the CSIT is either inexistent or perfect; in general, its knowledge is partial or imperfect for which hybrid communication schemes are more appropriate.

One basic approach is to start with a space–time code, for which no CSIT is required, and combine it with some type of signal processing technique to take advantage of the partial CSIT, e.g., [76].

Another different philosophy is to start with a signal processing approach, for which typically perfect CSIT is assumed, and make it robust to imperfections in the CSIT, e.g., [10, 101, 123, 159, 165].

1.3 A First Glimpse at Linear Transceivers: Beamforming

Beamforming is a term traditionally associated with array processing or *smart antennas* in wireless communications where an array of antennas exists either at the transmitter or at the receiver [75, 85, 99, 150, 154].

The concept of linear MIMO transceiver is closely related to that of classical beamforming as shown next.

1.3.1 Classical Beamforming for SIMO and MISO Channels

We consider the concept of beamforming over any arbitrary dimension, generalizing the traditional meaning that refers only to the space (antenna) dimension.

Consider a SIMO channel:

$$\mathbf{y} = \mathbf{h}x + \mathbf{n}, \tag{1.7}$$

where one symbol x is transmitted (normalized such that $\mathbb{E}\left[|x|^2\right] = 1$) and a vector $\mathbf{y}$ is received (the noise is assumed zero mean and white $\mathbb{E}\left[\mathbf{n}\mathbf{n}^\dagger\right] = \mathbf{I}$). The classical receive beamforming approach estimates the transmitted symbol by linearly combining the received vector via the beamvector $\mathbf{w}$:

$$\hat{x} = \mathbf{w}^\dagger \mathbf{y} = \mathbf{w}^\dagger \left(\mathbf{h}x + \mathbf{n}\right). \tag{1.8}$$

We can now design the receive beamvector $\mathbf{w}$ to maximize the SNR given by

$$\mathrm{SNR} = \frac{|\mathbf{w}^\dagger \mathbf{h}|^2}{\mathbf{w}^\dagger \mathbf{w}}. \tag{1.9}$$

The solution follows easily from the Cauchy–Schwarz's inequality:

$$|\mathbf{w}^\dagger \mathbf{h}| \leq \|\mathbf{w}\| \, \|\mathbf{h}\|, \tag{1.10}$$

where equality is achieved when $\mathbf{w} \propto \mathbf{h}$, i.e., when the beamvector is aligned with the channel. This is commonly termed *matched filter* or *maximum ratio combining*. The resulting SNR is then given by the squared-norm of the channel $\|\mathbf{h}\|^2$, i.e., fully utilizing the energy of the channel.

Consider now a MISO channel:

$$y = \mathbf{h}^\dagger \mathbf{s} + n, \tag{1.11}$$

where the vector signal $\mathbf{s}$ is transmitted and the scalar y is received (the noise is assumed zero mean and normalized $\mathbb{E}\left[|n|^2\right] = 1$). The classical transmit beamforming approach transmits on each antenna a weighted version of the symbol to be conveyed x via the beamvector $\mathbf{p}$:

$$\mathbf{s} = \mathbf{p}x, \tag{1.12}$$

where the transmitted power is given by the squared-norm of the beamvector $\|\mathbf{p}\|^2$ (assuming $\mathbb{E}\left[|x|^2\right] = 1$). The overall signal model is then

$$y = (\mathbf{h}^\dagger \mathbf{p})x + n. \tag{1.13}$$

We can now design the transmit beamvector $\mathbf{p}$ to maximize the SNR

$$\text{SNR} = |\mathbf{h}^\dagger \mathbf{p}|^2, \tag{1.14}$$

subject to a power constraint $\|\mathbf{p}\|^2 \leq P_0$. The solution again follows easily from the Cauchy–Schwarz's inequality:

$$|\mathbf{h}^\dagger \mathbf{p}| \leq \|\mathbf{h}\| \, \|\mathbf{p}\| \leq \|\mathbf{h}\| \sqrt{P_0}, \tag{1.15}$$

where both equalities are achieved when $\mathbf{p} = \sqrt{P_0}\mathbf{h}/\|\mathbf{h}\|$, i.e., when the beamvector is aligned with the channel and satisfies the power constraint with equality. An alternative way to derive this result is by rewriting the SNR as

$$\text{SNR} = \mathbf{p}^\dagger (\mathbf{h}\mathbf{h}^\dagger)\mathbf{p}, \tag{1.16}$$

from which the maximum value follows straightforwardly as the eigenvector of matrix $\mathbf{h}\mathbf{h}^\dagger$ corresponding to the maximum eigenvalue, which is precisely $\mathbf{h}/\|\mathbf{h}\|$, properly normalized to satisfy the power constraint. The resulting SNR is then given by $P_0 \|\mathbf{h}\|^2$, i.e., fully utilizing the energy of the channel and the maximum power at the transmitter.

1.3.2 Single Beamforming for MIMO Channels

We are now ready to extend the previous treatment of classical beamforming only at the receiver or only at the transmitter to both sides of the link as illustrated in Figure 1.4 (e.g., [3, 113]). Consider now a MIMO channel:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{1.17}$$

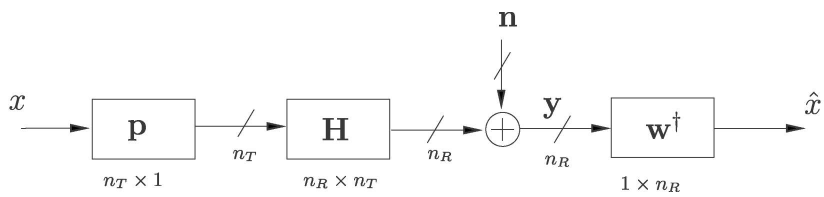

Fig. 1.4 Single beamforming scheme of a MIMO communication system.

where the vector signal $\mathbf{s}$ is transmitted and a vector $\mathbf{y}$ is received (the noise is assumed zero mean and white $\mathbb{E}\left[\mathbf{n}\mathbf{n}^\dagger\right] = \mathbf{I}$). The transmit beamforming generates the vector signal with beamvector $\mathbf{p}$ as

$$\mathbf{s} = \mathbf{p}x, \tag{1.18}$$

where one symbol x is transmitted (normalized such that $\mathbb{E}\left[|x|^2\right] = 1$), and the receive beamforming estimates the transmitted symbol by linearly combining the received vector with the beamvector $\mathbf{w}$:

$$\hat{x} = \mathbf{w}^\dagger\mathbf{y} = \mathbf{w}^\dagger\left(\mathbf{H}\mathbf{p}x + \mathbf{n}\right). \tag{1.19}$$

The SNR is given by

$$\mathrm{SNR} = \frac{|\mathbf{w}^\dagger\mathbf{H}\mathbf{p}|^2}{\mathbf{w}^\dagger\mathbf{w}}. \tag{1.20}$$

We can now maximize it with respect to the receive beamvector $\mathbf{w}$, for a given fixed $\mathbf{p}$, exactly as in the case of a classical receive beamforming. From the Cauchy–Schwarz's inequality we have that the optimum receiver is $\mathbf{w} \propto \mathbf{H}\mathbf{p}$, i.e., when the beamvector is aligned with the effective channel $\mathbf{h} = \mathbf{H}\mathbf{p}$. The resulting SNR is then given by

$$\mathrm{SNR} = \mathbf{p}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{p}. \tag{1.21}$$

The transmit beamvector $\mathbf{p}$ that maximizes this expression is, as in the classical transmit beamforming, the eigenvector of matrix $\mathbf{H}^\dagger\mathbf{H}$ corresponding to the maximum eigenvalue (or, equivalently, to the right singular vector of the channel matrix $\mathbf{H}$ corresponding to the maximum singular value, denoted by $\mathbf{v}_{H,\max}$), properly normalized to satisfy the power constraint with equality: $\mathbf{p} = \sqrt{P_0}\,\mathbf{v}_{H,\max}$. The final achieved SNR is $P_0\,\sigma_{H,\max}^2$, where $\sigma_{H,\max}$ denotes the maximum singular value.

Now that we know that the optimal transmitter is the best right singular vector, we can step back and elaborate on the optimal receiver $\mathbf{w} \propto \mathbf{Hp} = \sqrt{P_0}\,\mathbf{H}\mathbf{v}_{H,\max} = \sqrt{P_0}\,\sigma_{H,\max}\mathbf{u}_{H,\max}$ to realize that it is actually equal (up to an arbitrary scaling factor) to the best left singular vector of the channel matrix $\mathbf{H}$.

Summarizing, the best transmit–receive beamvectors correspond nicely to the right–left singular vectors of the channel matrix $\mathbf{H}$ associated to the largest singular value and the global communication process becomes

$$\hat{x} = \mathbf{w}^\dagger (\mathbf{Hp}x + \mathbf{n}) = (\sqrt{P_0}\,\sigma_{H,\max})x + n, \tag{1.22}$$

where n is an equivalent scalar noise with zero mean and unit variance.

1.3.3 Multiple Beamforming (Matrix Beamforming) for MIMO Channels: Problem Statement

As we have seen, obtaining the best transmit–receive beamvectors when transmitting one symbol over a MIMO channel is rather simple. However, precisely one of the interesting properties of MIMO channels is the multiplexing capability they exhibit. To properly take advantage of the potential increase in rate, we need to transmit more than one symbol simultaneously. We can easily extend the previous signal model to account for the simultaneous transmission of L symbols:

$$\mathbf{s} = \sum_{i=1}^{L} \mathbf{p}_i x_i = \mathbf{Px}, \tag{1.23}$$

where $\mathbf{P}$ is a matrix with columns equal to the transmit beamvectors $\mathbf{p}_i$ corresponding to the L transmitted symbols x_i stacked for convenience in vector $\mathbf{x}$ (normalized such that $\mathbb{E}\left[\mathbf{x}\mathbf{x}^\dagger\right] = \mathbf{I}$). The power constraint in this case is

$$\sum_{i=1}^{L} \|\mathbf{p}_i\|^2 = \mathrm{Tr}\left(\mathbf{P}\mathbf{P}^\dagger\right) \le P_0. \tag{1.24}$$

Similarly, each estimated symbol at the receiver is $\hat{x}_i = \mathbf{w}_i^\dagger \mathbf{y}$ or, more compactly,

$$\hat{\mathbf{x}} = \mathbf{W}^\dagger \mathbf{y}, \tag{1.25}$$

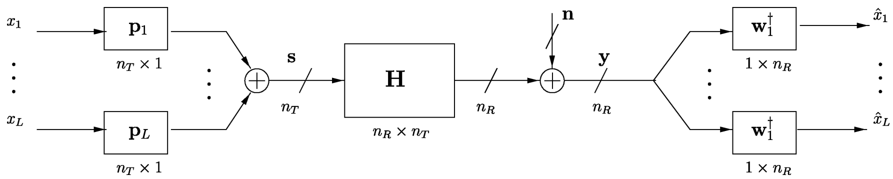

Fig. 1.5 Multiple beamforming interpretation of a MIMO communication system.

where $\mathbf{W}$ is a matrix with columns equal to the receive beamvectors $\mathbf{w}_i$ corresponding to the L transmitted symbols. We can either interpret this communication scheme as a *multiple beamforming* scheme as illustrated in Figure 1.5 or as a *matrix beamforming* scheme as shown in Figure 1.6. Both interpretations are actually natural extensions of the single beamforming scheme in Figure 1.4.

The design of the transmitter and receiver in the multiple beamforming case is fundamentally different from the single beamforming case. This happens because the L data streams are coupled and exhibit a tradeoff for two different reasons:

(i) The total power budget P_0 needs to be distributed among the different substreams.
(ii) Even for a given power allocation among the substreams, the design of the transmit "directions" is still coupled as the transmission of one symbol interferes the others, as can be seen from

$$\hat{x}_i = \mathbf{w}_i^\dagger \left(\mathbf{H}\mathbf{p}_i x_i + \mathbf{n}_i \right), \tag{1.26}$$

where $\mathbf{n}_i = \sum_{j \neq i} \mathbf{H}\mathbf{p}_j x_j + \mathbf{n}$ is the equivalent interference-plus-noise seen by the ith substream.

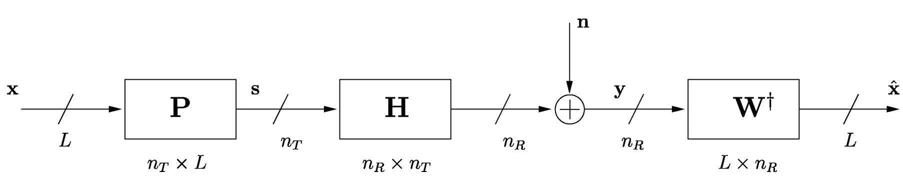

Fig. 1.6 Matrix beamforming interpretation of a MIMO communication system.

This inherent tradeoff among the substreams complicates the problem to the point that not even the problem formulation is clear: What objective should we consider to measure the system performance?

As a consequence, different authors have considered a variety of objective functions to design such systems (see the historical overview in Section 1.4). In some cases, deriving optimal solutions according to the selected objective and, in other cases, only giving suboptimal solutions due to the difficulty of the problem. It is important to mention that if we can assume the use of sufficiently long and good codes, then the problem formulation becomes rather simple as elaborated later in Section 1.4.

This text considers a general problem formulation based on an arbitrary objective function (alternatively, on individual constraints on the quality of each data stream) and develops a unified framework based on majorization theory that allows the simplification of the problem so that optimal solutions can be easily obtained.

1.3.4 Diagonal Transmission for MIMO Channels: A Heuristic Solution

Inspired by the solution in the single beamforming case, we can come up with a suboptimal strategy that simplifies the problem design a great deal. Recall that in the single beamforming scheme, the best transmit and receive beamvectors correspond to the right and left singular vectors of the channel matrix $\mathbf{H}$, respectively, associated to the largest singular value. In the multiple beamforming scheme, we can consider the natural extension and choose, for the ith substream, the right and left singular vectors of the channel matrix $\mathbf{H}$ associated to the ith largest singular value, $\mathbf{v}_{H,i}$ and $\mathbf{u}_{H,i}$, respectively:

$$\mathbf{p}_i = \sqrt{p_i}\,\mathbf{v}_{H,i} \quad \text{and} \quad \mathbf{w}_i = \mathbf{u}_{H,i}, \tag{1.27}$$

where p_i denotes the power allocated to the ith substream that must satisfy the power constraint $\sum_{i=1}^{L} p_i \leq P_0$.

With this choice of transmit–receive processing, the global communication process becomes diagonal or orthogonal (in the sense that the different substreams do not interfere with each other):

$$\hat{x}_i = \mathbf{w}_i^{\dagger} \left(\mathbf{H}\mathbf{p}_i x_i + \mathbf{n}_i\right) \tag{1.28}$$

$$= \sqrt{p_i}\,\sigma_{H,i}\,x_i + n_i \quad 1 \leq i \leq L, \tag{1.29}$$

or, more compactly,

$$\hat{\mathbf{x}} = \mathbf{W}^{\dagger} \left(\mathbf{H}\mathbf{P}\mathbf{x} + \mathbf{n}\right) \tag{1.30}$$

$$= \mathrm{diag}\left(p_1, \ldots, p_L\right) \mathbf{\Sigma}_H \mathbf{x} + \mathbf{n}, \tag{1.31}$$

where $\mathbf{\Sigma}_H$ is a diagonal matrix that contains the L largest singular values of $\mathbf{H}$ in decreasing order and $\mathbf{n}$ is an equivalent vector noise with zero mean and covariance matrix $\mathbb{E}\left[\mathbf{n}\mathbf{n}^{\dagger}\right] = \mathbf{I}$.

Since the substreams do not interfere with each other, we can nicely write the signal to interference-plus-noise ratios (SINRs) as

$$\mathrm{SINR}_i = p_i\,\sigma_{H,i}^2 \quad 1 \leq i \leq L \tag{1.32}$$

and the only remaining problem is to find the appropriate power allocation $\{p_i\}$ which will depend on the particular objective function chosen to measure the performance of the system.

Fortunately, as will be shown in this text, we do not need to content ourselves with suboptimal solutions and we can aim for the global solution.

1.4 Historical Perspective on MIMO Transceivers

The problem of jointly designing the transmit and receive signal processing is an old one. Already in the 1960s, we can easily find papers that jointly design transmit–receive filters for frequency-selective SISO channels to minimize the MSE (e.g., [11, 128] and references therein). The design of MIMO transceivers for communication systems dates back to the 1970s, where cable systems were the main application [90, 129].

The design of MIMO systems is generally quite involved since several substreams are typically established over MIMO channels (multiplexing property). Precisely, the existence of several substreams, each with its own performance, makes the definition of a global measure of the system performance not clear; as a consequence, a wide span of different design criteria has been explored in the literature as overviewed next.

At this point, it is important to emphasize that if one can assume the use of sufficiently long and good codes, i.e., if instead of a signal processing approach we adopt an information theoretic perspective, then the problem formulation simplifies drastically and the state of the art of the problem is very different. As first devised by Shannon in 1949 [140] for frequency-selective channels and rigorously formalized for a matrix channel in [21, 152, 153], the best transmission scheme that achieves the channel capacity consist of: (i) diagonalizing the channel matrix, (ii) using a waterfilling power allocation over the channel eigenmodes, and (iii) employing a Gaussian signaling (see also [33, 122, 148]). In many real systems, however, rather than with Gaussian codes, the transmission is done with practical discrete constellations and coding schemes.

In a more general setup, we can formulate the design of the MIMO system as the optimization of a global objective function based on the individual performance of each of the established substreams. Alternatively, we can consider the achievable set of individual performance of the substreams (e.g., in a CDMA system where each user has some minimum performance constraint). The classical aforementioned information-theoretic solution will be then a particular case of this more general setup.

The first linear designs (for cable systems) considered a mathematically tractable cost function as a measure of the system performance: the sum of the MSEs of all channel substreams or, equivalently, the trace of the MSE matrix [2, 90, 129, 171] (different papers explored variations of the problem formulation concerning the dimensions of the channel matrix, the channel frequency-selectivity, the excess-bandwidth, etc.). Decision-feedback schemes were also considered [128, 82, 170].

Due to the popularization of wireless multi-antenna MIMO systems in the late 1990s [43, 45, 122, 148], a new surge of interest on the design of MIMO transceivers appeared with a wireless rather than wired motivation. Different design criteria have been used by different authors as shown in the following. In [131] a unified signal model for block transmissions was presented as a MIMO system and different design criteria were considered such as the minimization of the trace of the MSE matrix and the maximization of the SINR with a zero-forcing (ZF) constraint. In [170], the minimization of the determinant of the MSE matrix was considered for decision-feedback (DF) schemes. In [130], a reverse-engineering approach was taken to obtain different known solutions as the minimization of the weighted trace of the MSE matrix with appropriate weights. In [3], the flat multi-antenna MIMO case was considered providing insights from the point of view of beamforming. Various criteria were considered in [132] under average power constraint as well as peak power constraint.

For the aforementioned design criteria, the problem is very complicated but fortunately it simplifies because the channel matrix turns out to be diagonalized by the optimal transmit–receive processing and the transmission is effectively performed on a diagonal or parallel fashion. Indeed, the diagonal transmission implies a *scalarization* of the problem (meaning that all matrix equations are substituted with scalar ones) with the consequent simplification (cf. Section 1.3.4). In light of the optimality of the diagonal structure for transmission in all the previous examples (including the capacity-achieving solution [33, 122, 148]), one might expect that the same would hold for any other criteria as well. However, as shown in [111], this is not the case.

More recently, the design of MIMO transceivers has been approached using the bit error rate (BER), rather than the MSE or the SINR, as basic performance measure. This approach is arguably more relevant as the ultimate performance of a system is measured by the (BER), but it is also more difficult to handle. In [106], the minimization of the BER (and also of the Chernoff upper bound) averaged over the channel substreams was treated in detail when a diagonal structure is imposed. The minimum BER design of a linear MIMO transceiver without the diagonal structure constraint was independently obtained in [36] and [111], resulting in an optimal *nondiagonal* structure. This result, however, only holds when the constellations used in all the substreams are equal.

In [111], a general unifying framework was developed that embraces a wide range of different design criteria for linear MIMO transceivers;

in particular, the optimal design was obtained for the family of Schur-concave and Schur-convex cost functions which arise in majorization theory [97]. Interestingly, this framework gives a clear answer to the question of whether the diagonal transmission is optimal: when the cost function is Schur-concave then the diagonal structure is optimal, but when the cost function is Schur-convex then the optimal structure is not diagonal anymore.

From the previous unifying framework based on majorization theory, it follows that the minimization of the BER averaged over the substreams, considered in [36, 111], is a Schur-convex function, provided that the constellations used on the substreams are equal, and therefore it can be optimally solved. The general case of different constellations, however, is much more involved (in such a case, the cost function is neither Schur-convex nor Schur-concave) and was solved in [110] via a primal decomposition approach, a technique borrowed from optimization theory [12, 88, 142].

An alternative way to formulate the design of MIMO transceivers is to consider an independent requirement of quality for each of the substreams rather than a global measure of quality. This was considered and optimally solved in [114], again based on majorization theory.

Interestingly, the unifying framework based on the majorization theory was later extended to nonlinear DF MIMO transceivers in [74] (see also [141]). The extension in [74] is based on a new matrix decomposition, namely, the generalized triangular decomposition [70]. While the linear transceiver design relies on the concept of *additive* majorization, the nonlinear decision-feedback transceiver invokes the *multiplicative* majorization. One can see an intriguing mathematical symmetry between the linear and nonlinear designs.

As evidenced by the previous results, majorization theory is a mathematical tool that plays a key role in transforming the originally complicated matrix-valued nonconvex problem into a simple scalar problem. Other recent successful applications of majorization theory in communication systems, from either an information-theoretic or a signal processing perspective, include the design of signature sequences in CDMA systems to maximize the sum-rate or to satisfy Quality-of-Service (QoS) requirements with minimum power by Viswanath *et al.* [161, 162, 163] and the study of the impact of correlation of the transmit antennas in MISO systems by Boche *et al.* [18, 79].

1.5 Outline

This text considers the design of point-to-point MIMO transceivers (this also includes multiuser CDMA systems) with CSI at both sides of the link according to an arbitrary cost function as a measure of the system performance. A unified framework is developed that hinges on majorization theory as a key tool to transform the originally complicated matrix-valued nonconvex problem into a simple scalar problem in most cases convex which can be addressed under the powerful framework of convex optimization theory [12, 13, 20]. The framework allows the choice of any cost function as a measure of the overall system performance and the design is based then on the minimization of the cost function subject to a power constraint or vice versa. In addition, the framework embraces the possibility of imposing a set of QoS constraints for the data streams with minimum required power.

This chapter has already given the basic background on MIMO channels and MIMO communication systems, including a natural evolution from classical beamforming to MIMO transceivers and a historical perspective on MIMO transceivers.

Chapter 2 introduces majorization theory on which the rest of the text is based.

Chapter 3 is fully devoted to *linear* MIMO transceivers composed of a linear precoder at the transmitter and a linear equalizer at the receiver. In particular, the key simplification relies on the *additive* majorization relation. Different types of design are considered in order of increasing conceptual and mathematical complexity: (i) based on a Schur-concave/convex cost function as a global measure of performance, (ii) based on individual QoS constraints, and (iii) based on an arbitrary cost function as a global measure of performance.

Then, Chapter 4 considers *nonlinear DF* MIMO transceivers, composed of a linear precoder at the transmitter and a decision-feedback equalizer (DFE) at the receiver (consisting of a feedforward stage and a feedback stage) or the dual form based on dirty paper coding by uplink–

downlink duality. Interestingly, the key simplification relies in this case on a *multiplicative* majorization relation. As in the linear case, different types of design are considered: (i) based on an arbitrary cost function as a global measure of performance (including Schur-concave/convex cost functions) and (ii) based on individual QoS constraints.

Hence, from Chapters 3 and 4, both the linear and nonlinear cases are nicely unified under an additive and multiplicative majorization relation. The basic design of point-to-point linear and nonlinear DF MIMO transceivers with CSI is thus well understood. This is not to say that the general problem of MIMO transceivers with CSI is fully solved. On the contrary, there are still many unanswered questions and future lines of research.

Chapter 5 precisely describes unanswered questions and future lines of research, namely, (i) the design of *multiuser* MIMO transceivers for networks with interfering users, (ii) the design of *robust* MIMO transceivers to imperfect CSI, (iii) the design of nonlinear MIMO transceivers with ML decoding, and (iv) the design of MIMO transceivers from an information-theoretic perspective with arbitrary constellations.

Notation. The following notation is used. Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and italics denote scalars. $\mathbb{R}^{m\times n}$ and $\mathbb{C}^{m\times n}$ represent the set of $m \times n$ matrices with real- and complex-valued entries, respectively. $\mathbb{R}_+$ and $\mathbb{R}_{++}$ stand for the set of nonnegative and positive real numbers, respectively. The super-scripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^\dagger$ denote matrix transpose, complex conjugate, and Hermitian operations, respectively. $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real and imaginary part, respectively. $\mathrm{Tr}(\cdot)$ and $\det(\cdot)$ (also $|\cdot|$) denote the trace and determinant of a matrix, respectively. $\|\mathbf{x}\|$ is the Euclidean norm of a vector $\mathbf{x}$ and $\|\mathbf{X}\|_F$ is the Frobenius norm of a matrix $\mathbf{X}$ (defined as $\|\mathbf{X}\|_F \triangleq \sqrt{\mathrm{Tr}(\mathbf{X}^\dagger\mathbf{X})}$). $[\mathbf{X}]_{i,j}$ (also $[\mathbf{X}]_{ij}$) denotes the (ith, jth) element of matrix $\mathbf{X}$. $\mathbf{d}(\mathbf{X})$ and $\boldsymbol{\lambda}(\mathbf{X})$ denote the diagonal elements and eigenvalues, respectively, of matrix $\mathbf{X}$. A block-diagonal matrix with diagonal blocks given by the set $\{\mathbf{X}_k\}$ is denoted by $\mathrm{diag}(\{\mathbf{X}_k\})$. The operator $(x)^+ \triangleq \max(0,x)$ is the projection onto the nonnegative orthant.

References

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[2] N. Amitay and J. Salz, "Linear equalization theory in digital data transmission over dually polarized fading radio channels," *At&T Bell Labs Technical Journal*, vol. 63, no. 10, pp. 2215–2259, December 1984.

[3] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2172–2178, November 2000.

[4] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. Wiley, third ed., 2003.

[5] S. L. Ariyavisitakul, "Turbo space-time processing to improve wireless channel capacity," *IEEE Transactions on Communications*, vol. 48, no. 8, pp. 1347–1358, August 2000.

[6] E. Beltrami, "Sulle funzioni bilineari," *Giornale De Matematiche*, vol. 11, pp. 98–106, 1873.

[7] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Society for Industrial and Applied Mathematics, 2001.

[8] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. New York, NY, USA: Kluwer Academic Publishers, 1999.

[9] M. Bengtsson, "A pragmatic approach to multi-user spatial multiplexing," in *Proceedings of 2nd IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM-2002)*, (Rosslyn, VA), August 4–6 2002.

[10] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, (L. C. Godara, ed.), Boca Raton, FL: CRC Press, 2001.

[11] T. Berger and D. W. Tufts, "Optimum pulse amplitude modulation. Part I: Transmitter-receiver design and bounds from information theory," *IEEE Transactions on Information Theory*, vol. IT-13, no. 2, pp. 196–208, April 1967.

[12] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, second ed., 1999.

[13] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Scientific, 2003.

[14] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[15] R. Bhatia, *Matrix Analysis*. New York: Springer-Verlag, 1997.

[16] E. Biglieri, *Coding for Wireless Channels*. New York, NY: Springer, 2005.

[17] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communication Magazine*, vol. 28, no. 5, pp. 5–14, May 1990.

[18] H. Boche and E. A. Jorswieck, "On the ergodic capacity as a function of the correlation properties in systems with multiple transmit antennas without CSI at the transmitter," *IEEE Transactions on Communications*, vol. 52, no. 10, pp. 1654–1657, October 2004.

[19] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, *A Tutorial on Geometric Programming*. Optimization and Engineering, Springer, April 2007.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[21] L. H. Brandenburg and A. D. Wyner, "Capacity of the Gaussian channel with memory: The multivariate case," *The Bell System Technical Journal*, vol. 53, no. 5, pp. 745–778, May–June 1974.

[22] A. R. Calderbank, "The art of signaling: Fifty years of coding theory," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2561–2595, October 1998.

[23] M. Chiang, "Geometric programming for communication systems," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 1–2, pp. 1–154, August 2005.

[24] K. Cho and D. Yoon, "On the general BER expression of one and two-dimensional amplitude modulations," *IEEE Transactions on Communications*, vol. 50, no. 7, pp. 1074–1080, July 2002.

[25] M. T. Chu, "Inverse eigenvalue problems," *SIAM Review*, vol. 40, no. 1, pp. 1–39 (electronic), 1998.

[26] M. T. Chu, "A fast recursive algorithm for constructing matrices with prescribed eigenvalues and singular values," *SIAM Journal of Numerical Analysis*, vol. 37, pp. 1004–1020, 2000.

[27] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: A unified view," *IEEE Transactions on Communications*, vol. 49, no. 9, pp. 1561–1571, September 2001.

[28] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, Jr., "MMSE Decision-feedback equalizers and coding. Parts I-II: Equalization results and coding results," *IEEE Transactions on Communications*, vol. 43, no. 10, pp. 2582–2604, October 1995.

[29] J. M. Cioffi and G. D. Forney, Jr., "Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise," in *Communications, Computation, Control and Signal Processing*, (A. Paulraj, V. Roychowdhury, and C. D. Schaper, eds.), ch. 4, Boston, MA, USA: Kluwer Academic Publishers, 1997.

[30] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1639–1667, June 2002.

[31] L. Collin, O. Berder, P. Rostaing, and G. Burel, "Optimal minimum distance-based precoder for MIMO spatial multiplexing systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 617–627, March 2004.

[32] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439–441, May 1983.

[33] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[34] M. O. Damen, H. El 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.

[35] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm, "Linear matrix inequality formulation of spectral mask constraints with applications to FIR filter design," *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2702–2715, November 2002.

[36] Y. Ding, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2410–2423, September 2003.

[37] H. El Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 968–985, June 2004.

[38] P. Elia, K. R. Kumar, S. A. Pawar, P. V. Kumar, and H.-F. Lu, "Explicit construction of space-time block codes achieving the Diversity-multiplexing gain tradeoff," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3869–3884, September 2006.

[39] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice-strategies for cancelling known interference," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 3820–3833, November 2005.

[40] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3417–3432, October 2005.

[41] G. D. Forney, Jr., "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," *41st Allerton Conference on Communication, Control, and Computing, Monticello, IL*, p. 2003, October 2003.

[42] G. D. Forney, Jr., "Shannon meets Wiener II: On MMSE estimation in successive decoding schemes," *Proceedings of 2004 Allerton Conference on Communication, Control, and Computing, Monticello, IL*, p. 2004, September 2004.

[43] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[44] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multiple-element arrays," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, March 1999.

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

[46] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[47] G. Ginis and J. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1085–1104, June 2002.

[48] W. Givens, "Computation of plane unitary rotations transforming a general matrix to triangular form," *Journal of SIAM*, vol. 6, no. 1, pp. 26–50, March 1958.

[49] R. H. Gohary, T. N. Davidson, and Z.-Q. Luo, "An efficient design method for vector broadcast systems with common information," in *Proceedings of IEEE 2003 Global Communications Conference (GLOBECOM-2003)*, (San Francisco, CA), December 1–5 2003.

[50] A. Goldsmith, *Wireless Communications*. New York: Cambridge University Press, 2005.

[51] G. H. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," *SIAM Journal on Numerical Analysis*, vol. 2, pp. 205–224, 1965.

[52] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins University Press, third ed., 1996.

[53] M. Grant, S. Boyd, and Y. Ye, "Disciplined convex programming," in *Global Optimization: from Theory to Implementation, Nonconvex Optimization and Its Applications*, (L. Liberti and N. Maculan, eds.), pp. 155–210, New York: Springer Science+Business Media, Inc., 2006.

[54] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Transactions on Information Theory*, vol. IT-18, no. 6, pp. 725–730, November 1972.

[55] T. Guess, "Optimal sequence for CDMA with decision-feedback receivers," *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 886–900, April 2003.

[56] W. W. Hager, *Applied Numerical Linear Algebra*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[57] R. J. Hanson and C. L. Lawson, "Extensions and applications of the householder algorithm for solving linear least squares problems," *Mathematics of Computation*, vol. 23, pp. 787–812, March 1969.

[58] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, pp. 774–780, August 1972.

[59] G. H. Hardy, J. E. Littlewood, and G. Pólya, "Some simple inequalities satisfied by convex functions," *Messenger Mathematics*, vol. 58, pp. 145–152, 1929.

[60] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. London and New York: Cambridge University Press, second ed., 1952.

[61] B. Hassibi, "A fast square-root implementation for BLAST," *Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*, pp. 1255–1259, November 2000.

[62] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Transactions on Information Theory*, vol. 34, no. 3, pp. 380–388, May 1988.

[63] M. L. Honig, K. Steiglitz, and B. Gopinath, "Multichannel signal processing for data communications in the presence of crosstalk," *IEEE Transactions on Communications*, vol. 38, no. 4, pp. 551–558, April 1990.

[64] A. Horn, "On the eigenvalues of a matrix with prescribed singular values," *Proceedings of American Mathematical Society*, vol. 5, pp. 4–7, 1954.

[65] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge: Cambridge University Press, 1991.

[66] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

[67] A. Householder, "Unitary triangularization of a nonsymmetric matrix," *Journal of ACM 6*, vol. 6, pp. 339–342, 1958.

[68] A. T. James, "Normal multivariate analysis and the orthogonal group," *The Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 40–75, March 1954.

[69] Y. Jiang, W. Hager, and J. Li, "The geometric mean decomposition," *Linear Algebra and Its Applications*, vol. 396, pp. 373–384, February 2005.

[70] Y. Jiang, W. Hager, and J. Li, "The generalized triangular decomposition," *Mathematics of Computation*, November 2006.

[71] Y. Jiang, W. Hager, and J. Li, "Tunable channel decomposition for MIMO communications using channel state information," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4405–4418, November 2006.

[72] Y. Jiang, J. Li, and W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3791–3803, October 2005.

[73] Y. Jiang, J. Li, and W. Hager, "Uniform channel decomposition for MIMO communications," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4283–4294, November 2005.

[74] Y. Jiang, D. P. Palomar, and M. K. Varanasi, "Precoder optimization for nonlinear MIMO transceiver based on arbitrary cost function," in *Proceedings of 41st IEEE Annual Conference on Information Sciences and Systems (CISS-2007)*, The John Hopkins University, Baltimore, MD. pp. 14–16, March 2007.

[75] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ, USA: Prentice Hall, 1993.

[76] G. Jöngren, M. Skoglund, and B. Ottersen, "Combining beamforming and orthogonal space-time block coding," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 611–627, March 2002.

[77] C. Jordan, "Mémoire sur les formes bilinéaires," *Journal of Mathematiques pure et Appliquees*, vol. 19, pp. 35–54, 1874.

[78] E. Jorswieck and H. Boche, "Transmission strategies for the MIMO MAC with MMSE receiver: average MSE optimization and achievable individual MSE region," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2872–2881, November 2003.

[79] E. A. Jorswieck and H. Boche, "Optimal transmission strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3440–3453, December 2004.

[80] I. Kalet, "The multitone channel," *IEEE Transactions on Communications*, vol. 37, no. 2, pp. 119–124, February 1989.

[81] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Transactions on Communications*, vol. 1, no. 1, pp. 46–55, January 2002.

[82] M. Kavehrad and J. Salz, "Cross-polarization cancellation and equalization in digital transmission over dually polarized multipath fading channels," *At&T Technical Journal*, vol. 64, no. 10, pp. 2211–2245, December 1985.

[83] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[84] P. Kosowski and A. Smoktunowicz, "On constructing unit triangular matrices with prescribed singular values," *Computing*, vol. 64, pp. 279–285, 2000.

[85] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, July 1996.

[86] P. Lancaster, *Theory of Matrices*. New York: Academic Press, 1969.

[87] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, UK: Cambridge University Press, 2003.

[88] L. S. Lasdon, *Optimization Theory for Large Systems*. New York: Macmillian, 1970.

[89] H. Lebret and S. Boyd, "Antenna array pattern synthesis via convex optimization," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 526–532, March 1997.

[90] K. H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Transactions on Communications*, vol. COM-24, no. 12, pp. 1283–1290, December 1976.

[91] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and Applications*, vol. 284, no. 1–3, pp. 193–228, July 1998.

[92] A. Lozano, A. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.

[93] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[94] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1037–1052, April 2004.

[95] W. K. Ma, T. N. Davidson, K. M. Wong, Z. Q. Luo, and P. C. Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Transactions on Signal Processing*, vol. 50, no. 4, pp. 912–922, April 2002.

[96] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: Wiley, 1999.

[97] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic Press, 1979.

[98] J. Milanovic, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Design of robust redundant precoding filter banks with zero-forcing equalizers for unknown frequency-selective channels," in *Proceedings of 2000 International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2000)*, Istanbul, Turkey. pp. 2761–2764, June 2000.

[99] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York, NY, USA: Wiley, 1980.

[100] C. T. Mullis and R. A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Transactions on Circuits and Systems*, vol. CAS-23, no. 9, pp. 551–562, September 1976.

[101] 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. 1423–1435, October 1998.

[102] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.

[103] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 969–996, November 2006.

[104] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. Vol. 13, Philadelphia, PA: SIAM, Studies in Applied Mathematics, 1994.

[105] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space-time block codes," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3903–3912, September 2006.

[106] E. N. Onggosanusi, A. M. Sayeed, and B. D. Van Veen, "Efficient signaling schemes for wideband space-time wireless channels using channel state information," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 1, pp. 1–13, January 2003.

[107] D. P. Palomar and M. Chiang, "Alternative distributed algorithms for network utility maximization: Framework and applications," *IEEE Transactions on Automatic Control*, to appear 2007.

[108] D. P. Palomar, *A Unified Framework for Communications through MIMO Channels*. PhD thesis, Technical University of Catalonia (UPC), Barcelona, Spain, May 2003.

[109] D. P. Palomar, "Convex primal decomposition for multicarrier linear MIMO transceivers," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, December 2005.

[110] D. P. Palomar, M. Bengtsson, and B. Ottersten, "Minimum BER linear transceivers for MIMO channels via primal decomposition," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2866–2882, August 2005.

[111] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, September 2003.

[112] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 686–695, February 2005.

[113] D. P. Palomar and M. A. Lagunas, "Simplified joint transmit-receive space-time equalization on spatially correlated MIMO channels: A beamforming approach," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 730–743, June 2003.

[114] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1179–1197, May 2004.

[115] A. Paulraj, R. Nabat, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, UK: Cambridge University Press, 2003.

[116] W. W. Peterson and J. E. J. Weldon, *Error-Correcting Codes*. Cambridge, Massachusetts, and London, England: The MIT Press, second ed., 1978.

[117] B. Picinbono, "On Circularity," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3473–3482, December 1994.

[118] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, August 1995.

[119] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 858–871, May 1997.

[120] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, third ed., 1995.

[121] S. U. H. Qureshi, "Adaptive Equalization," *Proceedings of the IEEE*, vol. 73, no. 9, pp. 1349—1387, September 1985.

[122] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 357–366, March 1998.

[123] F. Rey, M. Lamarca, and G. Vázquez, "Robust power allocation algorithms for MIMO OFDM systems with imperfect CSI," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 1070–1085, March 2005.

[124] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton Univ. Press, second ed., 1970.

[125] R. T. Rockafellar, "Lagrange multipliers and optimality," *SIAM Review*, vol. 35, no. 2, pp. 183–238, 1993.

[126] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multi-access space-time block coded MIMO systems: A probabilistically constrained approach," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1560–1570, August 2006.

[127] M. Rupf and J. L. Massey, "Optimal sequence multisets for synchronous code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 40, pp. 1261–1266, July 1994.

[128] J. Salz, "Optimum mean-square decision-feedback equalization," *Bell System of Techncial Journal*, vol. 52, pp. 1341–1373, October 1973.

[129] J. Salz, "Digital transmission over cross-coupled linear channels," *At&T Technical Journal*, vol. 64, no. 6, pp. 1147–1159, July–August 1985.

[130] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2198–2206, December 2001.

[131] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. Part I: Unification and optimal designs," *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 1988–2006, July 1999.

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

[133] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

[134] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 714–725, March 2003.

[135] M. Schubert and H. Boche, "QoS-based resource allocation and transceiver optimization," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 6, pp. 383–529, 2006.

[136] M. Schubert, S. Shuying, E. A. Jorswieck, and H. Boche, "Downlink sum-MSE transceiver optimization for linear multi-user MIMO systems," in *Proceedings of 39th Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA), pp. 1424–1428, October 28–November 1 2005.

[137] I. Schur, "On the characteristic roots of a linear substitution with an application to the theory of integral equations," *Mathematique Annalen*, vol. 66, pp. 488–510, 1909.

[138] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 214–226, January 2004.

[139] S. Shamai and L. Laroia, "The inter-symbol interference channel: Lower bounds on capacity and channel precoding loss," *IEEE Transactions on Information Theory*, vol. 42, no. 9, pp. 1388–1404, September 1996.

[140] C. E. Shannon, "Communication in the presence of noise," *Proceeding IRE*, vol. 37, no. 1, pp. 10–21, January 1949.
[141] M. B. Shenouda and T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding," in *Proceedings of 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'07)*, (Honolulu, Hawaii), pp. 15–20, April 2007.
[142] N. Z. Shor, *Minimization Methods for Non-Differentiable Functions*. Berlin, Germany: Springer-Verlag, 1985.
[143] T. Söderström and P. Stoica, *System Identification*. London, UK: Prentice Hall International, 1989.
[144] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*. Upper Saddle River, NJ: Prentice Hall, 1999.
[145] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11 and 12, pp. 625–653, 1999.
[146] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.
[147] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T Bell Labs, Internal Tech. Memo*, June 1995.
[148] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November–December 1999.
[149] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronic Letters*, vol. 7, pp. 138–139, March 1971.
[150] H. V. Trees, *Detection, Estimation, and Modulation Theory. Part IV: Optimum Array Processing*. New York: Wiley, 2002.
[151] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge University Press, 2005.
[152] B. S. Tsybakov, "Capacity of a vector Gaussian channel without memory," *Problems of Information Transmission*, vol. 1, pp. 26–40, 1965.
[153] B. S. Tsybakov, "Capacity of a discrete-time Gaussian channel with a filter," *Problems of Information Transmission*, vol. 6, pp. 253–256, July/September 1970.
[154] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, April 1988.
[155] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, March 1996.
[156] M. K. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," *Proceedings of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1405–1409, November 2–5 1997.

[157] S. Verdú, *Multiuser Detection*. New York, NY, USA: Cambridge University Press, 1998.

[158] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, October 2003.

[159] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, September 2001.

[160] 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, pp. 1912–1921, August 2003.

[161] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1984–1991, September 1999.

[162] P. Viswanath and V. Anantharam, "Optimal sequences for CDMA under colored noise: A schur-saddle function property," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1295–1318, June 2002.

[163] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1968–1983, September 1999.

[164] E. Viterbo and J. Boutros, "A universal lattice decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 1639–1642, July 2000.

[165] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 313–324, February 2003.

[166] H. Weyl, "Inequalities between two kinds of eigenvalues of a linear transformation," *Proceedings of the Natuaral Academic Science USA*, vol. 35, pp. 408–411, 1949.

[167] J. H. Wilkinson and C. Reinsch, "Linear Algebra," in *Handbook for Automatic Computation*, (F. L. Bauer, ed.), Berlin: Springer-Verlag, 1971.

[168] S.-P. Wu, S. Boyd, and L. Vandenberghe, "FIR filter design via semidefinite programming and spectral factorization," in *Proceedings of IEEE Conference on Decision and Control*, pp. 271–276, December 1996.

[169] F. Xu, T. N. Davidson, J. K. Zhang, and K. M. Wong, "Design of block transceivers with decision feedback detection," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 965–978, March 2006.

[170] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Transactions on Information Theory*, vol. 40, no. 5, pp. 1334–1347, September 1994.

[171] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Transactions on Communications*, vol. 42, no. 12, pp. 3221–3231, December 1994.

[172] W. Yu and J. M. Cioffi, "Sum capacity of a Gaussian vector broadcast channel," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, September 2004.

[173] W. Yu, D. P. Varodayan, and J. M. Cioffi, "Trellis and convolutional precoding for transmitter-based interference presubtraction," *IEEE Transactions on Communications*, vol. 53, no. 7, pp. 1220–1230, July 2005.

[174] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 145–152, January 2004.

[175] J.-K. Zhang, A. Kavčić, X. Ma, and K. M. Wong, "Design of Unitary Precoders for ISI Channels," in *Proceedings IEEE International Conference on Acoustics Speech and Signal Processing*, pp. 2265–2268, Orlando, Florida, 2002.

[176] J.-K. Zhang, A. Kavčić, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 154–172, January 2005.

[177] X. Zhang, D. P. Palomar, and B. Ottersten, "Robust design of linear MIMO transceivers," *submitted to IEEE Transactions on Signal Processing*, July 2006.

[178] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.