Scientific
Research

Advances in MIMO Techniques for Mobile Communications—A Survey

Farhan Khalid, Joachim Speidel
Institute of Telecommunications, University of Stuttgart, Stuttgart, Germany
Email: {*khalid, speidel*}*@inue.uni-stuttgart.de*

Abstract

This paper provides a comprehensive overview of critical developments in the field of multiple-input multiple-output (MIMO) wireless communication systems. The state of the art in single-user MIMO (SU-MIMO) and multiuser MIMO (MU-MIMO) communications is presented, highlighting the key aspects of these technologies. Both open-loop and closed-loop SU-MIMO systems are discussed in this paper with particular emphasis on the data rate maximization aspect of MIMO. A detailed review of various MU-MIMO uplink and downlink techniques then follows, clarifying the underlying concepts and emphasizing the importance of MU-MIMO in cellular communication systems. This paper also touches upon the topic of MU-MIMO capacity as well as the promising convex optimization approaches to MIMO system design.

1. Introduction

Multiple-input multiple-output (MIMO) wireless systems employ multiple transmit and receive antennas to increase the transmission data rate through spatial multiplexing or to improve system reliability in terms of bit error rate (BER) performance using space-time codes (STCs) for diversity maximization [1]. MIMO systems exploit multipath propagation to achieve these benefits, without the expense of additional bandwidth. More recent MIMO techniques like the geometric mean decomposition (GMD) technique proposed in [2] aim at combining the diversity and data rate maximization aspects of MIMO in an optimal manner. These advantages make MIMO a very attractive and promising option for future mobile communication systems especially when combined with the benefits of orthogonal frequency-division multiplexing (OFDM) [3,4].

The capacity of an $M \times N$ single-user MIMO (SU-MIMO) system with M transmit and N receive antennas, in terms of the spectral efficiency i.e. bits per second per Hz, is given by [1]

$$C = \log_2\left[\det\left(\mathbf{I}_N + \frac{\rho}{M}\mathbf{H}\mathbf{H}^H\right)\right] \qquad (1)$$

where $\mathbf{H}$ is the $N \times M$ MIMO channel matrix and ρ is the signal to noise ratio (SNR) at any receive antenna. Equation (1) assumes that the M information sources are uncorrelated and have equal power. Expressed in terms of the eigenvalues, Equation (1) can be written as [1]

$$C = \sum_{i=1}^{m} \log_2\left(1 + \frac{\rho}{M}\lambda_i\right) \qquad (2)$$

where λ_i represent the nonzero eigenvalues of $\mathbf{H}\mathbf{H}^H$ or $\mathbf{H}^H\mathbf{H}$ for $N \leq M$ and $M < N$ respectively and $m = \min(M, N)$. Therefore, MIMO systems are capable of achieving several-fold increase in system capacity as compared to single-input single-output (SISO) systems by transmitting on the spatial eigenmodes of the MIMO channel.

Equation (2) also shows that the performance of MIMO systems is dependent on the channel eigenvalues. Very low eigenvalues indicate weak transmission channels which may make it difficult to recover the information from the received signals. Optimal power allocation based on the water-filling algorithm can be used to maximize the system capacity subject to a total transmit power constraint. Water-filling provides substantial capacity gain when the eigenvalue spread, i.e., the condition number $\lambda_{max}/\lambda_{min}$ is sufficiently large.

The MIMO concept becomes even more attractive in multiuser scenarios where the network capacity can be increased by simultaneously accommodating several users

without the expense of valuable frequency resources.

This paper is arranged as follows: Section 2 provides an overview of the current wireless standards which support MIMO technologies. Sections 3 and 4 include detailed discussion and performance analysis of various important SU-MIMO and multiuser MIMO (MU-MIMO) techniques respectively that are proposed for the next generation wireless communication systems. In-depth description of several MU-MIMO uplink and downlink schemes is given in Section 4 followed by a brief discussion of the MU-MIMO capacity. Section 5 provides an overview of convex optimization which has become an important tool for designing optimal MIMO beamforming systems. Section 6 concludes this work and identifies the areas for future research.

2. Current Implementation Status

There has been a lot of research on MIMO systems and techniques. MIMO-OFDM WLAN products based on the IEEE 802.11n standard are already available. The IEEE 802.16 wireless MAN standard known as WiMAX also includes MIMO features. Fixed WiMAX services are being offered by operators worldwide. Mobile WiMAX networks based on 802.16e are also being deployed while 802.16m is under development. IEEE 802.20 mobile broadband wireless access (MBWA) standard is also being formulated which will have complete support for mobility including high-speed mobile users e.g., on train networks. For other applications like cellular mobile communications which supports both voice and data traffic, MIMO systems are yet to be deployed. However, the 3GPP's long term evolution (LTE) is under development and adopts MIMO-OFDM, orthogonal frequency-division multiple access (OFDMA) and single-carrier frequency-division multiple access (SC-FDMA) transmission schemes. The following text presents a more detailed discussion of the various technical aspects of these standards and technologies.

2.1. IEEE 802.11n Wi-Fi

The IEEE 802.11n WLAN standard incorporates MIMO-OFDM as a compulsory feature to enhance data rate. Initial target was to achieve data rates in excess of 100 Mb/s [5]. However, current WLAN devices based on 802.11n Draft 2.0 are capable of achieving throughput up to 300 Mb/s utilizing two spatial streams in a 40 MHz channel in the 5 GHz band [6].

Initially, there were two main proposals one form the WWiSE consortium and the other from the TGnSync consortium competing for adoption by the IEEE 802.11 TGn. However, another proposal by the Enhanced Wireless Consortium (EWC) was finally accepted as the first draft for IEEE 802.11n [7].

The IEEE 802.11n standard proposes the use of the legacy 20 MHz channel and also an optional 40 MHz channel. The available modulation schemes include BPSK, QPSK, 16-QAM and 64-QAM [5,6]. Convolutional coding with different code rates is specified and use of low-density parity-check (LDPC) codes is also supported [5,8]. The MIMO techniques adopted include both spatial multiplexing and diversity techniques. Open-loop MIMO (OL-MIMO) techniques which do not require channel state information (CSI) at the transmitter seem to have been preferred [9]. Non-iterative linear minimum mean square error (LMMSE) detection has primarily been considered so as to minimize the complexity associated with MIMO detection while ensuring reasonably good performance [10].

Spatial spreading mentioned in [11] is an open-loop MIMO spatial multiplexing technique where multiple data streams are transmitted such that the diversity is maximized for each of the streams. The MIMO diversity techniques introduced in the standard include space-time block coding (STBC) and cyclic shift diversity (CSD) which extend the range and reception of 802.11n devices. In addition, conventional receiver spatial diversity techniques like maximum ratio combining (MRC) are also specified. Transmit beamforming is also specified as an optional feature [6]. The Cisco Aironet 1250 series access point based on 802.11n draft 2.0 supports open-loop transmit beamforming [12].

802.11n draft 2.0 specifies a maximum of 4 spatial streams per channel. Thus, a maximum throughput of 600 Mb/s can be achieved by using 4 spatial streams in a 40 MHz channel. In addition to spatial multiplexing and doubled channel bandwidth, more efficient OFDM with shorter guard interval (GI) and new medium access control (MAC) layer enhancements (e.g. closed-loop rate adaptation [13]) have also contributed to the increased throughput of 802.11n [6].

2.2. IEEE 802.16 WiMAX

The IEEE 802.16 worldwide interoperability for microwave access (WiMAX) is a recently developed wireless MAN standard that employs MIMO spatial multiplexing and diversity techniques. In addition to fixed WiMAX, the IEEE 802.16e Mobile WiMAX standard has also been developed and was approved in December 2005 [14]. Fixed WiMAX networks have already been deployed around the world and Mobile WiMAX deployments have also started.

802.16e-2005 is basically an amendment to the 802.16-2004 standard for fixed WiMAX with addition of new features to support mobility. 802.16e specifies the 2–6 GHz frequency band for mobile applications and the 2–11 GHz band for fixed applications (The single-carrier WirelessMAN-SC PHY specification for fixed wireless

access however specifies the 10–66 GHz frequency band [15].). It also specifies a license-exempt band between 5–6 GHz. A cellular network structure is specified with support for handoffs and mobile users moving at vehicular speeds are also supported, thus enabling mobile wireless internet access [14,16].

In addition to single carrier transmission, the standard specifies OFDM transmission scheme with 128, 256, 512, 1024 or 2048 subcarriers. Both TDD and FDD duplexing is specified while the multiplexing/multiple access schemes include OFDMA in addition to burst TDM/TDMA. However, scalable OFDMA is specified in all mobile WiMAX profiles as the physical layer multiple access technique. The various channel bandwidths specified in the standard include 1.25, 1.75, 3.5, 5, 7, 10, 8.75, 10, 14 and 15 MHz. WiMAX supports adaptive modulation and coding schemes. The supported modulation schemes include BPSK, QPSK, 16-QAM and 64-QAM [14,16]. Optional 256-QAM support is provided in the WirelessMAN-SCa PHY [15]. Convolutional codes at rate1/2, 2/3, 3/4 or 5/6 are specified as mandatory for both uplink and downlink. In addition, convolutional turbo codes, repetition codes, LDPC and concatenated Reed-Solomon convolutional code (RS-CC) are specified as optional. The supported data rates range from 1 Mb/s to 75 Mb/s [14,16].

IEEE 802.16e supports both open-loop and closed-loop MIMO. Open-loop MIMO techniques include spatial multiplexing (SM) and space-time coding (STC) [14,17,18]. 802.16e includes support for up to four spatial streams and therefore a maximum of 4×4 MIMO configuration [14,18]. STC is based on the Alamouti scheme (also STBC) and is also called space-time transmit diversity (STTD). It is an optional feature and may be used to provide higher order transmit diversity on the downlink [14].

In closed-loop MIMO, full or partial CSI is available at the transmitter through feedback. Eigenvector steering is employed to approach full capacity of the MIMO channel and water filling can be used to maximize throughput by allocating power in an optimal manner [9,19]. IEEE 802.16e supports closed-loop MIMO precoding for SM and also closed-loop STC [14,17]. However, closed-loop MIMO is not yet supported in the latest WiMAX Forum Wave 2 profiles [18]. Another MIMO mode called "collaborative spatial multiplexing" is also specified where two subscriber stations (SS), each having a single antenna, use the same subchannel for uplink transmission in order to increase the throughput [14,15, 17,20].

The adaptive antenna systems (AAS) supported in 802.16e also include closed-loop adaptive beamforming, which uses feedback from the SS to the base station (BS) to optimize the downlink transmission [14,15,18].

IEEE 802.16 Task Group m (TGm) has also been set up to develop the IEEE 802.16m standard which will enable interoperability between WiMAX and 3GPP's Long Term Evolution (LTE) standard for next generation mobile communications [21,22]. 802.16m is expected to support high-speed mobile wireless access (up to 350 km/h) and peak data rates of over 300 Mb/s using 4×4 MIMO [22].

2.3. IEEE 802.20 MBWA

The IEEE 802.20 working group was established to draft the IEEE 802.20 Mobile Broadband Wireless Access (MBWA) standard which is also nicknamed as MobileFi. IEEE 802.20 proposes a complete cellular structure and is designed and optimized for mobile data services at speeds up to 250 km/h. However, it can also support voice services due to very low transmission latency of 10–30 ms (better than the 25–40 ms for 802.16e). User data rates in excess of 1 Mb/s can be supported at 250 km/h [23–25].

MBWA is designed to operate in the licensed bands below 3.5 GHz [24,25]. 2.5 MHz to 20 MHz of uplink/downlink transmission bandwidth can be allocated per cell [25]. For a bandwidth of 5 MHz, peak aggregate data rate of around 16 Mb/s can be supported in the downlink [23,24] which obviously would be much greater for higher bandwidths.

The transmission scheme is based on OFDM, with OFDMA used for downlink transmission while both OFDMA and code-division multiple access (CDMA) are specified for the uplink. Rotational OFDM is specified as an optional scheme. The standard supports both FDD and TDD operation. The supported modulation schemes include QPSK, 8-PSK, 16-QAM and 64-QAM. Support of hierarchical (layered) modulation involving the superposition of two modulation schemes is also included for broadcast and multicast services. The specified FEC coding schemes include convolutional codes, turbo codes and LDPC codes [25].

Various MIMO schemes are also supported. STTD (based on STBC) and SM are specified for SU-MIMO transmission, utilizing up to 4 transmit antennas. STTD is particularly important for high speed mobile access. Two different stream multiplexing schemes namely single codeword (SCW) and multiple codeword (MCW) may be employed for MIMO transmission. These schemes also support closed-loop MIMO downlink transmission with rank adaptation. Both schemes utilize linear precoding at the BS for transmit beamforming based on the feedback of a suitable precoding matrix from the user equipment's (UE's) codebook to the BS. The standard also supports MU-MIMO or space-division multiple access (SDMA) transmission in the downlink which involves multiuser scheduling and precoding at the BS depending upon the feedback of the preferred precoding matrix index and differential channel quality indicator

(CQI) reports from the UEs [25].

The IEEE 802.20 standard was supposed to be available in 2006 but was delayed due to lack of support from some of the key vendors and the political turmoil within the standards forum [23]. However, it was finally approved in June 2008 and made available by the end of August 2008 [25].

2.4. 3GPP LTE

The 3rd generation partnership project's (3GPP) long term evolution (LTE) project is aimed at developing a new mobile communications standard for gradual migration from 3G to 4G. LTE physical layer is almost near completion. It specifies an OFDM based system with support for MIMO. Downlink transmission is based on OFDMA while SC-FDMA is used for the uplink due to its low PAPR characteristics. It supports both TDD and FDD operation. A packet switching architecture is specified for LTE [26,27].

LTE supports scalable bandwidths of 1.25, 2.5, 5, 10 and 20 MHz. Peak data rates of 100 Mb/s and 50 Mb/s are supported in the downlink and the uplink respectively, in 20 MHz channel. The standard specifies full performance within a cell up to 5 km radius and slight degradation from 5–30 km. Operation up to 100 km may be possible. It also supports high-speed mobility with high performance at speeds up to 120 km/h while the E-UTRAN (Evolved Universal Terrestrial Radio Access Network i.e., LTE's RAN) should be able to maintain the connection up to 350 km/h, or even up to 500 km/h. LTE also specifies very low latency operation with control plane (C-plane) latency of < 50-100ms and user plane (U-plane) latency of < 10 ms [27,28].

The single-user MIMO techniques supported include STBC and SM. Closed-loop multiple codeword (MCW) SM with codebook based precoding and with support for cyclic delay diversity (CDD) is specified. A maximum of two downlink spatial streams are specified. LTE also supports MU-MIMO in the downlink as well as in the uplink. Closed-loop transmit diversity using MIMO beamforming with rank adaptation is also supported. The supported antenna configurations for the downlink include 4 × 2, 2 × 2, 1 × 2 and 1 × 1 whereas 1 × 2 and 1 × 1 configurations are supported in the uplink [27,29,30]. However, multiple UE antennas in the uplink may be supported in future.

3. Single-User MIMO Techniques

Various open-loop and closed-loop SU-MIMO techniques are discussed in the following text along with performance analysis and comparison. Some of the techniques mentioned herein have already been adopted for the current standards while other advanced methods are

likely candidates for the next generation wireless systems.

3.1. V-BLAST

The vertical Bell Laboratories Layered Space-Time (V-BLAST) [31] is one of the very first open-loop spatial multiplexing MIMO systems which has been practically demonstrated to achieve much higher spectral efficiencies than SISO systems, in rich scattering environments. In V-BLAST, a single data stream is demultiplexed into multiple substreams which are mapped on to symbols and then transmitted through multiple antennas. Inter-substream coding is not employed in V-BLAST, however channel coding can be applied to the individual substreams for reduction of bit error rate (BER). CSI in a V-BLAST system is available at the receiver only by means of channel estimation. **Figure 1** shows the simple block diagram of a V-BLAST system.

V-BLAST detection can be accomplished by using linear detectors like zero-forcing (ZF) or minimum mean square error (MMSE) detector along with symbol cancellation (also called successive interference cancellation). Symbol cancellation is a nonlinear technique which enhances the detection performance by subtracting the detected components of the transmit vector from the received symbol vector [31]. This technique, however, is prone to error propagation.

The QR decomposition of the MIMO channel matrix **H** can be used to represent the ZF nulling in V-BLAST [2]. Assuming a frequency-flat fading MIMO channel, the corresponding sampled baseband received signal for a V-BLAST system with M transmit and N receive antennas ($M \leq N$) is therefore given by

$$\begin{aligned} \mathbf{y} &= \mathbf{Hx} + \mathbf{n} \\ &= \mathbf{QRx} + \mathbf{n} \end{aligned} \tag{3}$$

where **Q** is an $N \times M$ unitary matrix with orthonormal columns, **R** is a $M \times M$ upper triangular matrix, **x** is the transmitted signal and **n** represents the noise vector. The
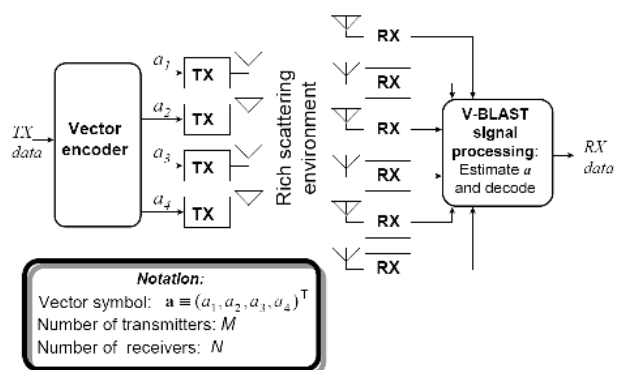

Figure 1. V-BLAST system block diagram [31].

discrete-time index is dropped to simplify notation. Multiplying both sides of Equation (3) by $\mathbf{Q}^H$ gives

$$\tilde{\mathbf{y}} = \mathbf{R}\mathbf{x} + \tilde{\mathbf{n}} \qquad (4)$$

The sequential signal detection in V-BLAST can be accomplished as follows [2]:
for $i = M : -1 : 1$

$$\hat{x}_i = C\left[\left(\tilde{y}_i - \sum_{j=i+1}^{M} r_{ij}\hat{x}_j\right) / r_{ii}\right]$$

end

where C represents mapping to the nearest modulation symbol.

The results for an initial V-BLAST prototype mentioned in [31] yielded spectral efficiencies of 20–40 bps/Hz in indoor scenarios which is quite impressive. However, later has shown that V-BLAST also performs reasonably well in mobile scenarios and can be employed for MIMO-OFDM systems as well and further improvements have been suggested in the literature. [32] proposes an extension of V-BLAST incorporating power and rate feedback which approaches closed-loop MIMO capacities. Equal power allocation with per-antenna rate control (PARC) produces the best results for the proposed system. PARC enables the transmitter to select the appropriate data rate and the associated modulation and coding scheme (MCS) for each transmit antenna based on the feedback of channel quality information from the receiver [33].

It presents a comparison between a modified V-BLAST system with limited feedback (including the modulation index and the number of streams to be used) and closed-loop MIMO (CL-MIMO) in [34]. CL-MIMO shows 15.1% throughput improvement for Rayleigh fading channel, 48.1% for spatially correlated channel and 104% for the case of a realistic channel model, at SNR of 25 dB. **Figure 2** shows these results.
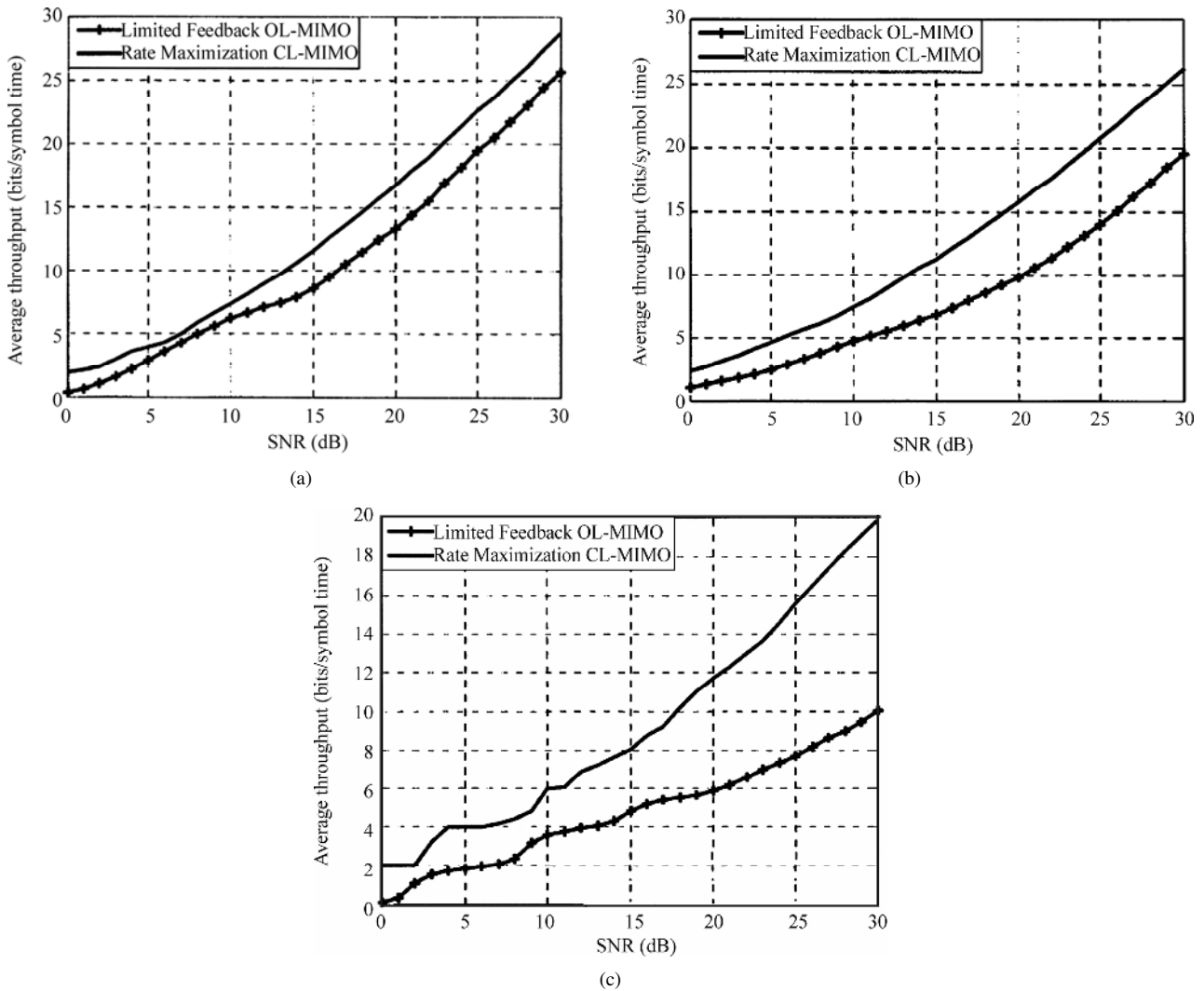
Figure 2. Throughput for (a) Rayleigh fading channel, (b) Spatial correlation channel and (c) Realistic channel model [34].

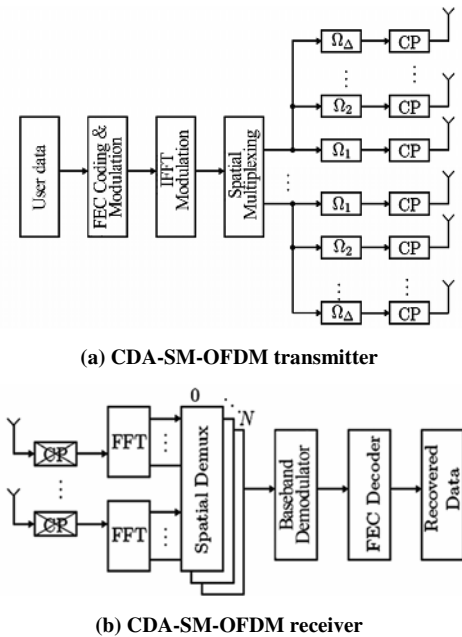## 3.2. Spatial Multiplexing with Cyclic Delay Diversity

Spatial multiplexing (SM) can be combined with a simple diversity technique such as cyclic delay diversity (CDD) to obtain much better performance as compared to regular SM systems like V-BLAST. Such a system which combines SM and MIMO diversity is referred to as a joint diversity and multiplexing (JDM) system [35]. SM with CDD is also specified in the 3GPP LTE standard [30].

It proposes a cyclic delay assisted SM-OFDM (CDA-SM-OFDM) system which does not require any CSI at the transmitter, however complete CSI is required at the receiver [35]. **Figure 3** shows the transmitter and receiver block diagram.

The blocks denoted $\Omega_1, \Omega_2, \ldots, \Omega_\Delta$ perform the cyclic delay operation which involves cyclic shifting of the signal within each group of Δ transmit antennas per SM branch. If there are P SM branches then the total number of transmit antennas is $P\Delta$. The receiver for CDA-SM-OFDM system is similar to V-BLAST.

CDD increases the channel frequency-selectivity since cyclic shifting of the OFDM signal and then adding those shifted signals linearly at the receiver inserts virtual echoes on the channel response. The resulting higher order frequency diversity can be exploited by any coded OFDM (COFDM) system [35].

Figure 4 shows a comparison of the CDA-SM-OFDM system capacity with 2×2 and 4×2 SM-OFDM systems. Here it can be seen that the capacity of the CDA-SM-
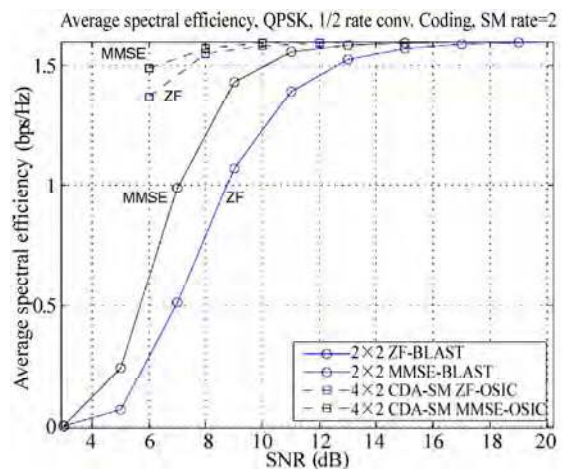
OFDM system lies between that for the two SM-OFDM systems. However, the capacities for the SM-OFDM systems are plotted for the ideal case i.e. with the best possible STC and channel coding schemes. It can also be seen that the outage capacity i.e. the capacity obtained below 10% of the times, for the CDA-SM-OFDM system is much higher than the 2×2 SM-OFDM and closer to the 4×2 SM-OFDM. Thus the system performance for the CDA-SM-OFDM system shows a significant increase just by employing a simple STC i.e. CDD.

It has also be shown in [35] that the eigenvalue spread for the CDA-SM-OFDM system is generally higher than both the SM-OFDM schemes and this means that eigen-beamfoming can be employed for CDD based SM systems. In fact, the 3GPP LTE standard incorporates CDD based SM with precoding and specifies precoding matrices for small and large delay CDD [30].

Figure 5 provides a comparison of the average spectral
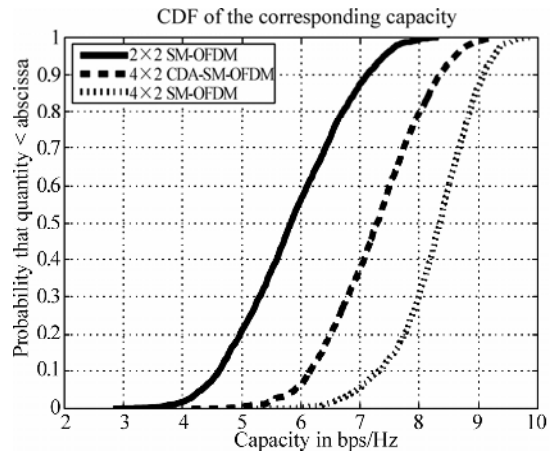
Figure 4. Comparison of system capacity for 4×2 CDA-SM-OFDM system with 2×2 and 4×2 SM-OFDM system [35].

(a) CDA-SM-OFDM transmitter

(b) CDA-SM-OFDM receiver

Figure 3. CDA-SM-OFDM system transmitter and receiver [35].

Figure 5. Average spectral efficiencies in bps/Hz [35].

efficiencies of 2 × 2 SM-OFDM systems and 4 × 2 CDA-SM-OFDM systems for a low user mobility indoor WLAN scenario. Here it can be seen that the 4 × 2 CDA-SM-OFDM systems provide much higher spectral efficiencies at low SNR values.

3.3. Singular Value Decomposition Based MIMO Precoding

Singular value decomposition (SVD) based MIMO precoding is a closed-loop MIMO scheme where the precoding filter at the transmitter is designed by taking the SVD of the MIMO channel matrix $\mathbf{H}$.

[36] provides an analysis of the classical SVD based MIMO precoding scheme, SVD based precoding with ZF equalization, SVD based precoding with MMSE equalization and also an improved SVD based precoding technique. All of these schemes are analyzed with realistic channel knowledge at the transmitter. **Figure 6** shows the block diagram of the SVD based MIMO-OFDM transmitter and receiver.

3.3.1. Classical SVD Precoding and Equalization
In SVD based techniques, the channel matrix $\mathbf{H}$ of a MIMO system with N_t transmit antennas and N_r receive antennas, is decomposed as

$$\mathbf{H} = \mathbf{UDV}^H \qquad (5)$$

where $\mathbf{U} \in \mathbb{C}^{N_r \times N_t}$ and $\mathbf{V} \in \mathbb{C}^{N_t \times N_t}$ are unitary matrices while $\mathbf{D} \in \mathbb{R}^{N_t \times N_t}$ is a diagonal matrix consisting of the ordered singular values d_i.

The classical SVD approach utilizes matrix $\mathbf{V}$ for precoding at the transmitter. The columns v_i of matrix $\mathbf{V}$ are the eigenvectors of $\mathbf{HH}^H$. The received signal is given by

$$\mathbf{r} = \mathbf{HVs} + \mathbf{n} \qquad (6)$$

where $\mathbf{s}$ is a vector of information symbols s_i and $\mathbf{n}$ is the noise vector corresponding to an additive white Gaussian noise (AWGN) process with variance σ_n^2 for each element. At the receiver, matrix $\mathbf{U}^H$ is employed for equalization and the detected signal vector is given by

$$\begin{aligned} \mathbf{y} = \mathbf{U}^H \mathbf{r} &= \mathbf{U}^H \mathbf{HVs} + \mathbf{U}^H \mathbf{n} \\ &= \mathbf{U}^H \mathbf{UDV}^H \mathbf{Vs} + \mathbf{U}^H \mathbf{n} \\ \mathbf{y} &= \mathbf{Ds} + \mathbf{U}^H \mathbf{n} \end{aligned} \qquad (7)$$

Each individual received signal can be written as

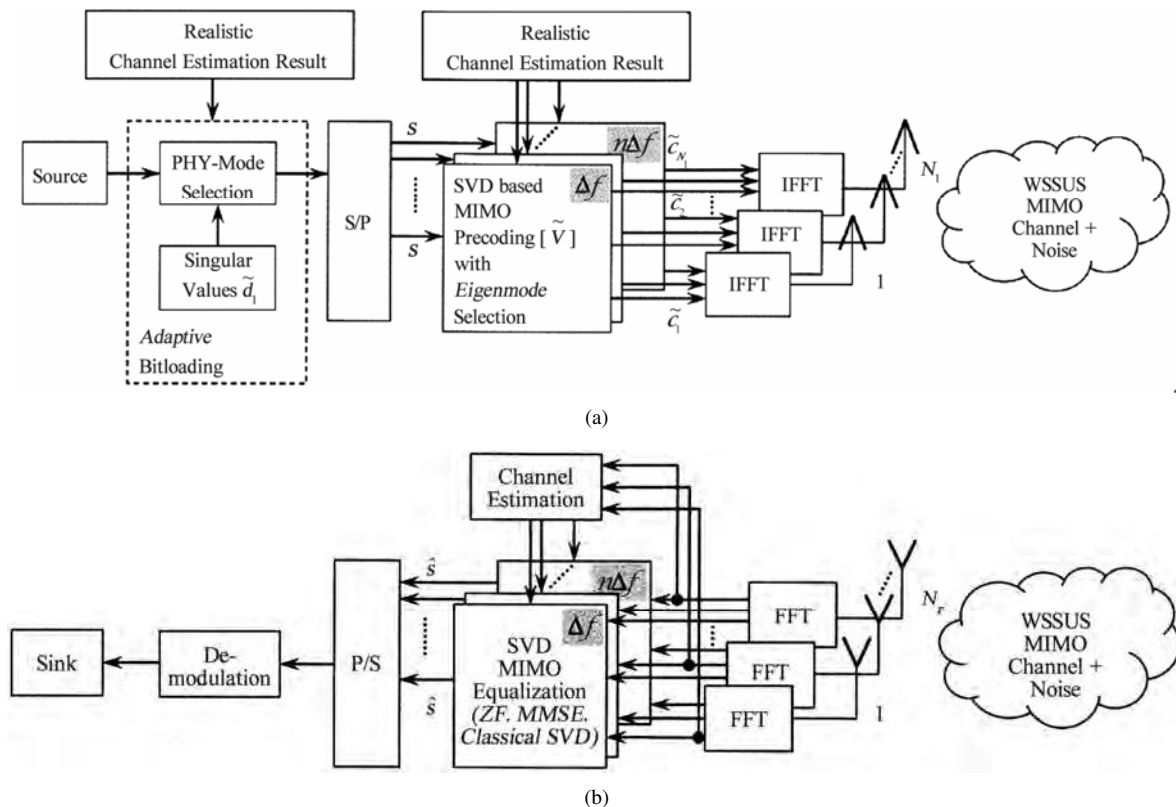$$y_i = d_i s_i + n_i' \qquad (8)$$

(a)

(b)

Figure 6. SVD based MIMO-OFDM system (a) Transmitter and (b) Receiver [36].

where $n_i^{'}$ represents the i-th element of $\mathbf{U}^H\mathbf{n}$. The corresponding SISO SNR values are then given by

$$SNR_i = d_i^2 \frac{|s_i|^2}{2\sigma_n^2} \qquad (9)$$

Equations (8) and (9) show that the singular values represent the MIMO processing gain for each of the eigenmodes. Therefore, SVD based MIMO precoding requires adaptive modulation and bit loading techniques for capacity maximization [36].

3.3.2. SVD Precoding with ZF Equalization

Linear ZF equalization can also be used at the receiver which is based on the inversion of the estimated MIMO channel matrix. ZF equalization requires the estimation of the product of $\mathbf{HV}$ at the receiver. Assuming ideal channel knowledge at the receiver, the detected signal can then be given by

$$\begin{aligned} \mathbf{y} &= (\mathbf{HV})^+ \mathbf{r} \\ &= (\mathbf{HV})^+ \mathbf{HVs} + (\mathbf{HV})^+ \mathbf{n} \end{aligned} \qquad (10)$$

$$\mathbf{y} = \mathbf{s} + (\mathbf{UD})^{-1}\mathbf{n} \qquad (11)$$

where $(\cdot)^+$ represents the pseudo inverse.

3.3.3. SVD Precoding with MMSE Equalization

MMSE equalization is based on minimizing the mean square error (MSE) between the transmitted and detected symbols. The minimum mean square error is given by

$$e = \min\left(\left\{|\hat{s}_i - s_i|^2\right\}\right) \qquad (12)$$

where s_i is the transmitted symbol and $\hat{s}_i$ represents the received symbol. The detected signal for SVD based MIMO MMSE equalization is given by

$$\mathbf{y} = \left[K\sigma_n^2\mathbf{I} + (\mathbf{HV})^H(\mathbf{HV})\right]^{-1}(\mathbf{HV})^H\mathbf{r} \qquad (13)$$

with $K = N_t$ utilized eigenmodes. As seen from Equation (13), MMSE MIMO equalization also requires the precoding matrix $\mathbf{V}$ at the receiver.

3.3.4. Improved SVD Precoding Technique with Realistic Channel Knowledge

An improved SVD based MIMO precoding technique is also proposed in [36] which maximizes MIMO capacity while considering realistic channel knowledge at the transmitter rather than the ideal one. The MIMO capacity for realistic channel knowledge is given by

$$C_H = \sum_{i=1}^{N_t} \log_2\left(1 + \tilde{d}_i^2 \frac{|s_i|^2}{2\sigma_n^2}\right) \qquad (14)$$

where $\tilde{d}_i$ represent the singular values for the case of

realistic channel knowledge. The improved technique considers the K strongest eigenmodes for transmission with $K < N_t$ if the following statement is fulfilled.

$$\sum_{i=1}^{N_t} \log_2\left(1 + \tilde{d}_i^2 \frac{|s_i|^2}{2\sigma_n^2}\right) \leq \sum_{i=1}^{K} \log_2\left(1 + \frac{N_t}{K}\tilde{d}_i^2 \frac{|s_i|^2}{2\sigma_n^2}\right) \qquad (15)$$

The remaining eigenmodes which correspond to the $N_t - K$ unused eigenvectors of the precoding matrix $\tilde{\mathbf{V}}$ are not utilized.

3.3.5. Performance Comparison

Figure 7 shows the BER performance comparison of the classical SVD, ZF and MMSE equalization schemes for a 4×4 MIMO-OFDM system. A curve for ideal ZF equalization is also provided for reference. It is clear from the comparison that MMSE equalization provides the best results with realistic channel estimation.

Figure 8 shows the performance comparison of the
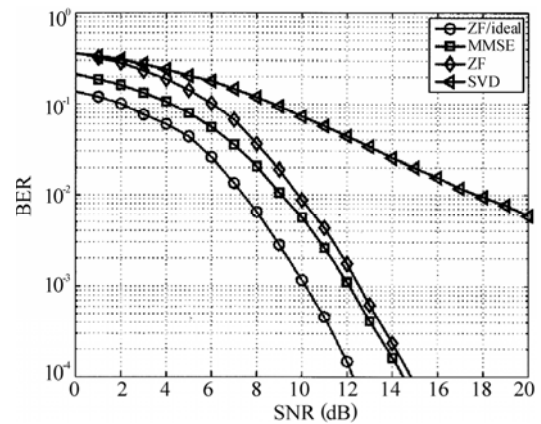

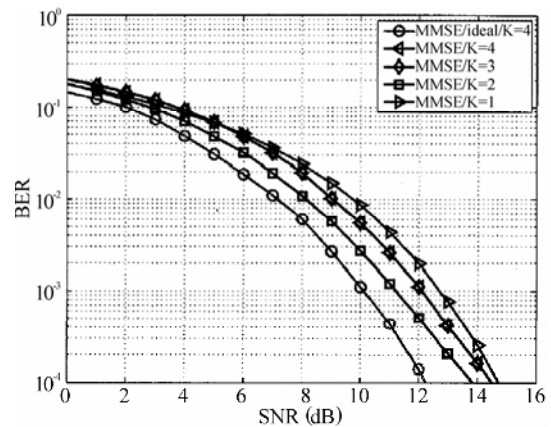**Figure 7. BER performance of uncoded 4×4 SVD based MIMO systems [36].**

Figure 8. BER performance of an uncoded 4×4 SVD based MIMO system with MMSE equalization [36].

MMSE equalization scheme with different values of K (utilized eigenmodes) for an uncoded 4×4 MIMO system with realistic channel knowledge at the transmitter. The BER curve for the case of ideal channel knowledge is also provided. The MMSE equalization scheme provides the best performance for $K = 2$ utilized eigenmodes selected according to Equation (15).

3.4. Geometric Mean Decomposition (GMD) Based MIMO

GMD based MIMO [2] is also a closed-loop joint transceiver design scheme which aims at optimally combining the benefits of MIMO diversity and spatial multiplexing. This technique utilizes the GMD of the MIMO channel matrix for precoder and equalizer design when the CSI is available at both the transmitter and the receiver. It is also applicable to MIMO-OFDM systems.

GMD calculation algorithm in [2] starts from the SVD of the channel matrix $\mathbf{H}$ which is given according to Equation (5)

$$\mathbf{H} = \mathbf{UDV}^H$$

The GMD is then given by

$$\begin{aligned} \mathbf{H} &= \mathbf{UU}_R^H \mathbf{RV}_R \mathbf{V}^H \\ &\triangleq \mathbf{QRP}^H \end{aligned} \tag{16}$$

where $\mathbf{Q}$ and $\mathbf{P}$ are semi-unitary matrices, $\mathbf{P}$ being the linear precoder at the transmitter. $\mathbf{R} \in \mathbb{R}^{K \times K}$ is an upper triangular matrix whose diagonal elements are the geometric mean of the K nonzero singular values of $\mathbf{H}$. The GMD scheme thus decomposes the MIMO channel into identical parallel subchannels which makes the symbol constellation selection and the overall system design much simpler. GMD can also be seen as an extended QR decomposition.

GMD MIMO can be implemented with the V-BLAST receiver and also with the zero-forcing dirty paper precoder (ZFDP). The V-BLAST technique has been discussed earlier in the text. The ZFDP technique also involves sequential nulling and cancellation but at the transmitter and utilizes CSI at the transmitter only.

The ZFDP scheme combines QR decomposition and "dirty paper" precoding. The QR decomposition for ZFDP is given by

$$\mathbf{H}^H = \breve{\mathbf{Q}} \breve{\mathbf{R}} \tag{17}$$

The sampled baseband received signal is then given by

$$\mathbf{y} = \breve{\mathbf{R}}^H \breve{\mathbf{Q}}^H \mathbf{x} + \mathbf{n} \tag{18}$$

Substituting $\mathbf{x} = \breve{\mathbf{Q}}\breve{\mathbf{x}}$ we have

$$\mathbf{y} = \breve{\mathbf{R}}^H \breve{\mathbf{x}} + \mathbf{n} \tag{19}$$

Let $\mathbf{s} \in \mathbb{C}^{K \times 1}$ be the transmitted symbol vector then

$\breve{\mathbf{x}}$ should satisfy

$$\mathrm{diag}\left\{\breve{\mathbf{R}}\right\}\mathbf{s} = \breve{\mathbf{R}}^H \breve{\mathbf{x}} \tag{20}$$

where the left-hand side represents the element-wise multiplication of the diagonal elements of $\breve{\mathbf{R}}$ with the elements of $\mathbf{s}$. The solution to Equation (20) is then

$$\breve{\mathbf{x}} = \left(\breve{\mathbf{R}}^H\right)^{-1} \mathrm{diag}\left\{\breve{\mathbf{R}}\right\}\mathbf{s} \tag{21}$$

The ZFDP scheme, unlike V-BLAST, does not suffer from the error propagation problem. However, due to the matrix inversion in Equation (21) the norm of $\breve{\mathbf{x}}$ can be significantly amplified resulting in increased transmitter power consumption. This problem can be resolved by using the Tomlinson-Harashima precoder to restrict the transmit signal level within acceptable limits [2].

3.4.1. Combining GMD with V-BLAST and ZFDP

The GMD-VBLAST scheme can be implemented beginning with the GMD of the channel matrix, $\mathbf{H} = \mathbf{QRP}^H$. The information symbol vector $\mathbf{s}$ is then encoded by the linear precoder $\mathbf{P}$ resulting in the transmit signal $\mathbf{x} = \mathbf{Ps}$. The resulting signal at the receiver is then given by

$$\mathbf{y} = \mathbf{QRs} + \mathbf{n} \tag{22}$$

which can be decoded simply by using the V-BLAST receiver. GMD-ZFDP scheme can be also be implemented in a similar way. The resulting K independent and identical subchannels are given by

$$y_i = \bar{\lambda}_H x_i + n_i \quad ; i = 1, \ldots, K \tag{23}$$

where $\bar{\lambda}_H$ represent the subchannel gain and are in fact the identical diagonal elements of the matrix $\mathbf{R}$ [2].

3.4.2. Performance

Some simulation results from [2] depicting the performance of GMD based MIMO schemes are presented in the following text, assuming independent identically distributed (i.i.d) Rayleigh flat fading channels. **Figure 9** shows a comparison of the capacity of GMD-MIMO with other schemes for 4×4 MIMO configuration. The informed transmitter (IT) curve corresponds to the Shannon channel capacity when CSI is available at both the transmitter and the receiver while the uninformed transmitter (UT) curve corresponds to the channel capacity when CSI is not available at the transmitter. MTM and MMD are both linear precoder design schemes for linear transceivers. MTM is based on the minimization of the trace of the MSE matrix while MMD minimizes the maximum diagonal elements of the MSE matrix resulting in near-optimal performance. Clearly, GMD outperforms both MTM and MMD at high SNR and approaches optimal capacity. The capacity loss of GMD at low SNR is due to

the ZF receiver. Based on GMD, the authors of [2] have also proposed another scheme called uniform channel decomposition (UCD) which can decompose a MIMO channel into identical subchannels in a strictly capacity lossless manner [37].

Figures 10 and 11 show the BER performance comparison of GMD-MIMO with ordered MMSE-VBLAST, MTM and MMD for 2 × 4 and 4 × 4 MIMO configurations respectively. GMD achieves much higher performance particularly at high SNR.

Figure 12 shows a performance comparison of GMD-VLBAST and GMD-ZFDP when combined with OFDM for ISI suppression. GMD-VBLAST results in performance loss of about 2 dB because of error propagation.

3.5. Turbo-MIMO Systems

Turbo-MIMO systems represent a class of MIMO com-
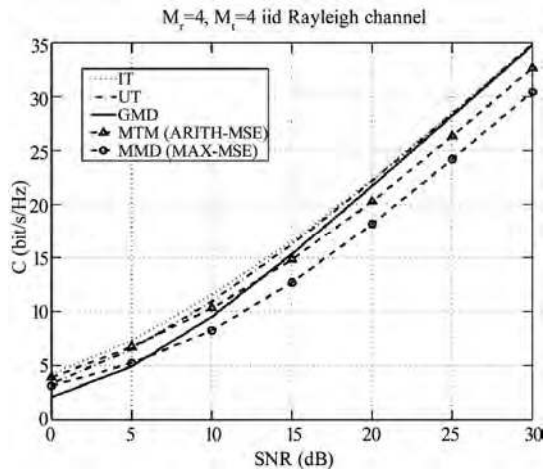
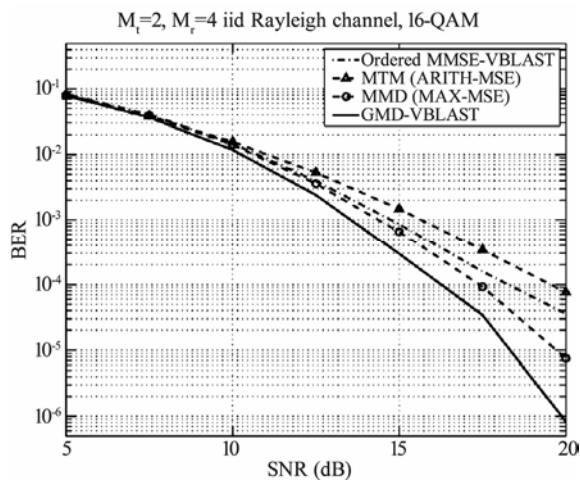Figure 9. Average capacity for 4 × 4 MIMO configuration [2].

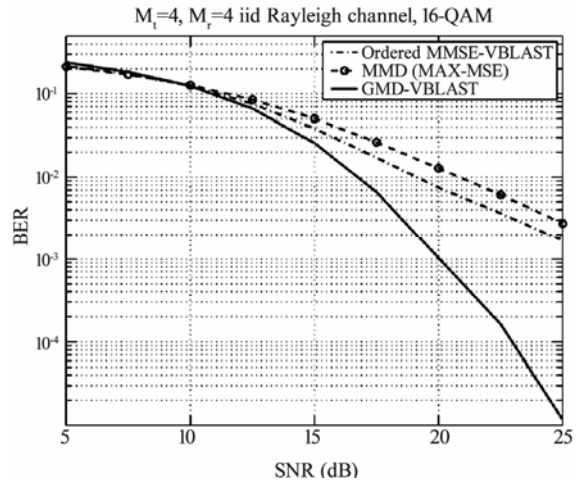Figure 10. BER performance for 2 × 4 MIMO configuration [2].

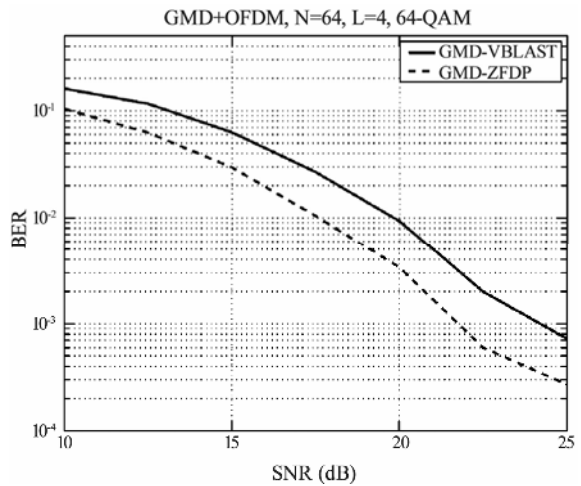Figure 11. BER performance for 4 × 4 MIMO configuration [2].

Figure 12. BER performance of GMD based MIMO-OFDM systems [2].

munication systems that combine the turbo-processing principle used in turbo coding with MIMO. These systems aim at attaining channel capacity close to the Shannon limit for MIMO channels with manageable complexity and can be implemented from diversity maximization or SM aspects [38].

A turbo-MIMO architecture known as TurboBLAST is presented in [39]. This MIMO system is based on random layered space-time (RLST) coding which is a combination of independent block-time coding and space-time interleaving. The receiver uses iterative turbo-processing for RLST decoding and estimation of the flat fading MIMO channel matrix. A similar turbo-MIMO system based on space-time bit-interleaved coded modulation (ST-BICM) is presented in [38]. ST-BICM codes are formed by concatenation of a turbo encoded sequence and ST interleaving.

Figure 13 shows the block diagram of a ST-BICM MIMO system transmitter. The information bits are turbo encoded based on a linear forward error correction (FEC) code represented as the outer code. The encoded sequence is then bit-interleaved using a space-time pseudo-random interleaver denoted by Π in the figure. Each interleaved substream is then independently mapped onto *M*-ary PSK or QAM symbols and transmitted using a separate antenna. The inner code basically represents a linear space-time mapper which allows for a flexible MIMO design with optimal diversity order and multiplexing gain or a desired tradeoff between the two. STBCs can be used to obtain the maximum diversity order while a symbol multiplexer can be used if full multiplexing gain is desired [38].

Figure 14 shows a double iterative decoding receiver for the ST-BICM MIMO system. It operates in two stages consisting of inner and outer iterative decoding loops. The inner and outer decoders are separated by an interleaver and a deinterleaver represented by Π and Π^{-1}

respectively. This arrangement decorrelates the correlated outputs between the two stages. The decorrelator compensates for the interleaving operation at the transmitter. The two stages iteratively exchange information, producing a better estimate of the transmitted symbols after each iteration, until the receiver converges [38].

The inner decoder is in fact a MIMO detector, the optimal choice being the maximum a posteriori probability (MAP or APP) detector/decoder. However, due to the excessive computational complexity of APP detection, reduced-complexity near-optimal detectors like MMSE-SIC or reduced-complexity APP detectors e.g. the list-sphere detector (LSD), iterative tree search (ITS) and multilevel bit mapping ITS (MLM-ITS) detectors can be used.

The outer decoder consists of a channel turbo decoder with two decoding stages separated by an interleaver and a deinterleaver denoted by α and α⁻¹ respectively in **Figure 14**. This arrangement forms the outer iterative decoding loop of the ST-BICM MIMO receiver [38].
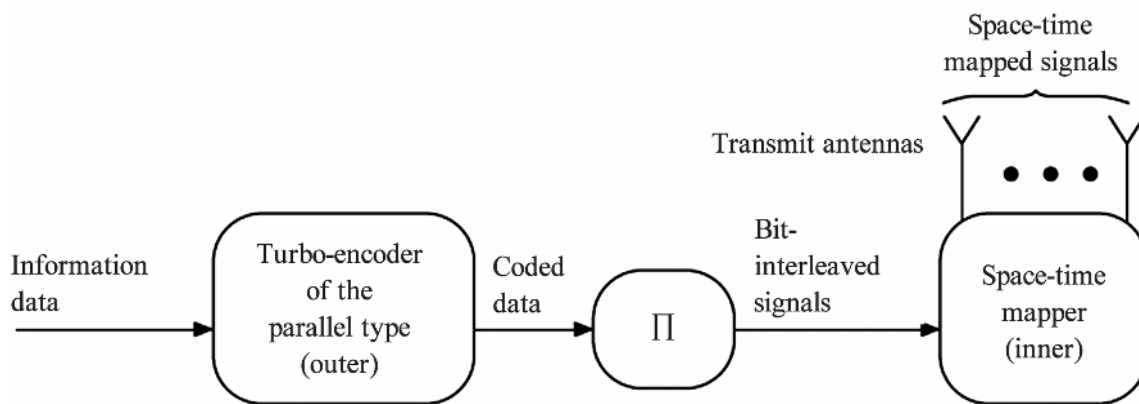
Figure 13. ST-BICM MIMO system transmitter [38].
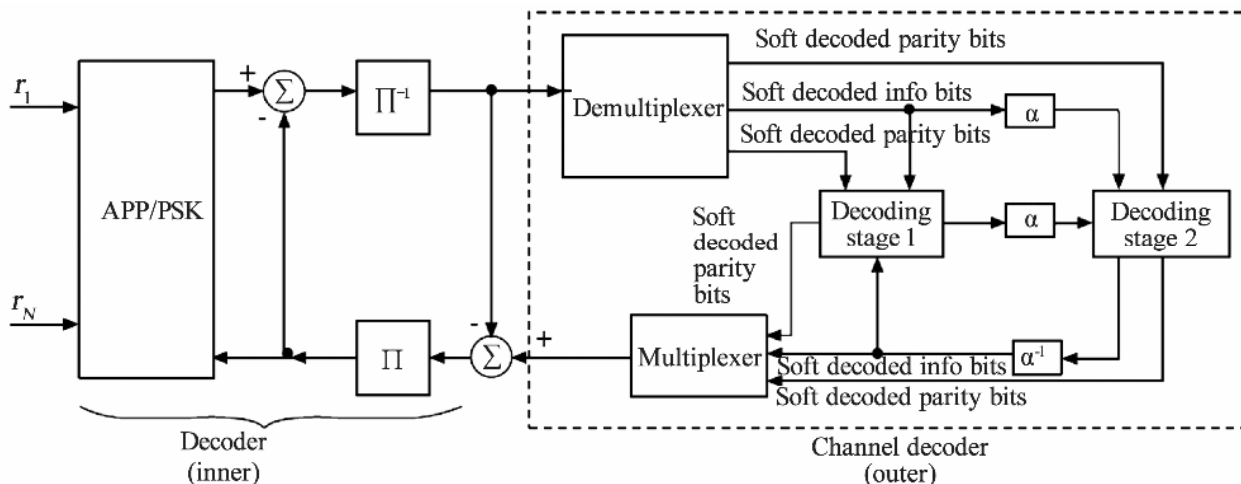
Figure 14. Receiver structure for the ST-BICM MIMO system [38].

3.5.1. Performance

Figure 15 shows the BER performance of a simulated 8 × 8 ST-BICM MIMO system using a rate-1/2, memory 2 turbo code as the outer channel code, with feed-forward and feedback generators 5 and 7 (octal) respectively. A block fading channel is assumed for the inner encoder which remains constant for a block size of 192 information bits with each block representing a statistically independent channel realization. A rich scattering Rayleigh MIMO model is used to select the elements of the MIMO channel matrix. 4 iterations are used in the inner decoder loop while 8 iterations are used in the outer channel decoder loop. The figure shows a comparison for different modulation schemes with MMSE-SIC and MLM-ITS inner detectors. The performance of MLM-ITS detection increases with larger list size M however, at the cost of increased complexity. The respective capacity limits for QPSK, 16-QAM and 64-QAM are also shown. At BER = 10^{-5} and $M = 64$, The ST-BICM systems using QPSK, 16-QAM and 64-QAM operate 1, 4 and 6 dB away from their respective capacity limits [38].

Figure 16 shows the BER performance of the ST-BICM system using MLM-ITS detector as the no. of iterations in the outer decoder increase from 1 to 5. Clearly the performance improves with the no. of iterations which pertains to only a linear increase in complexity. However, it can also be seen that the performance gain between successive iterations diminishes somewhat due to the feedback of correlated noise. Further increase in iterative gain can be achieved by using larger interleavers [38].

3.6. Limited Feedback Strategies for Closed-loop MIMO Systems

Certain closed-loop MIMO systems like the SVD and GMD based systems assume the availability of full CSI at the transmitter. Full CSI is available at the transmitter in a TDD system with duplex time less than the channel coherence time due to the reciprocity of the channel while in a FDD system a feedback channel for CSI is required thus consuming additional bandwidth. However, in practical scenarios the extra load resulting from large

CSI feedback is not desirable and may not even be possible e.g. in case of rapidly varying mobile channels. Furthermore, results have shown that performance close to that with full CSI can be achieved by using limited feedback strategies utilizing only a few bits of feedback. **Figure 17** shows the block diagram of a limited feedback MIMO system [40].
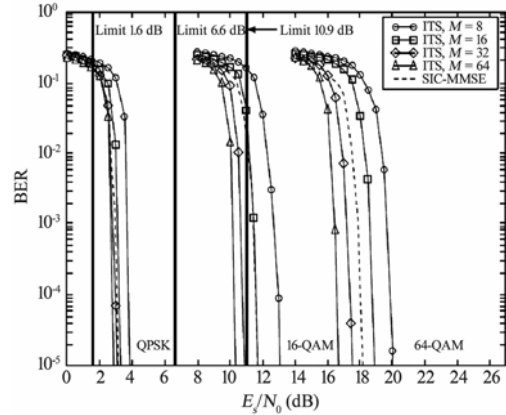
Figure 15. BER performance of 8 × 8 ST-BICM MIMO system with different modulation and inner detection schemes [38].
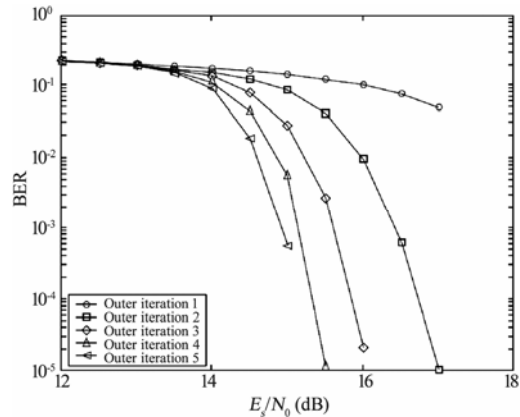
Figure 16. BER performance of 8 × 8 ST-BICM MIMO system with different no. of iteration in the outer decoder [38].
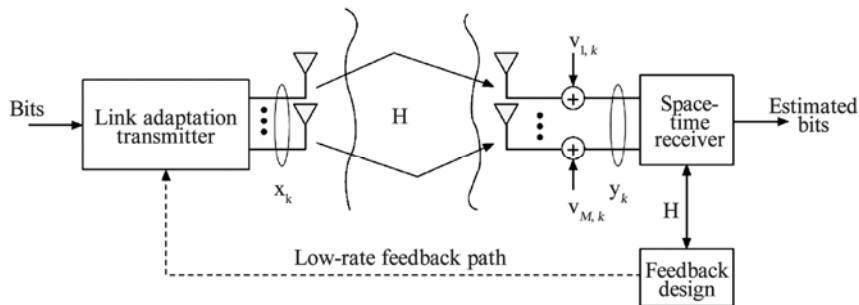

**Figure 17. Limited feedback closed-loop MIMO system [40].**

The feedback may be based on channel quantization or quantization of some properties of the transmitted signal. Channel quantization involves vector quantization (VQ) of the channel matrix **H** as depicted in **Figure 18**. The quantized version of the MIMO channel can then be fed back to the transmitter. However, it has been observed that quantization of the entire channel may not be necessary and it may be sufficient to include only some part of the channel structure like the channel singular vectors. For example, the optimal precoding matrix for an i.i.d MIMO channel consists of the eigenvectors of the channel covariance matrix, as columns. The feedback overhead may be further reduced by using only a limited number of quantized weighting vectors or matrices for precoding. This collection of precoding matrices is known as a precoding codebook and is shared by the transmitter and the receiver. The feedback consists of bits representing a particular precoding matrix within the codebook [40,41].

A large codebook length for vector quantization schemes results in increased complexity at the receiver due to the exhaustive search required for selecting a precoding matrix. In such cases, when the codebook length and therefore the corresponding no. of feedback bits B is large, scalar quantization of the elements of the precoding matrix can be employed instead. However, for small values of B, scalar quantization may become too inaccurate. In such cases, performance of scalar quantization can be improved by using the *reduced rank* approach where the columns of the precoding matrix are constrained to lie within a subspace of dimension less than the no. of transmit antennas N_t [40].

Figure 19 shows the symbol error rate (SER) performance of a simulated 4×5 limited feedback MIMO beamformer with different feedback strategies. The system uses 16-QAM modulation for transmission and MRC at the receiver. Optimal BF in the figure represents the optimal beamformer with unquantized feeback and full CSI at the transmitter. Grassmannian BF (6-bit) represents signal adaptive beamforming using a 6-bit feedback VQ codebook and results in the best performance, lying within 0.7 dB of the optimal BF and approximately 1dB better than the 40-bit channel quantization which suffers from large quantization error. The 6-bit quantized reduced rank (RR) beamformer with dimension $D = 3$, performs close to the 40-bit channel quantization [40].
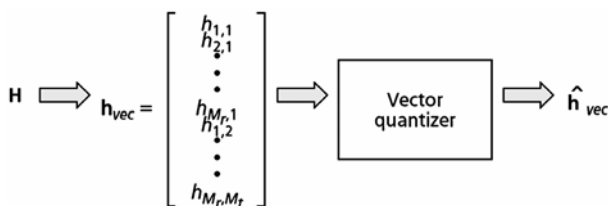
Figure 18. Channel Quantization [40].
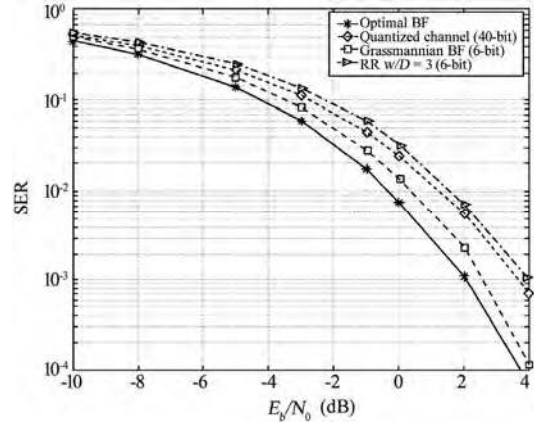
Figure 19. Limited feedback beamformer performance for 4×5 MIMO configuration [40].

3.6.1. Link Adaptation without Precoding

In addition to the precoding matrix, other information e.g. the received signal to interference and noise ratio (SINR) may also be included in the feedback for link adaptation. However, some MIMO schemes like the modified V-BLAST schemes in [32,34] rely solely on this type of feedback without any precoding information.

Another example is the 2-codeword multiple codewords (2CW-MCW) scheme for FDD MIMO-OFDM cellular systems proposed in [41] that uses SINR feedback for each stream to select a suitable modulation and coding scheme (MCS) for each of the two simultaneously transmitted codewords. The two codewords are mapped onto 2 and 4 streams respectively for 2×2 and 4×4 antenna configurations. The mapping may either be fixed or adaptive. Adaptive mapping also makes use of the SINR feedback. Precoding is not used in this scheme resulting in reduced feedback overhead.

3.6.2. Partial Feedback Schemes

Partial feedback schemes for MIMO systems are based on the feedback of statistical channel information along with some instantaneous channel quality indicator (CQI) e.g. SNR, SINR etc. to the transmitter. A partial feedback scheme for MIMO-OFDM systems involving the decomposition of MIMO channel covariance matrix is presented in [42].

The covariance matrix **R** is calculated from the estimated MIMO channel matrix **H** (for the k-th subcarrier) at the receiver and is given by

$$\mathbf{R} = \mathrm{E}\left\{ \mathbf{H}^H \mathbf{H} \right\} \qquad (24)$$

The matrix **R** is then decomposed using SVD which is given by

$$\mathbf{R} = \mathbf{U}_{Stat} \mathbf{\Lambda}_{Stat} \mathbf{V}_{Stat}^H \qquad (25)$$

where $\mathbf{\Lambda}_{Stat}$ is a diagonal matrix containing the singular

values while $\mathbf{U}_{Stat}$ and $\mathbf{V}_{Stat}$ are unitary matrices. The feedback includes the matrix $\mathbf{\Lambda}_{Stat}$ and the column vectors of $\mathbf{V}_{Stat}$ for power allocation and spatial processing (precoding) at the transmitter [42].

Figure 20 shows the block diagram of a $N_T \times N_R$ MIMO-OFDM system based on this partial feedback scheme which transmits N_S spatial streams using N_C OFDM subcarriers. The received signal vector for the k-th subcarrier is given by

$$\mathbf{y} = \mathbf{HQAx} \qquad (26)$$

where $\mathbf{x}$ is the transmit data vector, $\mathbf{H}$ is the MIMO channel matrix for the k-th subcarrier, $\mathbf{A}$ is a diagonal matrix with N_S diagonal elements determined by the matrix $\mathbf{\Lambda}_{Stat}$ feedback for power allocation to the active spatial streams, and the $N_T \times N_S$ matrix $\mathbf{Q}$ represents a spatial processing transformation which maps the spatial streams to the transmit antennas. The $\mathbf{Q}$ matrix is constructed from the vectors of $\mathbf{V}_{Stat}$ received at the transmitter via feedback and is used to maximize the received energy for each transmitted spatial stream. This enables

maximum ratio transmission (MRT) and SVD beamforming along with tracking of spatial variations of the MIMO channel [42].

The MIMO channel covariance and the corresponding channel singular values do not vary rapidly with time even at vehicular speeds around 100 km/h [42]. This greatly reduces the feedback load on the system and makes this closed-loop MIMO-OFDM system suitable for mobile environments.

Figures 21 and **22** show the simulated frame error rate (FER) performance of the proposed MIMO-OFDM system in comparison with open-loop SM and perfect CSI feedback MIMO-OFDM systems, for 2×2 and 4×4 MIMO configurations respectively. The figures include FER performance curves for QPSK and 64-QAM modulation in a low speed mobile scenario using the ITU PB channel profile with vehicular speeds of 3 km/h. The OFDM scheme is based on 512-point FFT with 15 subchannels for data transmission each consisting of 20 continuous subcarriers, for a total bandwidth of 5 MHz. The frame duration is about 0.5ms. Turbo coding is employed for FEC and MMSE detection is used at the receiver.
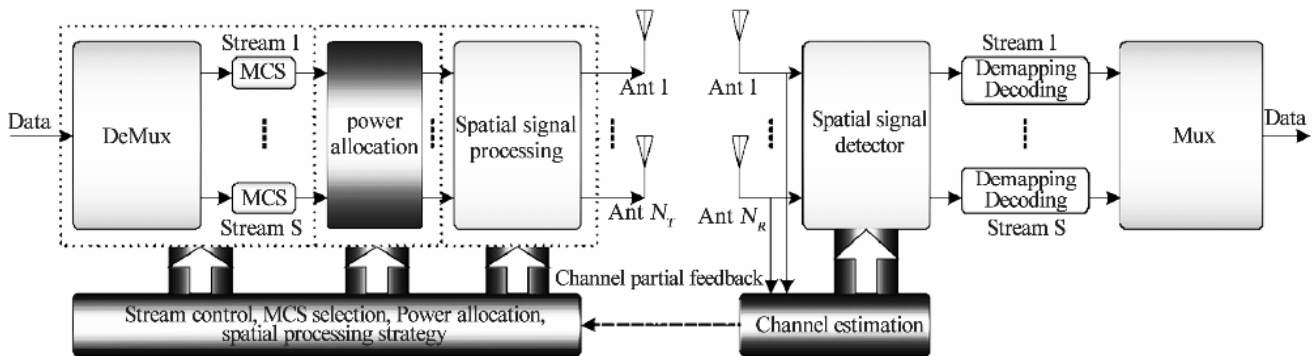

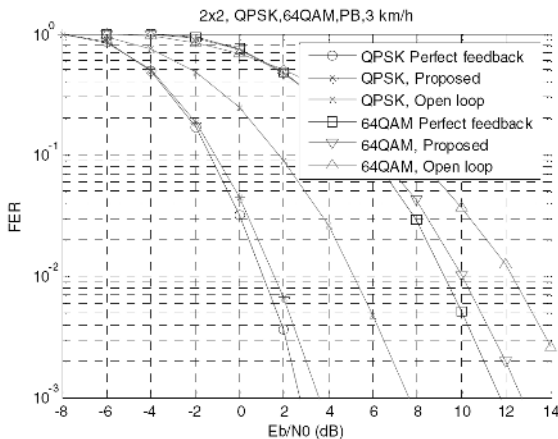**Figure 20. MIMO-OFDM system with partial feedback [42].**

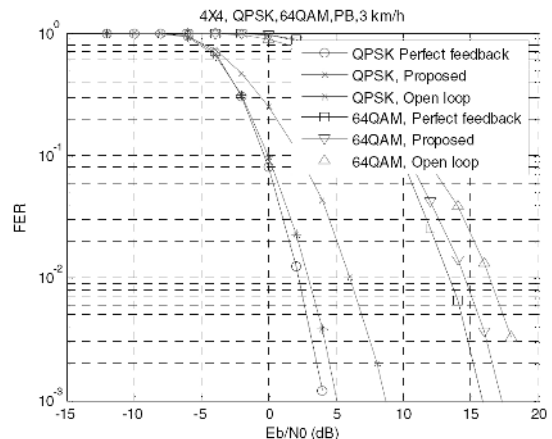**Figure 21. FER performance for coded 2×2 MIMO configuration [42].**

Figure 22. FER performance for coded 4×4 MIMO configuration [42].

Equal power allocation is used for the open-loop SM system while water-filling is used for the closed-loop systems. Ideal channel knowledge is assumed at the receiver for all systems and antenna correlations are not considered [42].

As seen from the results, the proposed system operates quite close to the perfect CSI feedback system and shows substantial performance gain over the open-loop system. The small performance loss in comparison with the perfect feedback system is primarily due to the quantization error associated with limited feedback [42].

Efficient feedback reconstruction algorithms for improvement of the closed-loop transmit diversity scheme constituting the mode 1 of 3GPP's wideband code-division multiple access (WCDMA) 3G standard are presented in [43]. These algorithms efficiently reconstruct the beamforming weights at the transmitter while considering the effect of feedback error. Performance results for vehicular speeds up to 100 km/h are provided. The proposed techniques are applicable to closed-loop MIMO diversity systems and may possibly be extended to 4G systems.

In [44], the optimal MIMO precoder designs for frequency-flat and frequency-selective fading channels are presented, assuming partial CSI at the transmitter consisting of transmit and receive correlation matrices. The elements of transmit and receive correlation matrices are determined from the respective transmit and receive antenna spacing and angular spread. It is shown that from the capacity maximization perspective, the optimal precoder for a frequency-flat fading channel is an eigen-beamformer. On the other hand, the optimal precoder for a frequency-selective fading channel represented by L uncorrelated effective paths consists of $P + L$ parallel eigen-beamformers where P is an arbitrary value depending on the no. of vectors in a transmission data block [44].

A closed-loop limited feedback MIMO scheme called multi-beam MIMO (MB-MIMO) is proposed in [45] for 3GPP LTE E-UTRA downlink. MB-MIMO employs multiple fixed beams at the base station (Node B) to transmit multiple data streams. The no. of beams and data streams to be used are adaptively selected using a codebook at the UE. The selected precoding vectors or beam indices constituting a precoding matrix are then fed back to the Node B. The MB-MIMO scheme can adaptively switch between MIMO SM and transmit beamforming (Tx-BF) modes. Tx-BF is used if a single beam is selected and SM is used if multiple beams are selected.

The proposed scheme eliminates the need for a hardware calibrator (HW-CAL) at Node B that was required for a previously proposed MB-MIMO implementation. HW-CAL compensates the phase variations caused by RF components and was needed to align the phase condition of each transmit antenna element for maximizing the transmit beamforming gain. The proposed scheme

uses a larger codebook based on an extended precoding matrix which includes phase terms that can be controlled to align the phase condition of the 4 node B antenna elements. This results in high beamforming gain even without HW-CAL. However, 4 additional bits or a total of 8 bits are required for feedback, which is still a small number.

3.7. MIMO over High-Speed Mobile Channels

Open-loop MIMO diversity techniques like STC and space frequency coding (SFC) are appropriate choices for high-speed mobile channels that vary rapidly with time. In such scenarios, maintaining a reliable link becomes the foremost priority rather than maximizing system throughput.

High-speed mobile channels undergo fast fading which may cause time variation of the fading channel within an OFDM symbol period. This results in the loss of subchannel orthogonality and leads to interchannel interference (ICI) due to the distribution of leakage signals over other OFDM subcarriers. The error floor associated with ICI increases with the speed of the mobile terminal [46].

An improved MIMO-OFDM technique for high-speed mobile access in cellular environments is proposed in [46]. This technique reduces ICI and provides diversity gain as well as noise averaging even for highly correlated channels. ICI is reduced by transmitting weighted data on adjacent subcarriers. The weights are selected such that the mean ICI power is minimized. The adopted weight selection procedure however results in suboptimal weights. Diversity gain in [46] is achieved by using space-frequency block coding (SFBC) which is based on Alamouti code but the coding is applied in frequency domain i.e. to OFDM subcarriers rather than to OFDM symbols in time domain [47]. **Figure 23** shows the data assignment scheme for the 2 × 1 SFBC-OFDM system, without the weighting factors. The transmit data is assigned to subcarrier groups each consisting of two adjacent subcarriers, as shown in the figure.

Instead of SFBC, other diversity techniques such as STBC, space-frequency trellis coding (SFTC), maximal-
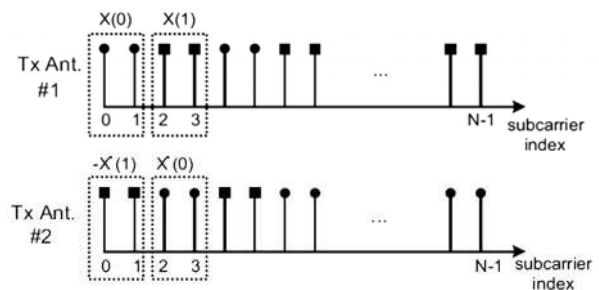

**Figure 23. Data assignment scheme for ICI reduction in 2 × 1 SFBC-OFDM system [46].**

ratio receive combining (MRRC) etc. can also be used. The proposed technique operates without CSI at the transmitter and does not require any pilot signals for channel tracking. However, it is suitable for OFDM systems with subcarrier group spacing less than the channel coherence bandwidth because the channel coefficients are assumed to be identical for adjacent subcarriers.

Figure 24 shows the simulated BER performance of the proposed SFBC-OFDM scheme in comparison with conventional SFBC MIMO-OFDM schemes for 2×1 antenna configuration using I-METRA MIMO channel model Case A for downlink transmission with mobile speed of 250 km/h. Case A corresponds to a frequency-flat Rayleigh fading channel with uncorrelated antennas. 25 MHz of downlink channel bandwidth is used at 2 GHz with 2048 OFDM subcarriers. Performance results for QPSK and 16-QAM are provided.

Figure 25 shows the performance comparison using
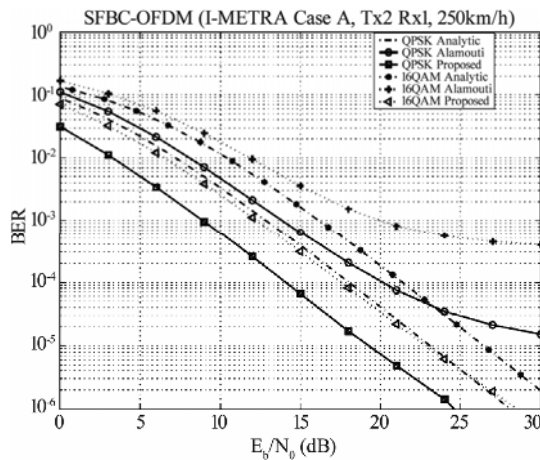
Figure 24. BER Performance of SFBC-OFDM systems using I-METRA Case A channel [46].
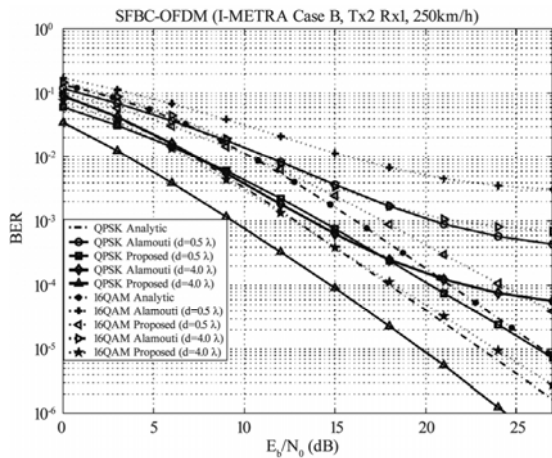
Figure 25. BER Performance of SFBC-OFDM systems using I-METRA Case B channel [46].

I-METRA Case B which corresponds to a frequency-selective fading channel with correlated transmit antennas in an urban macro cellular environment.

The proposed SFBC-OFDM scheme clearly outperforms the conventional SFBC-OFDM schemes in both cases. The conventional SFBC-OFDM scheme referred to as Alamouti in the figures is severely performance limited due to the error floor phenomenon resulting from ICI introduced by the high-speed mobile user at 250 km/h.

4. Multiuser MIMO

Multiuser MIMO (MU-MIMO) systems consist of multiple antennas at the BS and a single or multiple antennas at each UE. MU-MIMO enables space-division multiple access (SDMA) in cellular systems which increases the system capacity by exploiting the spatial dimension (i.e. the location of UEs) to accommodate more users within a cell. It also provides beamforming or array gain as well as diversity gain due to the use of multiple antennas. In case of multiple antennas at the UE, spatial multiplexing can also be employed to further enhance the spectral efficiency [48].

The uplink and the downlink of a MU-MIMO system represent two different problems which are discussed in the following text.

4.1. The MU-MIMO Uplink

The MU-MIMO uplink channel is a MIMO multiple access channel (MIMO-MAC) [49] where the users simultaneously transmit data over the same frequency channel to the BS equipped with multiple antennas. The BS must separate the received user signals by means of array processing, multiuser detection (MUD), or some other method [48]. **Figure 26** shows various linear and nonlinear MUD schemes for MIMO-OFDM systems, some of which are discussed in the later sections.

4.1.1. Classic SDMA-OFDM MUDs
An overview of some classic MUDs for MU-MIMO-OFDM is presented in [3]. The discussion is based on the
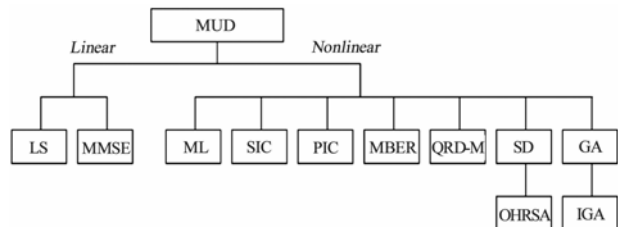
Figure 26. Various multiuser detectors (MUDs) for MIMO-OFDM systems [3].

uplink MIMO SDMA-OFDM system model of **Figure 27** where each of the L UEs uses a single transmit antenna while the BS is equipped with P antennas.

The complex-valued $P \times 1$ received signal vector at the BS antenna array for the k-th subcarrier of the n-th OFDM symbol is given by

$$\mathbf{x} = \mathbf{Hs} + \mathbf{n} \qquad (27)$$

where $\mathbf{s}$ is the $L \times 1$ transmitted signal vector, $\mathbf{n}$ is the $P \times 1$ AWGN noise vector and $\mathbf{H}$ is the $P \times L$ channel transfer function matrix consisting of L column vectors, each containing the transfer functions for a particular UE. Therefore, $\mathbf{H}$ can be represented as

$$\mathbf{H} = \left[\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \cdots, \mathbf{H}^{(L)} \right] \qquad (28)$$

where

$$\mathbf{H}^{(l)} = \left[H_1^{(l)}, H_2^{(l)}, \cdots, H_P^{(l)} \right]^T, \quad l = 1, \cdots, L \quad (29)$$

is a $P \times 1$ vector whose elements are the channel transfer functions for the transmission paths between the transmit antenna of the l-th UE and the P BS antennas.

It is assumed that the complex signal $s^{(l)}$ transmitted by the l-th user has zero mean and variance σ_l^2 while the AWGN signal n_p has zero mean and variance σ_n^2. The channel transfer functions $H_p^{(l)}$ are assumed to be independent, stationary, complex Gaussian distributed processes with zero mean and unit variance.

The classic MUD schemes [3] are discussed in the following text.

1) *MMSE MUD*:

Figure 28 shows the schematic diagram of a MMSE SDMA-OFDM MUD. The multiuser signals received at each BS antenna are multiplied by a complex-valued array weight $w_p^{(l)}$ and then summed up. The superscript l represents a particular user which means that a separate set of weights is used for detection of each user's signal. The combiner output $y(t)$ is subtracted from a user specific reference signal $r(t)$ known at the BS and the UE, resulting in an error signal $\epsilon(t)$. The error signal is used for weight estimation according to the MMSE criterion. The steepest descent algorithm can be used in this regard for stepwise weight adjustment for each subcarrier of each user. The performance of the MMSE MUD improves as the no. of antennas P in the BS antenna array is increased and degrades when the no. of users increase.
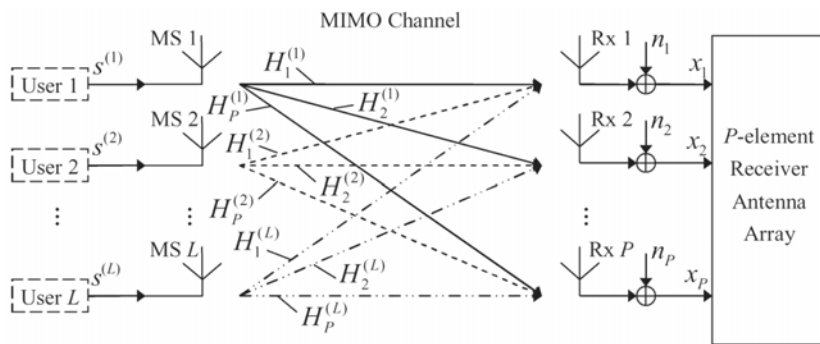
Figure 27. Uplink MIMO SDMA-OFDM system model with single antenna at each UE [3].
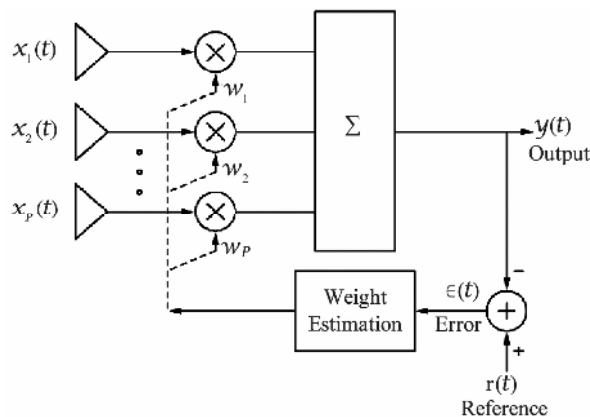
Figure 28. MMSE SDMA-OFDM MUD [3].

2) *Successive Interference Cancellation (SIC) MUD*:

The successive interference cancellation (SIC) MUD enhances the MMSE MUD using SIC. For each subcarrier, the detection order of the users is arranged according to their estimated total received signal power at the BS antenna array and the strongest user's signal with the least multiuser interference (MUI) is detected using the MMSE MUD. The detected user signal is then subtracted from the composite multiuser signal and the next strongest user is detected by the same procedure. This process continues till the detection is completed for all users. SIC results in high diversity gain at the MMSE combiner, which mitigates the effects of MUI as well as channel fading. The SIC MUD is also effective in near-far scenarios that result from inaccurate power control. However, it is prone to errors in power classification of user signals and also to interuser error propagation. **Figure 29** shows the BER performance comparison of MMSE MUD and SIC MUD (*M*-SIC with *M* = 2) for an SDMA-OFDM scenario with four single-antenna UEs and a four-antenna BS antenna array using QPSK modulation. The indoor short wireless asynchronous transfer mode (SWATM) channel model is used.

3) *Parallel Interference Cancellation (PIC) MUD*:

The PIC MUD does not require any power classification of the received user signals. The detection procedure consists of two iterations for all subcarriers. In the first iteration, MMSE detection is used to estimate all user signals $y^{(l)}$ from the received composite multiuser signal vector **x**. In case of channel encoded transmission, all user signals must be decoded, sliced, channel encoded

again and also remodulated onto subcarriers. In the second detection iteration, signal vectors $\breve{\mathbf{x}}^{(l)}$ for all L users are reconstructed and an estimate $\tilde{y}^{(k)}$ of each user signal is generated by subtracting the signal vectors $\breve{\mathbf{x}}^{(l)}$, $l \neq k$ of all other users followed by MMSE combining. The estimated user signals are then channel decoded and sliced. The PIC MUD scheme is also vulnerable to interuser error propagation.

4) *Maximum Likelihood (ML) MUD*:

The ML MUD employs the ML detection principle to find the most likely transmitted user signals through an exhaustive search. It provides the optimal detection performance but also has the highest complexity of any other MUD. For an OFDM-SDMA system with L simultaneous users, the ML MUD produces the estimated $L \times 1$ symbol vector $\hat{\mathbf{s}}_{\mathrm{ML}}$ consisting of the most likely transmitted symbols of the L users for a particular OFDM subcarrier, as given by

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg \min_{\breve{\mathbf{s}} \in M^L} \left\| \mathbf{x} - \mathbf{H}\breve{\mathbf{s}} \right\|^2 \tag{30}$$

where M^L is a set containing 2^{mL} trial vectors, m being the no. of bits per symbol depending on the modulation scheme used resulting in 2^m constellation points. Therefore, the computational complexity of the ML MUD increases exponentially with the no. of users L thus making it prohibitive for practical implementation.

5) *Sphere Decoding (SD) aided MUD*:

SD-aided MUDs use SD for reduced-complexity ML multiuser detection with near-optimal performance. SD reduces the ML search to within a hypersphere of a certain radius around the received signal. The radius of this search sphere determines the complexity of the MUD. Various SD algorithms have been proposed in literature like the complex-valued SD (CSD) and multistage SD (MSD) which significantly reduce the complexity by reducing the required search radius [3].

4.1.2. Layered Space-Time MUD

A V-BLAST based MUD scheme referred to as layered space-time MUD (LAST-MUD) is presented in [50] for CDMA uplink. This scheme is somewhat similar to the SIC MUD since V-BLAST detection also incorporates SIC.

Figure 30 shows the layered space-time MU-MIMO system block diagram. Here the single antenna users are arranged in G groups each containing M users for a total of $K = G \times M$. The UEs within each group are treated as the multiple transmit antennas of a V-BLAST system. The users within each group share the same unique spreading code which distinguishes the groups from one another. Therefore, out of the K total spreading codes, only G are unique. The $N \times K$ random spreading matrix consisting of K length N code vectors is denoted by

Fr.-Inv. Fad. SWATM, 4Rec.-Antennas, 4 Users, QPSK.
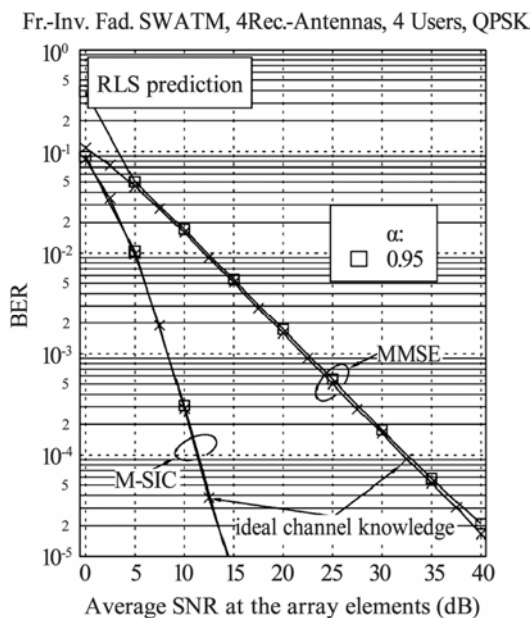
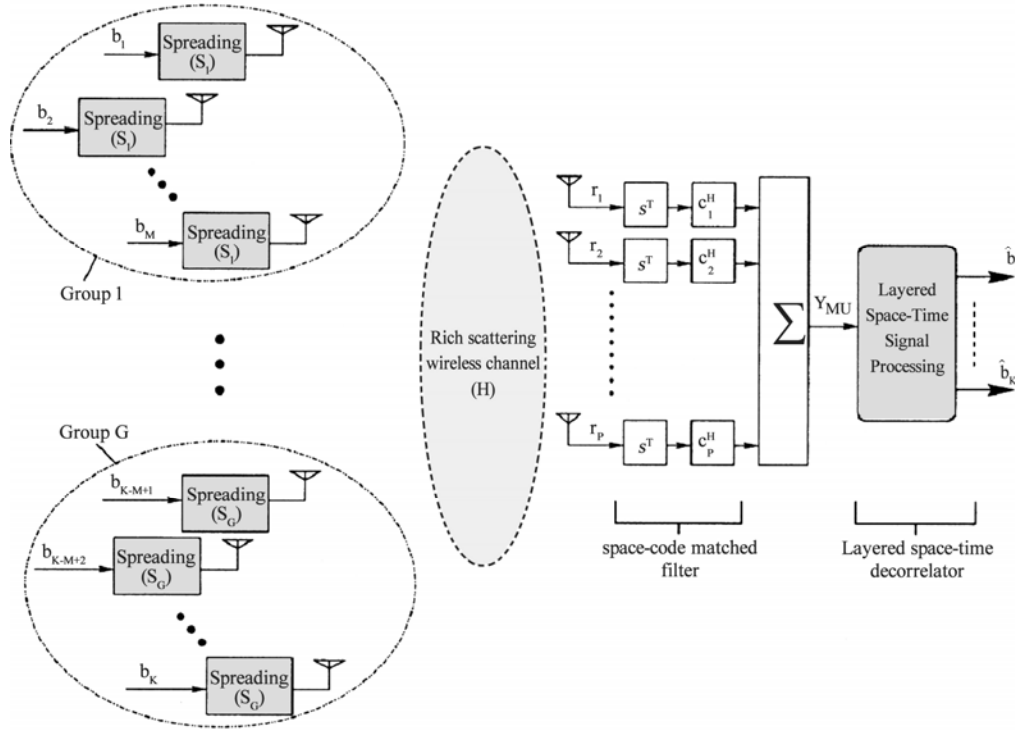**Figure 29. BER performance of MMSE MUD and SIC MUD [3].**

Figure 30. LAST MU-MIMO system [50].

$\mathbf{S} = \left[\mathbf{S}_1, \cdots, \mathbf{S}_1, \mathbf{S}_2, \cdots, \mathbf{S}_2, \cdots, \mathbf{S}_G, \cdots, \mathbf{S}_G \right]$. The proposed system can also accommodate users with multiple antennas thus enabling spatial multiplexing for achieving high data rates. In that case each user with multiple transmit antennas will be considered as one group. The $K \times 1$ transmitted symbol vector is represented as $\mathbf{b} = \left[b_1, \cdots, b_K \right]^T$ where each element represents the bit transmitted by a particular user. The channel between the users and the BS is considered to be a frequency-flat fading MIMO channel and is denoted by the channel matrix $\mathbf{H} = \left[\mathbf{h}_1, \cdots, \mathbf{h}_P \right]^T$ where $\mathbf{h}_p$ is the $K \times 1$ channel coefficient vector between all K users and the p-th BS antenna. The BS is equipped with a total of P antennas.

The $N \times 1$ received baseband signal at the p-th BS antenna for a certain symbol period after chip-matched filtering is given by

$$\mathbf{r}_p = \mathbf{S}\mathbf{C}_p\mathbf{b} + \mathbf{n}_p \tag{31}$$

where $\mathbf{C}_p = \mathrm{diag}\left(\mathbf{h}_p \right)$ is the complex diagonal channel matrix for the p-th BS antenna and $\mathbf{n}_p$ is the corresponding complex-valued AWGN noise vector with zero mean and variance σ^2. A frequency-flat fading MIMO channel is assumed as well as perfect channel estimation and symbol synchronization. The users are assumed to be

separated by a considerable distance so that the antennas of different users are not correlated. The channel estimation and symbol synchronization at the BS is also assumed to be ideal.

The BS employs space-code matched filtering to separate the different user groups. The $K \times 1$ sufficient statistic vector $\mathbf{Y}_{\mathrm{MU}}$ is then fed to the layered space-time decorrelator which eliminates the remaining inter-user interference to produce the estimated symbol vector $\hat{\mathbf{b}} = \left[\hat{b}_1, \cdots, \hat{b}_K \right]$. The vector $\mathbf{Y}_{\mathrm{MU}}$ is given by

$$\mathbf{Y}_{\mathrm{MU}} = \sum_{p=1}^{P} \mathbf{C}_p^H \mathbf{S}^T \mathbf{r}_p = \tilde{\mathbf{R}}_{\mathrm{MU}}\mathbf{b} + \tilde{\mathbf{n}} \tag{32}$$

where $\tilde{\mathbf{R}}_{\mathrm{MU}}$ is the $K \times K$ space-code cross-correlation matrix,

$$\tilde{\mathbf{R}}_{\mathrm{MU}} = \sum_{p=1}^{P} \mathbf{X}_p^H \mathbf{X}_p \tag{33}$$

with $\mathbf{X}_p = \mathbf{S}\mathbf{C}_p$. The $K \times 1$ real Gaussian noise vector $\tilde{\mathbf{n}}$ with covariance matrix $\sigma^2 \mathbf{R}_{\mathrm{MU}}$ is given by

$$\tilde{\mathbf{n}} = \sum_{p=1}^{P} \mathbf{X}_p^H \mathbf{n}_p \tag{34}$$

The detection algorithm is iterative and consists of three steps: 1) computation of the nulling vector, 2) user

IJCNS

signal estimation and 3) interference cancellation (SIC).

For the *i*-th iteration, the first step consists of calculating the pseudoinverse $\left[\tilde{\mathbf{R}}_{MU}(i)\right]^{+}$ of $\tilde{\mathbf{R}}_{MU}(i)$. The user signals are then ranked according to their post detection SNRs (following the space-code matched filtering) and the user having the highest SNR, given by

$$k_i = \arg \max_{j \notin \{k_1,\cdots,k_{i-1}\}} \frac{\left\| b_j \right\|^2}{\sigma^2 \left[\tilde{\mathbf{R}}_{MU}(i) \right]^{+}_{(j,j)}} \qquad (35)$$

is selected. The subscripts *i* and *j* in this equation denote the elements of an array, vector or matrix. The nulling vector for the selected user is $\mathbf{w}_{k_i} = \left[\tilde{\mathbf{R}}_{MU}(i)\right]^{+}_{k_i}$, i.e., the k_i-th column of $\left[\tilde{\mathbf{R}}_{MU}(i)\right]^{+}$. The slicer output is then given by

$$z_{k_i} = \mathbf{w}_{k_i}^{T} \mathbf{Y}_{MU}(i) \qquad (36)$$

resulting in the estimated symbol $\hat{b}_{k_i}$. The final step is interference cancellation where the detected symbol is subtracted from the received signal vector resulting in the symbol vector for the next iteration $i+1$, given by

$$\mathbf{r}_p(i+1) = \mathbf{r}_p(i) - \left(\mathbf{X}_p(i)\right)_{k_i} \hat{b}_{k_i} \qquad (37)$$

where $\left(\mathbf{X}_p(i)\right)_{k_i}$ is the k_i-th column of $\mathbf{X}_p(i)$. Similarly, $\mathbf{X}_p(i+1)$ and $\tilde{\mathbf{R}}_{MU}(i+1)$ are obtained by striking out the k_i-th column of $\mathbf{X}_p(i)$ and the k_i-th row and column of $\tilde{\mathbf{R}}_{MU}(i)$ respectively. $\mathbf{Y}_{MU}(i+1)$ is then given by

$$\mathbf{Y}_{MU}(i+1) = \sum_{p=1}^{P} \mathbf{X}_p(i+1)^{H} \mathbf{r}_p(i+1) \qquad (38)$$

This process is repeated until all *K* user signals are detected. Two reduced complexity versions of LAST-MUD called serial layered space-time group multiuser detector (LASTG-MUD) and parallel LASTG-MUD are also presented in [50].
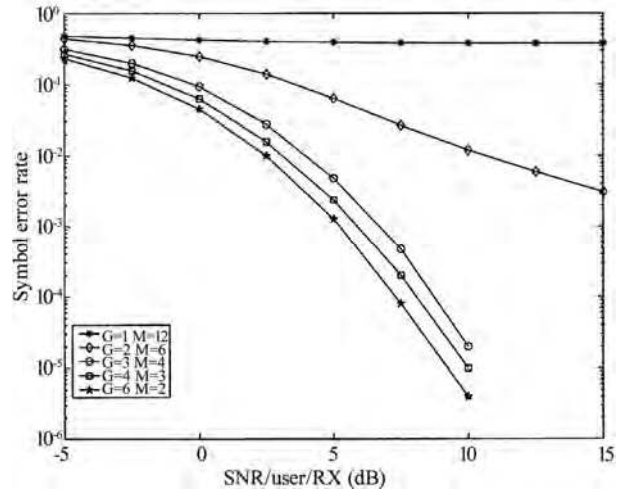
Figure 31 shows the SER performance of the LAST-MUD scheme using 4-QAM modulation with 12 simultaneous users (single-antenna) and 6 BS antennas, as the no. of user groups is increased. Fixed spreading factor of $N = 15$ is used. The performance improves as the users are distributed into more (smaller) groups since the no. of unique spreading codes also increases. For $G = 1$, the performance is equivalent to V-BLAST and represents the worst case.

Figure 32 shows the SER performance as the no. of users is increased by adding more user groups with $M =$

4 users per group. The LAST-MUD scheme provides substantial increase in network capacity by accommodating a large no. of simultaneous users with good SER performance.

4.1.3. SMMSE SIC MUD

A MUD scheme for MIMO-OFDM systems referred to as successive MMSE receive filtering with SIC (SMMSE SIC) is presented in [51]. The MMSE SIC MUD suffers from performance loss in scenarios with multiple closely spaced antennas located at the same UE. The proposed scheme tackles this problem by successively calculating the rows of the receive matrix at the BS for each of the UE transmit antennas, followed by SIC thus transforming the uplink MU-MIMO channel into a set of parallel SU-MIMO channels.

Figure 31. Grouping effect on SER performance of LAST-MUD [50].
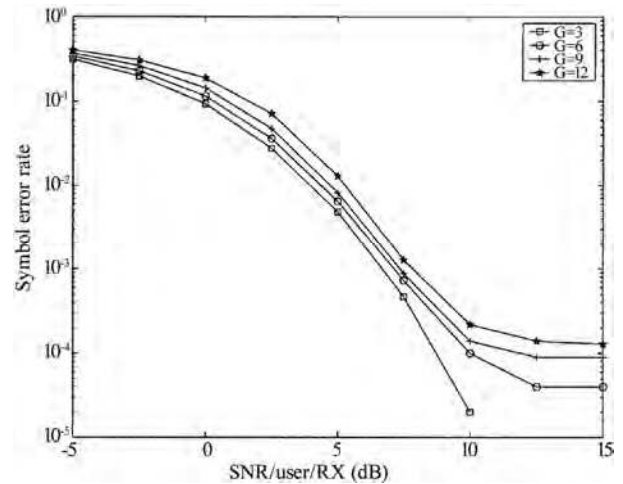
Figure 32. SER performance of LAST-MUD with increasing number of users [50].

Figure 33 shows the MU-MIMO uplink employing SMMSE SIC detection. The system consists of K simultaneous users, each equipped with M_{T_i} transmit antennas for $i = 1, \cdots, K$ and therefore, a total of $M_T = \sum_{i=1}^{K} M_{T_i}$ transmit antennas. The BS has M_R receive antennas. $\mathbf{s}_i$ and $\mathbf{r}_i$ represent the i-th user data vector and the receive vector respectively whereas $\mathbf{D}_i$ and $\mathbf{F}_i$ are the respective UE transmit matrix and the BS receive matrix. The MIMO channel matrix is represented as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \ \mathbf{H}_2 \ \cdots \ \mathbf{H}_K \end{bmatrix} \quad (39)$$

where $\mathbf{H}_i \in \mathbb{C}^{M_R \times M_{T_i}}$ is the MIMO channel matrix between user i and the BS for $i = 1, \cdots, K$. The $M_R \times M_T$ receive filter matrix at the BS is given by

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 \ \mathbf{F}_2 \ \cdots \ \mathbf{F}_K \end{bmatrix} \quad (40)$$

where $\mathbf{F}_i \in \mathbb{C}^{M_{T_i} \times M_R}$ corresponds to the i-th user. Each row of $\mathbf{F}_i$ corresponds to one of the M_{T_i} transmit antennas at the UE.

The proposed algorithm successively calculates the rows of each $\mathbf{F}_i$ using the MMSE criterion such that the users are ordered according to their respective total MSE in ascending order (starting with the minimum MSE). The total MSE of user i is obtained by summing up the MSE corresponding to each of its individual transmit antennas. Following the receive filtering by the receive filter matrix $\mathbf{F}$, SIC is performed to eliminate the MUI. This process has the effect of transforming the multiuser uplink channel into a set of parallel $M_{T_i} \times M_{T_i}$ SU-MIMO channels represented as $\mathbf{F}_i \mathbf{H}_i$.
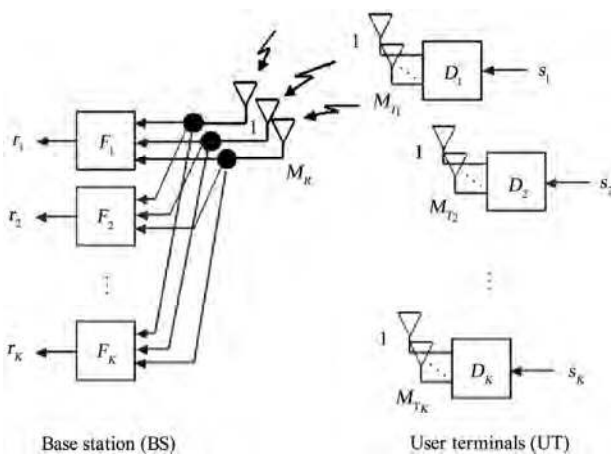
The SMMSE SIC detection in [51] has been considered for STBC based MIMO transmission i.e. the Alamouti scheme and also for dominant eigenmode transmission (DET). The open-loop Alamouti scheme provides MIMO diversity without the need for any CSI at the UEs. On the other hand, DET requires full CSI at both the transmitter and the receiver, which means that the UE transmit matrices $\mathbf{D}_i$ need to be computed at the BS and then fed forward to the UEs. However, beside the full diversity gain equal to that of STBC, DET also provides the maximum array gain by transmitting over the strongest eigenmode of the MIMO channel [49,51,52].

Figure 34 shows the BER performance of SMMSE SIC Alamouti in comparison with V-BLAST. The impact of channel estimation errors is also shown. The simulated MIMO-OFDM system consists of 6 BS antennas and 3 UEs equipped with 2 antennas each, denoted as 6 × {2, 2, 2} MU-MIMO configuration. The MIMO channel $\mathbf{H}$ is assumed to be frequency selective with the power delay profile defined by IEEE 802.11n - D for non-line-of-sight (NLOS) conditions. The channel model for each user's channel $\mathbf{H}_i$ takes into account the antenna correlation at the BS. However, the antennas at each UE are considered to have low spatial correlation assuming large angular spread at the user. A total of 64 OFDM subcarriers are used with subcarrier spacing of 150 kHz. The user data is encoded using a 1/2 rate convolutional code. 4-QAM modulation is used for each subcarrier of SMMSE SIC Alamouti system while BPSK is used for V-BLAST so that the data rate remains the same for both systems.

Clearly, the SMMSE SIC Alamouti system outperforms V-BLAST by a large margin particularly when the channel estimation errors are taken into account.

Figure 33. Block diagram of SMMSE SIC MUD based MU-MIMO system [51].
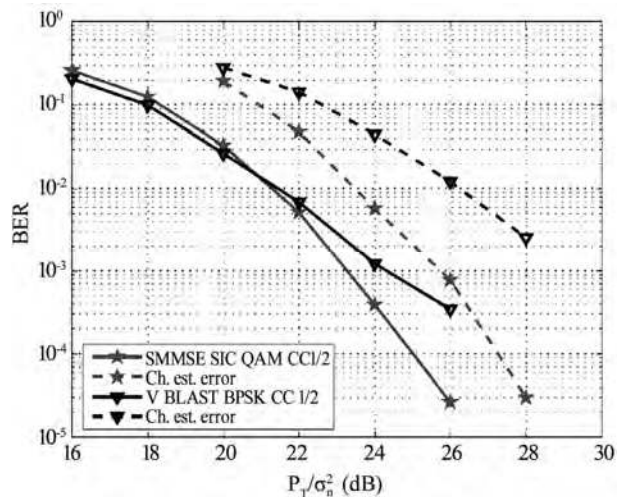
Figure 34. BER performance of SMMSE SIC Alamouti and V-BLAST for 6 × {2, 2, 2} configuration [51].

Figure 35 shows the performance gain of SMMSE SIC DET over SMMSE SIC Alamouti. SMMSE SIC DET shows substantial gains in the high SNR region and these become more significant as the no. of users in the system increases.

4.1.4. Turbo MUD

An iterative Turbo MMSE MUD scheme is presented in [53] for single-carrier (SC) space-time trellis-coded (STTC) SDMA MIMO systems in frequency-selective fading channels. This scheme can jointly detects multiple UE transmit antennas while cancelling the MUI from undetected users along with co-channel interference (CCI) and ISI through soft cancellation. Unknown co-channel interference (UCCI) from other interferers not known to the system is also considered. It is also shown that the no. of BS antennas required to achieve the corresponding lower performance bound of single-user detection is equal to the no. of users rather than the total no. of transmit antennas. The receiver derivations in [53] are provided for M-PSK modulation but can be extended to QAM as well.

Figure 36 shows the system model while the UE transmitter block diagram is given in **Figure 37**. The system has a total of $K + K_I$ simultaneous users, each indexed by $k = 1, \cdots K, K+1, \ldots K_I$ where the first K are the users to be detected at the BS while the remaining K_I are unknown interfering users representing the source of UCCI. Each UE is equipped with N_T transmit antennas. However, the system can also support unequal no. of antennas at the UEs. Each UE encodes the bit sequence $c_k(i)$ for $i = 1, \cdots, BN_T$ using a rate k_0 / N_T STTC code, where B represents the frame length in symbols. The encoded sequences $b_k(i) \in \mathbb{Q}$, $i = 1, \cdots, BN_T$ are
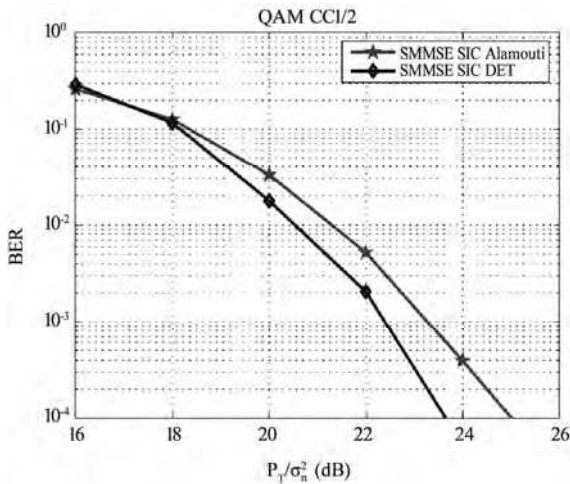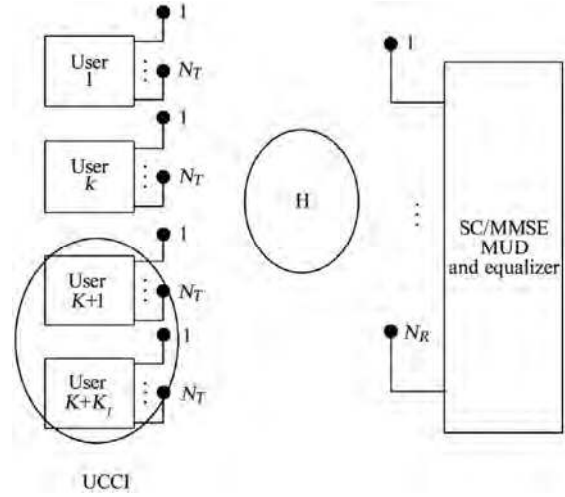

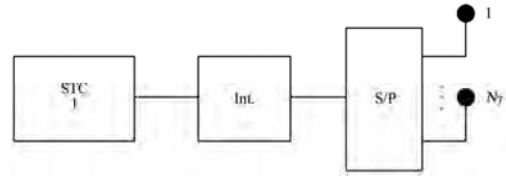**Figure 36. System model of Turbo MUD based STTC MU-MIMO system [53].**

Figure 37. UE transmitter block diagram [53].

grouped into B blocks of N_T symbols, where $\mathbb{Q} = \left\{ \alpha_1, \cdots, \alpha_{2^{k_0}} \right\}$ represents the modulation alphabet of M-PSK, and then interleaved by user-specific permutation of blocks of length N_T within a frame, such that the positions within each block remain unchanged. This interleaving process preserves the rank properties of the STTC code. User-specific training sequences of length TN_T are then attached at the start of the interleaved sequences. After serial-to-parallel (S/P) conversion of the entire frame, the resulting sequences $b_k^{(n)}(i)$ for $n = 1, \cdots, N_T$, $i = 1, \cdots, B+T$ are transmitted over the frequency-selective MIMO channel using the N_T transmit antennas.

The transmit signals from all users are received at the N_R BS receive antennas. The space-time sampled received signal vector $\mathbf{y}(i) \in \mathbb{C}^{LN_R \times 1}$ at time instant i is given by

$$\mathbf{y}(i) = \underbrace{\mathbf{H}\mathbf{u}(i)}_{\text{desired}} + \underbrace{\mathbf{H}_I \mathbf{u}_I(i)}_{\text{UCCI}} + \underbrace{\mathbf{n}(i)}_{\text{noise}}, \quad i = 1, \cdots, T+B \quad (41)$$

which can also be represented as

$$\mathbf{y}(i) = \left[\mathbf{r}^T(i+L-1), \cdots, \mathbf{r}^T(i) \right]^T \quad (42)$$

Figure 35. BER performance of SMMSE SIC Alamouti and SMMSE SIC DET for 6 × {2, 2, 2} configuration [51].

           IJCNS

where each vector $\mathbf{r}(i) \in \mathbb{C}^{N_R \times 1}$ consists of N_R elements $r_m(i)$, $m = 1, \cdots, N_R$ representing the received signal sample after matched filtering at the m-th receive antenna and L is the no. of paths of the frequency-selective MIMO channel. The channel matrix $\mathbf{H} \in \mathbb{C}^{LN_R \times KN_T(2L-1)}$ has the structure

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}(0) & \cdots & \mathbf{H}(L-1) & \cdots & \mathbf{0} \\ \vdots & \ddots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{H}(0) & \cdots & \mathbf{H}(L-1) \end{bmatrix} \quad (43)$$

Each element $\mathbf{H}(l) \in \mathbb{C}^{N_R \times KN_T}$ is given by

$$\mathbf{H}(l) = \begin{bmatrix} h_{1,1}^{(1)}(l) & \cdots & h_{1,1}^{(N_T)}(l) & \cdots & h_{K,1}^{(1)}(l) & \cdots & h_{K,1}^{(N_T)}(l) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ h_{1,N_R}^{(1)}(l) & \cdots & h_{1,N_R}^{(N_T)}(l) & \cdots & h_{K,N_R}^{(1)}(l) & \cdots & h_{K,N_R}^{(N_T)}(l) \end{bmatrix} \quad (44)$$

where $h_{k,m}^n(l)$ represents the complex gain for the l-th path between the n-th transmit antenna of user k and the m-th BS receive antenna. The channel matrix $\mathbf{H}_I \in \mathbb{C}^{LN_R \times K_I N_T(2L-1)}$ corresponding to UCCI has a similar structure as $\mathbf{H}$, and consists of matrices $\mathbf{H}_I(l) \in \mathbb{C}^{N_R \times K_I N_T}$.

The vectors $\mathbf{u}(i) \in \mathbb{Q}^{KN_T(2L-1) \times 1}$ and $\mathbf{u}_I(i) \in \mathbb{Q}^{K_I N_T(2L-1) \times 1}$ consist of the respective transmit sequences $b_k^{(n)}(i)$ of the desired and the unknown users, and are expressed as

$$\mathbf{u}(i) = \left[\mathbf{b}^T(i+L-1), \cdots, \mathbf{b}^T(i), \cdots, \mathbf{b}^T(i-L+1) \right]^T$$
$$\mathbf{u}_I(i) = \left[\mathbf{b}_I^T(i+L-1), \cdots, \mathbf{b}_I^T(i), \cdots, \mathbf{b}_I^T(i-L+1) \right]^T \quad (45)$$

where the vectors $\mathbf{b}(i) \in \mathbb{Q}^{KN_T \times 1}$ and $\mathbf{b}_I(i) \in \mathbb{Q}^{K_I N_T \times 1}$ are given by

$$\mathbf{b}(i) = \left[b_1^{(1)}(i), \cdots, b_1^{(N_T)}(i), \cdots b_K^{(1)}(i), \cdots b_K^{(N_T)}(i) \right]^T$$
$$\mathbf{b}_I(i) = \left[b_{K+1}^{(1)}(i), \cdots, b_{K+1}^{(N_T)}(i), \cdots b_{K+K_I}^{(1)}(i), \cdots b_{K+K_I}^{(N_T)}(i) \right]^T \quad (46)$$

$\mathbf{n}(i) \in \mathbb{C}^{LN_R \times 1}$ is the AWGN vector with covariance matrix $\sigma^2 \mathbf{I}$.

The Turbo receiver block diagram is shown in **Figure 38**. For the k-th user, the receiver first associates the signals from the user's N_T transmit antennas to N_T/n_0 sets of equal size n_0 such that the antennas $n = 1, \cdots, n_0$ represent the first set and so on. The training sequence $\mathbf{u}(i)$, $i = 1, \cdots, T$ is used to obtain an estimate $\hat{\mathbf{H}}$ of the channel matrix $\mathbf{H}$. The UCCI-plus-noise covariance matrix $\mathbf{R}$ is then estimated. The estimate for the first iteration is given by

$$\hat{\mathbf{R}} = \frac{1}{T} \sum_{i=1}^{T} \left(\mathbf{y}(i) - \hat{\mathbf{H}}\mathbf{u}(i) \right) \left(\mathbf{y}(i) - \hat{\mathbf{H}}\mathbf{u}(i) \right)^H \quad (47)$$
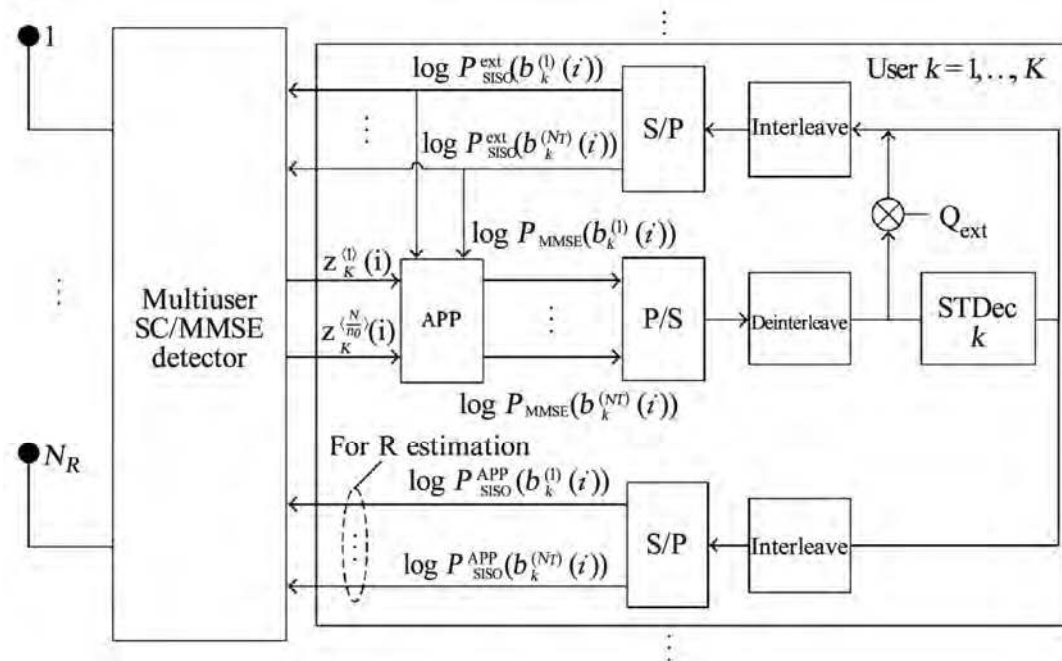

Figure 38. Block diagram of the iterative Turbo multiuser receiver [53].

From the second iteration onward, the soft feedback vector $\bar{\mathbf{u}}(i)$ from the APP (also MAP) SISO decoding is used to obtain the covariance matrix estimate

$$\hat{\mathbf{R}} = \frac{1}{T}\sum_{i=1}^{T}\left(\mathbf{y}(i) - \hat{\mathbf{H}}\mathbf{u}(i)\right)\left(\mathbf{y}(i) - \hat{\mathbf{H}}\mathbf{u}(i)\right)^{H} \\ + \frac{1}{B}\sum_{i=T+1}^{T+B}\left(\mathbf{y}(i) - \hat{\mathbf{H}}\bar{\mathbf{u}}(i)\right)\left(\mathbf{y}(i) - \hat{\mathbf{H}}\bar{\mathbf{u}}(i)\right)^{H}. \tag{48}$$

Vector $\bar{\mathbf{u}}(i)$ consists of the sequences $\bar{b}_k^{(n)}(i)$ calculated using the *a posteriori* probability $P_{\text{SISO}}^{\text{APP}}$. The signals $b_k^{(n)}(i)$, $n = 1, \cdots, n_0$ for user k are jointly detected using a linear MMSE MUD which filters the signal vector

$$\mathbf{y}_k^{\langle 1\rangle} = \mathbf{y}(i) - \hat{\mathbf{H}}\mathbf{u}_k^{\langle 1\rangle}(i), \quad i = T+1, \cdots, B+T \tag{49}$$

where the vector $\mathbf{u}_k^{\langle 1\rangle}(i)$ is calculated using the *extrinsic* probability $P_{\text{SISO}}^{\text{ext}}$ obtained after APP SISO decoding and $(\cdot)^{\langle 1\rangle}$ denotes the first set of n_0 antennas for user k. The weighting matrix $\mathbf{W}_k^{\langle 1\rangle}(i)$ for the MMSE MUD satisfies the criterion,

$$\left[\mathbf{W}_k^{\langle 1\rangle}(i), \mathbf{A}_k^{\langle 1\rangle}(i)\right] = \arg\min_{\mathbf{W},\mathbf{A}}\left\|\mathbf{W}^H\mathbf{y}_k^{\langle 1\rangle}(i) - \mathbf{A}^H\boldsymbol{\beta}_k^{\langle 1\rangle}(i)\right\|^2 \tag{50}$$

subject to the constraint $[\mathbf{A}]_{j,j} = 1$, $j = 1, \cdots, n_0$ to avoid the trivial solution $\left[\mathbf{W}_k^{\langle 1\rangle}(i), \mathbf{A}_k^{\langle 1\rangle}(i)\right] = [\mathbf{0},\mathbf{0}]$. The vector $\boldsymbol{\beta}_k^{\langle 1\rangle}(i) \in \mathbb{C}^{n_0 \times 1}$ is given by

$$\boldsymbol{\beta}_k^{\langle 1\rangle}(i) = \left[b_k^{(1)}(i), \cdots, b_k^{(n_0)}(i)\right]^T. \tag{51}$$
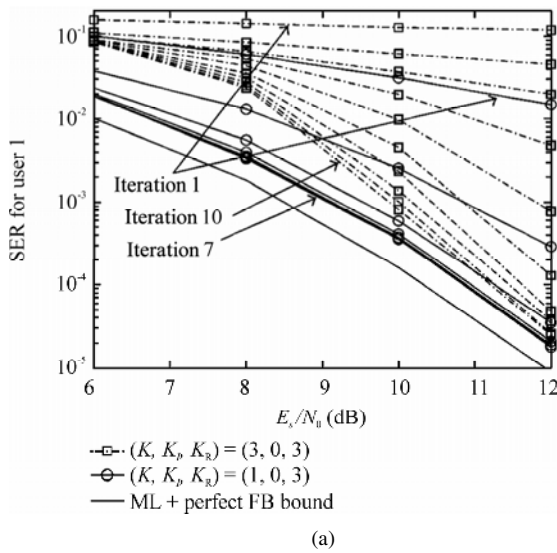
The corresponding output $\mathbf{z}_k^{\langle 1\rangle}(i) \in \mathbb{C}^{n_0 \times 1}$ of the MMSE MUD is given by

$$\mathbf{z}_k^{\langle 1\rangle}(i) = \mathbf{W}_k^{\langle 1\rangle}(i)\mathbf{y}_k^{\langle 1\rangle}(i) \\ = \boldsymbol{\Omega}_k^{\langle 1\rangle}(i)\boldsymbol{\beta}_k^{\langle 1\rangle}(i) + \boldsymbol{\psi}_k^{\langle 1\rangle}(i) \tag{52}$$

where the matrix $\boldsymbol{\Omega}_k^{\langle 1\rangle}(i) \in \mathbb{C}^{n_0 \times n_0}$ consists of the equivalent channel gains after filtering and $\boldsymbol{\psi}_k^{\langle 1\rangle}(i) \in \mathbb{C}^{n_0 \times 1}$ is the filtered AWGN vector. The MMSE MUD outputs $\mathbf{z}_k^{\langle \gamma\rangle}(i)$ along with the parameters $\boldsymbol{\Omega}_k^{\langle \gamma\rangle}(i)$ and $\boldsymbol{\psi}_k^{\langle \gamma\rangle}(i)$ for all antenna sets $\gamma = 1, \cdots, N_T/n_0$ of user k, are passed on to the APP SISO detector which calculates the *extrinsic* probabilities for SISO decoding. This iterative procedure is continued for all antenna sets of the remaining users until all users are detected. The complexity of this receiver is primarily associated with the MMSE and APP blocks and is on the order of $O\left\{\max\left(L^3 N_R^3, 2^{k_0 n_0}\right)\right\}$.

Two special cases of the proposed receiver architecture are considered in [53]. Receiver 1, with $n_0 = 1$, detects the transmit antennas of user k one by one. This receiver type has the lowest complexity since the complexity depends exponentially on n_0. Receiver 2, with $n_0 = N_T$, jointly detects all N_T transmit antennas of user k and has the highest complexity.

The simulated SER and FER performance of the two receiver types vs. per-antenna E_s/N_0 is shown in **Figure 39** and **Figure 40**, for a particular user referred to as user 1. Performance comparison with optimal joint ML
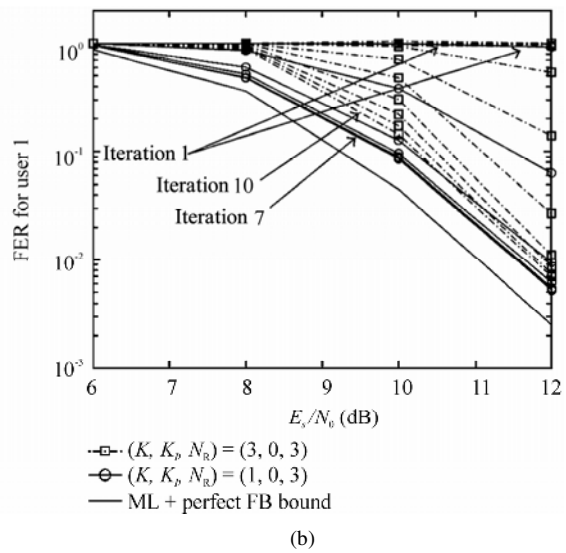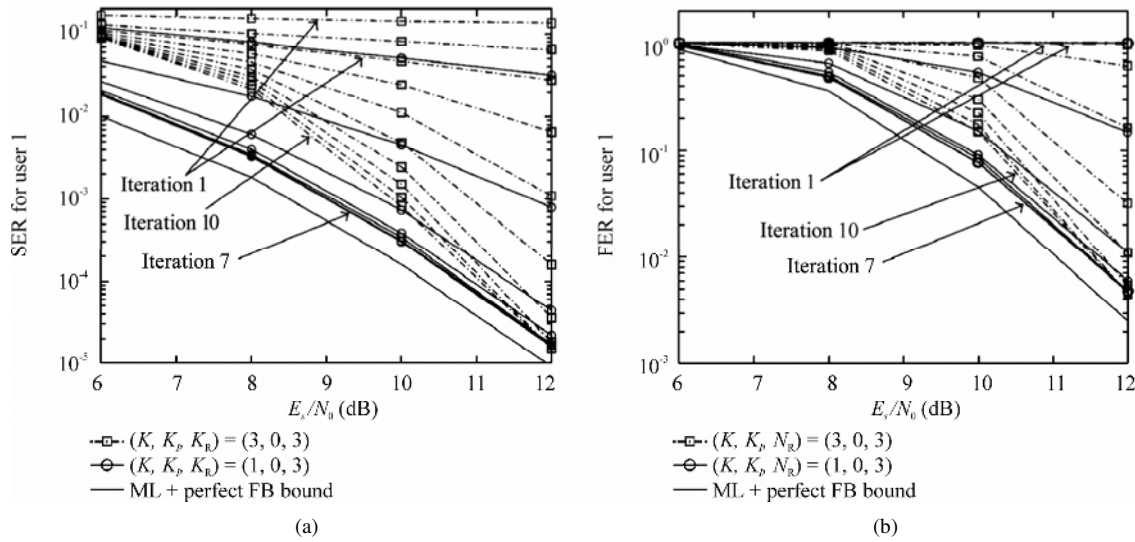

Figure 39. Receiver 1's (a) SER and (b) FER performance, $(K, K_I, N_R) = (3, 0, 3)$ and $(1, 0, 3)$, $N_T = 2$ [53].

detection of all N_T antennas of user 1 followed by MAP SISO decoding assuming perfect feedback (FB), is also provided. The results are provided for 1 and 10 iterations for $(K, K_I, N_R) = (3, 0, 3)$ and for 1 and 7 iterations for $(K, K_I, N_R) = (1, 0, 3)$ with $N_T = 2$ transmit antennas per user. K, K_I and N_R represent the no. of desired users, unknown interferers (UCCI) and BS receive antennas. QPSK modulation is used for transmission and frequency-selective fading is assumed with $L = 5$ uncorrelated Rayleigh distributed paths. For a sufficient no. of iterations, both receivers perform reasonably

close to the ML receiver, particularly in the high SNR region. Receiver 2 shows slightly better performance than receiver 1.

Figure 41 shows the SER and FER performance of the two receivers for $(K, K_I, N_R) = (2, 1, 3)$ with $N_T = 2$ transmit antennas per desired user and a single transmit antenna for the unknown interferer (UCCI) i.e. $N_I = 1$. Two cases of signal-to-UCCI interference ratio (SIR) are considered. For SIR = 3 dB, the signal transmitted from UCCI's antenna is assumed to have the same power as that of the signal from one antenna of the desired user whereas, for SIR = 0 dB, UCCI's antenna transmits at

Figure 40. Receiver 2's (a) SER and (b) FER performance, $(K, K_I, N_R) = (3, 0, 3)$ and $(1, 0, 3)$, $N_T = 2$ [53].
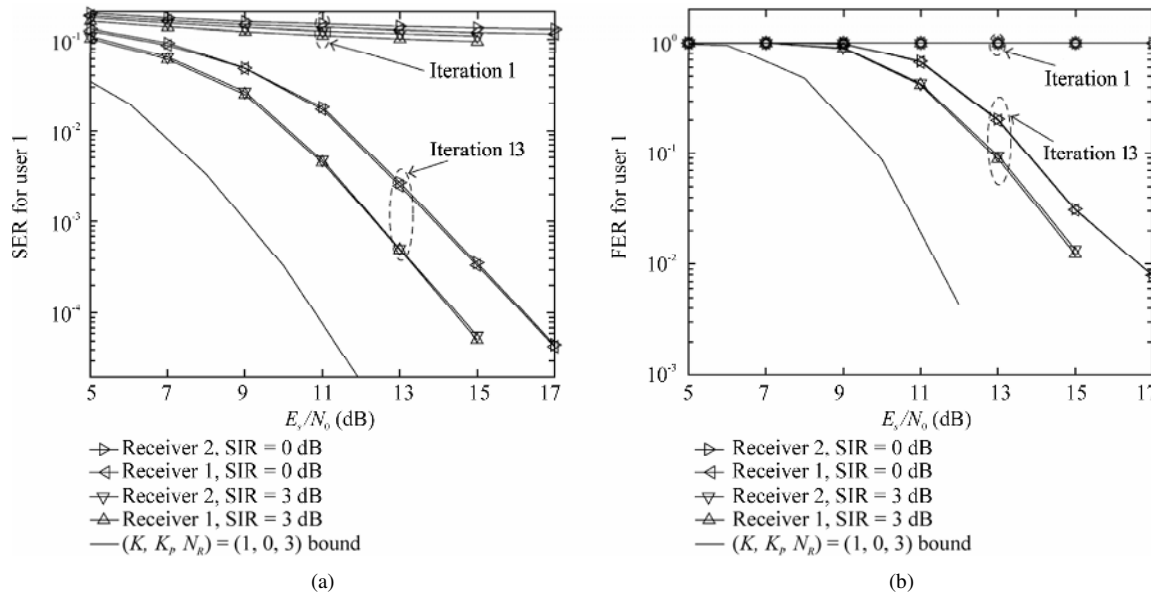
Figure 41. Receiver 1's and 2's (a) SER and (b) FER performance, $(K, K_I, N_R) = (2, 1, 3)$, $N_T = 2$, $N_I = 1$ [53].

twice the transmit power of a desired user's antenna. The performance of both receivers is obviously better for the 3 dB SIR case. However, the UCCI has a considerable impact on performance and more iterations are needed to achieve a reasonably low SER or FER. The performance will degrade further in case of multiple UCCI antennas and also as the UCCI sources i.e. the no. of unknown interferers increase.

4.1.5. Genetic Algorithm (GA) Assisted MUDs

Genetic algorithms (GAs) are based on the theory of evolution's concept of survival of the fittest, where the genes from the fittest individuals of a species are passed on to the next generation through the process of "natural selection". When applied to MUDs, an "individual" represents the L-dimensional MUD weight vector corresponding to the L users. These MUD weights are then optimized using GA by genetic operations of "mating" and "mutation" to get a new generation of individuals i.e. the MUD weights. The initial "population" (MUD weights) is typically obtained from the MMSE solution which is retained throughout the GA search process as an alternate solution in case of poor convergence [3].

Considering the SDMA-OFDM system model of **Figure 27** with a single transmit antenna at each UE and P receive antennas at the BS, the ML-based decision metric or objective function (OF) for a GA-assisted MUD corresponding to the p-th receive antenna can be written as

$$\Omega_p(\breve{\mathbf{s}}) = \left| x_p - \mathbf{H}_p \breve{\mathbf{s}} \right|^2 \qquad (53)$$

where x_p is the received symbol corresponding to the p-th BS antenna for a specific OFDM subcarrier and

$\mathbf{H}_p$ is the p-th row of the channel transfer function matrix $\mathbf{H}$. The estimated symbol vector of the L users corresponding to the p-th BS antenna is then given by

$$\hat{\mathbf{s}}_{\mathrm{GA}_p} = \arg\left\{ \min_{\breve{\mathbf{s}}} \left[\Omega_p(\breve{\mathbf{s}}) \right] \right\} \qquad (54)$$

The combined decision metric for the P receive antennas can therefore be written as

$$\Omega(\breve{\mathbf{s}}) = \sum_{p=1}^{P} \Omega_p(\breve{\mathbf{s}}) = \left\| \mathbf{x} - \mathbf{H}\breve{\mathbf{s}} \right\|^2 \qquad (55)$$

Therefore, the decision rule for the GA-assisted MUD is to find an estimate $\hat{\mathbf{s}}_{\mathrm{GA}}$ of the $L \times 1$ transmitted symbol vector such that $\Omega(\breve{\mathbf{s}})$ is minimized [3].

Review and analysis of various GA-assisted MUDs is provided in [3] for the SDMA-OFDM uplink consisting of a single transmit antenna at each UE. **Figure 42** shows the schematic diagram of the SDMA-OFDM uplink system based on the concatenated MMSE-GA MUD. The concatenated MMSE-GA MUD uses the MMSE estimate $\hat{\mathbf{s}}_{\mathrm{MMSE}} \in \mathbb{C}^{L \times 1}$ of the transmitted symbol vector of the L users as initial information for the GA. $\hat{\mathbf{s}}_{\mathrm{MMSE}}$ is given by

$$\hat{\mathbf{s}}_{\mathrm{MMSE}} = \mathbf{W}_{\mathrm{MMSE}}^{H} \mathbf{x} \qquad (56)$$

where $\mathbf{W}_{\mathrm{MMSE}} \in \mathbb{C}^{P \times L}$ is the MMSE MUD weight matrix, expressed as

$$\mathbf{W}_{\mathrm{MMSE}} = \left(\mathbf{HH}^{H} + \sigma_n^2 \mathbf{I} \right)^{-1} \mathbf{H} \qquad (57)$$
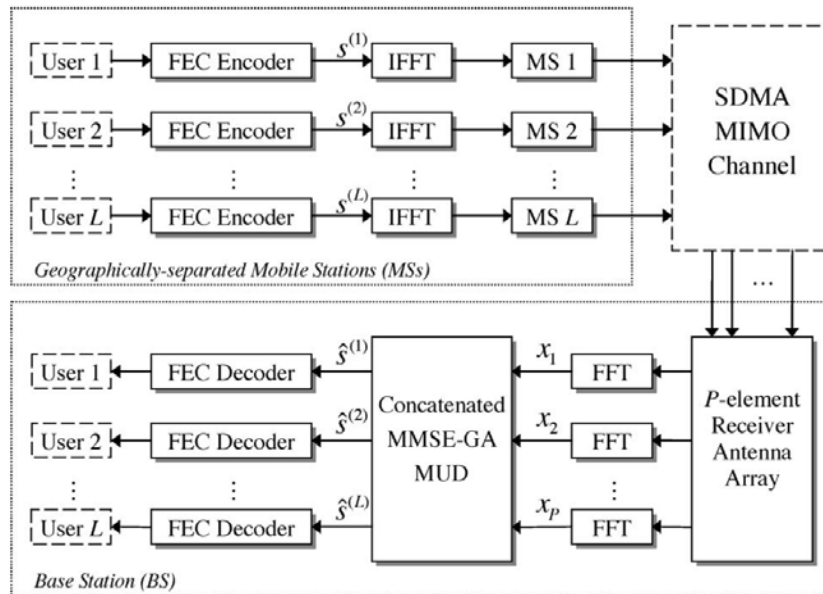
Figure 42. SDMA-OFDM uplink system based on concatenated MMSE-GA MUD [3].

Using this MMSE estimate, the 1st GA *generation*, $y = 1$ containing a *population* of X *individuals*, is created. The *x*-th individual is a symbol vector denoted as

$$\tilde{\mathbf{s}}_{(y,x)} = \left[\tilde{s}_{(y,x)}^{(1)}, \tilde{s}_{(y,x)}^{(2)}, \cdots, \tilde{s}_{(y,x)}^{(L)} \right], \quad x = 1, \cdots, X \quad (58)$$
$$y = 1, \cdots, Y$$

where each element $\tilde{s}_{(y,x)}^{(l)}$ called a *gene*, belongs to the set of complex-valued modulation symbols corresponding to the particular modulation scheme used.

The GA search procedure shown in **Figure 43** is then initiated, which involves several GA operations like *mating*, *mutation*, *elitism* etc. leading to the next generation. This process is repeated for Y generations and the individual with the highest *fitness value* is considered to be the detected $L \times 1$ multiuser symbol vector for the

corresponding OFDM subcarrier. All users are jointly detected by the concatenated MMSE-GA MUD and therefore, no error propagation exists between the detected users.

Enhanced GA MUDs, utilizing an advanced mutation technique called biased Q-function-based mutation (BQM) instead of the conventional uniform mutation (UM), and incorporating the iterative turbo trellis coded modulation (TTCM) scheme for FEC decoding, are also discussed in [3]. **Figure 44** shows the schematic diagram of an MMSE-initialized iterative GA (IGA) MUD incorporating TTCM. The $P \times 1$ received symbol vector **x** is detected by the MMSE MUD to get the estimated symbol vector $\hat{\mathbf{s}}_{\text{MMSE}}$ of the L users consisting of the symbols $\hat{s}_{\text{MMSE}}^{(l)}$, $l = 1, \cdots, L$. Each of these symbols are then
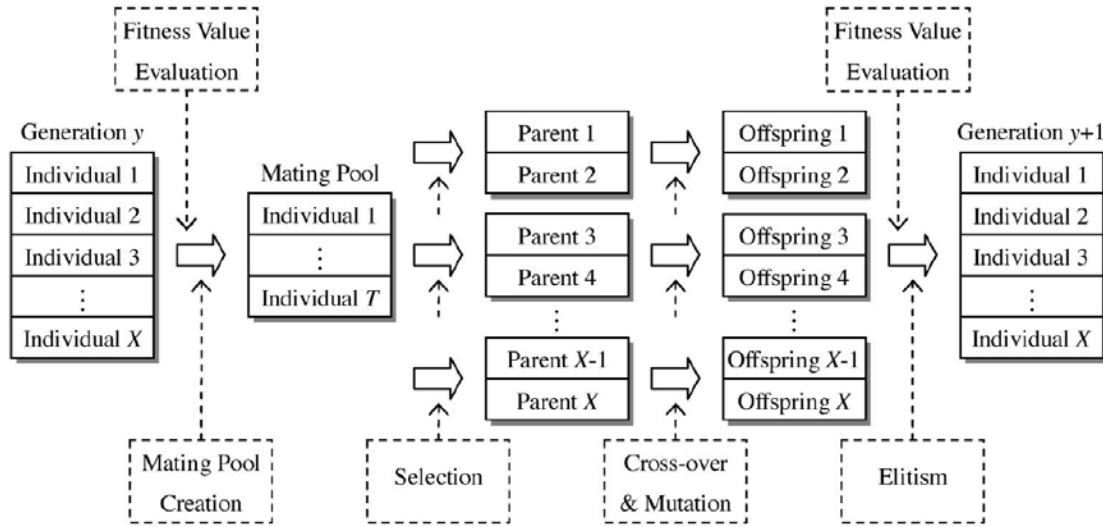

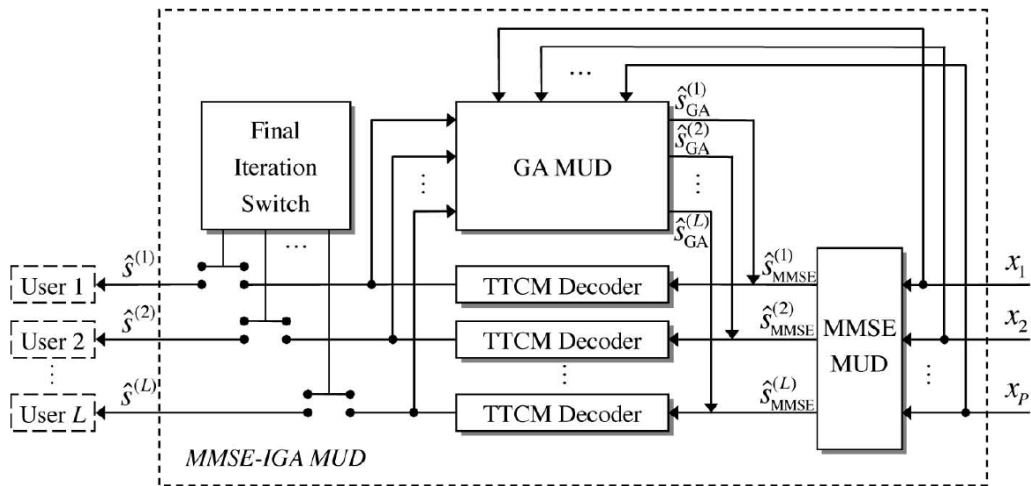**Figure 43. GA search procedure for one generation [3].**

Figure 44. Schematic diagram of an IGA MUD [3].

decoded by a TTCM decoder to get a more reliable estimate. The resulting symbol vector is then used as the initial information for the GA MUD. The GA-estimated symbol vector $\hat{\mathbf{s}}_{GA}$ is then fed back to the TTCM decoders for further improvement of the estimate. This optimization process involving the GA MUD and the TTCM decoders is continued for a desired no. of iterations. The final estimates $\hat{s}^{(l)}$ of the L users' symbols are then obtained at the output after the final iteration.

Figure 45 shows the BER performance comparison of various MMSE-initialized TTCM-assisted GA and IGA MUD based SDMA-OFDM systems consisting of $L = 6$ (single-antenna) users and $P = 6$ BS antennas. 4-QAM modulation and SWATM channel model is used. GA population size of $X = 20$ (also $X = 10$ for TTCM, MMSE-IGA (2)) and a total of $Y = 5$ generations are considered. UM and BQM mutation schemes are employed. Performance curves for 1×1 SISO AWGN, 1×6 MRC AWGN, TTCM-MMSE SDMA-OFDM and TTCM-ML SDMA-OFDM systems are also provided for reference. The TTCM-MMSE-GA MUDs and the IGA MUDs in particular provide exceptionally good BER performance. The performance of the TTCM-MMSE-IGA scheme with 2 iterations and $X = 20$ (represented as TTCM, MMSE-IGA (2) in the figure), is in fact identical to the optimum TTCM-ML MUD.

The BER performance comparison for rank-deficient scenarios where the no. of users exceeds the no. of BS antennas resulting in insufficient degrees of freedom for separating the users, is given in **Figure 46**. Performance curves for $L = 6$, 7 and 8 users and $P = 6$ BS antennas are provided. The IGA MUD schemes still perform reasonably good with relatively small performance degradation.

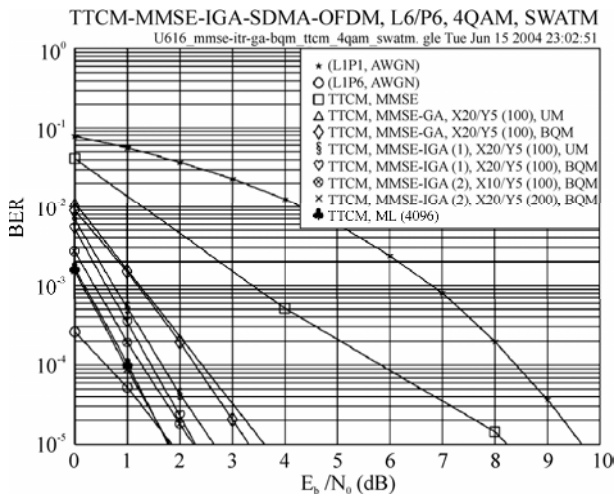**Figure 47** compares the complexity of the TTCM-MMSE-GA and TTCM-ML MUDs in terms of the no. of
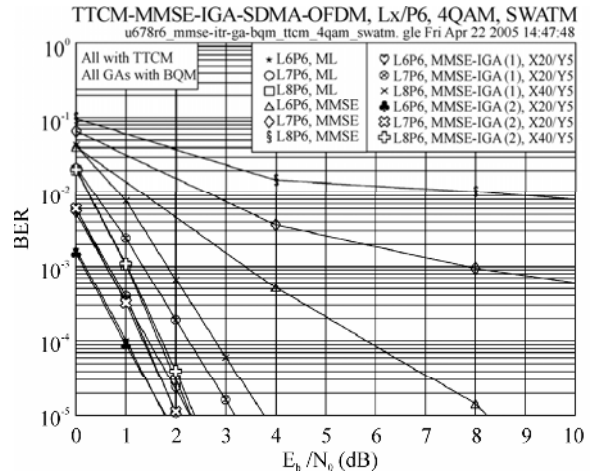

Figure 46. BER performance of TTCM-MMSE-GA/IGA SDMA-OFDM systems, $L = 6, 7, 8$ and $P = 6$ [3].
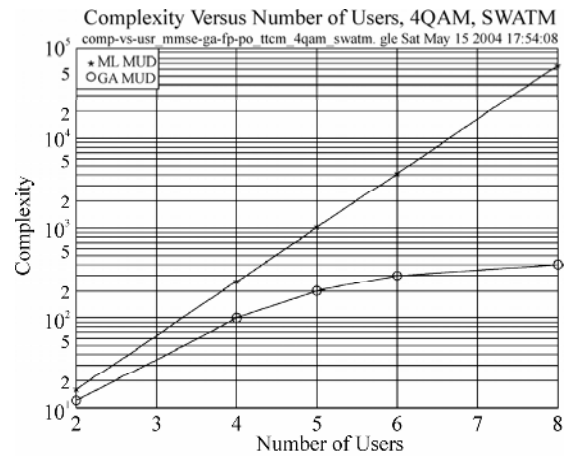
Figure 47. Complexity of TTCM-MMSE-GA and TTCM-ML SDMA-OFDM systems vs. no. of users, $L = P$ [3].

OF calculations, as the no. of users increase. The no. of BS antennas is kept equal to the no. of users i.e. $L = P$. The complexity of the GA MUD increases very slowly with the number of users as compared to the ML MUD, resulting in a huge difference as more users are added to the system.

4.2. The MU-MIMO Downlink

The MU-MIMO downlink channel is referred to as the MIMO broadcast channel (MIMO-BC) [49] where the BS equipped with multiple antennas, simultaneously transmits data to multiple UEs consisting of one or more antennas each, as shown in **Figure 48**. The multiuser interference (MUI) (also called multiple access interference, MAI) can be suppressed by means of transmit beamforming or "dirty paper" coding. Therefore, CSI

Figure 45. BER performance of TTCM-MMSE-GA/IGA SDMA-OFDM systems, $L = 6, P = 6$ [3].

feedback from each user is required for precoding at the BS. Various linear and nonlinear precoding techniques for MU-MIMO downlink systems are discussed in the following text.

4.2.1. Channel Inversion

Channel inversion is a linear precoding technique for MU-MIMO downlink systems where each UE is equipped with a single receive antenna. As the name suggests, channel inversion uses the inverse of the channel matrix for precoding to remove the MUI, as illustrated in **Figure 49**.

Assuming that the no. of receive antennas $M_R \le M_T$, the no. of transmit antennas, ZF precoding can be used for this purpose. The $M_T \times 1$ transmitted signal vector is then given by

$$\begin{aligned} \mathbf{x} &= \mathbf{H}^+ \mathbf{d} \\ &= \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H \right)^{-1} \mathbf{d} \end{aligned} \quad (59)$$
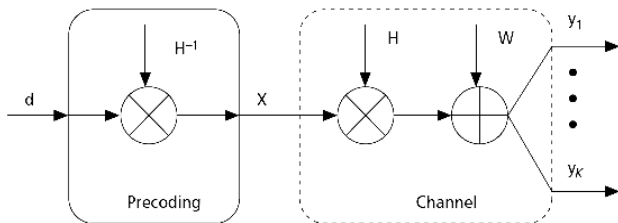
where $\mathbf{d}$ is the vector of data symbols to be precoded and $\mathbf{H}^+$ is the pseudoinverse of the $M_R \times M_T$ channel matrix $\mathbf{H}$. Vector $\mathbf{d}$ can have any dimension up to the rank of $\mathbf{H}$ [48]. The *i*-th column of the prefiltering or precoding matrix $\mathbf{P}_{ZF}$ is given by [49]

$$\mathbf{p}_{ZF,i} = \frac{\mathbf{h}_i^{(+)}}{\sqrt{\left\| \mathbf{h}_i^{(+)} \right\|^2}} \quad (60)$$

where $\mathbf{h}_i^{(+)}$ is the *i*-th column of $\mathbf{H}^+$. The combined received signal vector can be expressed as
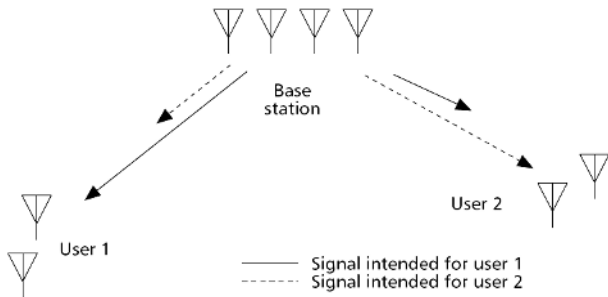
Figure 48. The MU-MIMO downlink [48].

Figure 49. Channel inversion [48].

$$\mathbf{y} = \mathbf{d} + \mathbf{w} \quad (61)$$

where $\mathbf{w}$ is the noise vector. Therefore, ZF precoding is only suitable for low-noise or high transmit power scenarios [48].

MMSE precoding, also called "regularized" channel inversion provides a better alternative. In this case, the transmitted signal vector is given by

$$\mathbf{x} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \alpha \mathbf{I} \right)^{-1} \mathbf{d} \quad (62)$$

where α is the loading factor. For a MU-MIMO downlink system with total transmit power P_T and K simultaneous users, $\alpha = K / P_T$ maximizes the SINR at the receivers [48].

4.2.2. Block Diagonalization

Block diagonalization (BD) or block channel inversion, which was first proposed in [54], is a generalization of channel inversion to multi-antenna UEs [48]. BD also requires the total no. of receive antennas M_R to be less than or equal to the no. of transmit antennas M_T i.e. $M_R \le M_T$.

Consider the system model of **Figure 50**. The system consists of K simultaneous users each having M_{R_i} receive antennas for $i = 1, \cdots, K$ such that the total no. of receive antennas $M_R = \sum_{i=1}^{K} M_{R_i}$. The combined channel matrix $\mathbf{H} \in \mathbb{C}^{M_R \times M_T}$ is given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1^T & \mathbf{H}_2^T & \cdots & \mathbf{H}_K^T \end{bmatrix}^T \quad (63)$$

where $\mathbf{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$ represents the MIMO channel from the M_T BS antennas to user *i*. The combined precoding matrix $\mathbf{P} \in \mathbb{C}^{M_T \times S}$ can be expressed as

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 & \cdots & \mathbf{P}_K \end{bmatrix} \quad (64)$$

where $\mathbf{P}_i \in \mathbb{C}^{M_T \times S_i}$ is the precoding matrix for the *i*-th user, $S \le M_R$ is the total no. of transmitted data streams and $S_i \le M_{R_i}$ is the no. of data streams transmitted to user *i*. $\mathbf{P}$ needs to be selected in such a way that $\mathbf{HP}$ becomes block diagonal. To this end, a matrix $\tilde{\mathbf{H}}_i$ is defined as
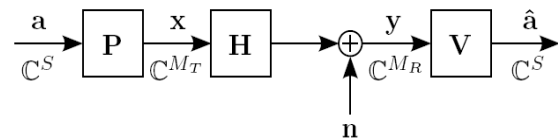
Figure 50. System model for MU-MIMO downlink transmission [55].

$$\tilde{\mathbf{H}}_i = \begin{bmatrix} \mathbf{H}_1^T & \cdots & \mathbf{H}_{i-1}^T & \mathbf{H}_{i+1}^T & \cdots & \mathbf{H}_K^T \end{bmatrix}^T \quad (65)$$

which contains all but the *i*-th user's channel matrix. Therefore, $\mathbf{P}_i$ lies in the null space of $\tilde{\mathbf{H}}_i$ and consists of unitary column vectors which are obtained by the SVD of $\tilde{\mathbf{H}}_i$, given by

$$\tilde{\mathbf{H}}_i = \tilde{\mathbf{U}}_i \tilde{\mathbf{D}}_i \begin{bmatrix} \tilde{\mathbf{V}}_i^{(1)} & \tilde{\mathbf{V}}_i^{(0)} \end{bmatrix}^H \quad (66)$$

The rightmost $M_T - \tilde{L}_i$ singular vectors $\tilde{\mathbf{V}}_i^{(0)} \in \mathbb{C}^{M_T \times M_T - \tilde{L}_i}$ form an orthogonal basis for the null space of $\tilde{\mathbf{H}}_i$ where $\tilde{L}_i$ is the rank of $\tilde{\mathbf{H}}_i$. The product $\mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)}$ with dimensions $M_{R_i} \times (M_T - \tilde{L}_i)$ represents the equivalent channel matrix for user *i* after eliminating the MUI. Thus, BD transforms the MU-MIMO downlink system into a set of *K* parallel $(M_T - \tilde{L}_i) \times M_{R_i}$ SU-MIMO systems. Using SVD, $\mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)}$ can be expressed as

$$\mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)} = \mathbf{U}_i \begin{bmatrix} \mathbf{D}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_i^{(1)} & \mathbf{V}_i^{(0)} \end{bmatrix}^H \quad (67)$$

where $\mathbf{D}_i$ is an $L_i \times L_i$ diagonal matrix, assuming L_i to be the rank of $\mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)}$. The product of $\tilde{\mathbf{V}}_i^{(0)}$ and the first L_i singular vectors $\mathbf{V}_i^{(1)}$ produces an orthogonal basis of dimension L_i and represent the transmission vectors which maximize the information rate for the *i*-th user while eliminating the MUI. Therefore, the precoding matrix $\mathbf{P}_i$ for user *i* consists of $\tilde{\mathbf{V}}_i^{(0)} \mathbf{V}_i^{(1)}$ with appropriate power scaling. Optimal power allocation is achieved by water-filling, using the diagonal elements of the matrices $\mathbf{D}_i$ and can either be implemented globally to maximize the overall information rate of the system or on a per-user basis [54,56,57].

4.2.3. Successive Optimization

Successive optimization (SO) [54,56,57] is a successive precoding algorithm which addresses the power control problem in BD where capacity loss occurs due to the nulling of overlapping subspaces of different users. First, an optimum ordering of the users is determined like in case of V-BLAST detection. The precoding matrix for each user is then designed in a successive manner so that it lies in the null space of the channel matrices of the previous users only. In other words, the transmit power of user *i* is optimized in such a manner that it does not interfere with users $1, \cdots, i-1$. However, interference with the successive users is allowed. The combined channel matrix for the previous $i-1$ users can be written as

$$\hat{\mathbf{H}}_i = \begin{bmatrix} \mathbf{H}_1^T & \mathbf{H}_2^T & \cdots & \mathbf{H}_{i-1}^T \end{bmatrix}^T \quad (68)$$

and its SVD is given by

$$\hat{\mathbf{H}}_i = \hat{\mathbf{U}}_i \hat{\mathbf{D}}_i \begin{bmatrix} \hat{\mathbf{V}}_i^{(1)} & \hat{\mathbf{V}}_i^{(0)} \end{bmatrix}^H \quad (69)$$

$\hat{\mathbf{V}}_i^{(0)}$ contains the $M_T - \hat{L}_i$ rightmost singular vectors where $\hat{L}_i$ is the rank of $\hat{\mathbf{H}}_i$. The precoding matrix $\mathbf{P}_i$ that lies in the null space of $\hat{\mathbf{H}}_i$ is then determined as $\mathbf{P}_i = \hat{\mathbf{V}}_i^{(0)} \mathbf{P}_i'$ for some choice of $\mathbf{P}_i'$.

4.2.4. Dirty Paper Coding

Dirty paper coding is a nonlinear precoding technique and is based on the concept introduced by Costa [58] where the AWGN channel is modified by adding interference which is known at the transmitter. This concept is analogous to "writing on dirty paper" where the writing is the desired signal and the dirt represents the interference. Since the transmitter knows where the "dirt" or interference is, writing on dirty paper is the same as writing on clean paper [48].

In addition to SU-MIMO systems, e.g. the GMD-ZFDP scheme mentioned in Subsection 3.4, dirty paper coding is also applicable to MU-MIMO downlink transmission. In a MU-MIMO downlink system, CSI feedback from the users is available at the BS and it can figure out the interference produced at a particular user by the signals meant for other users. Therefore, dirty paper coding can be applied to each user's signal at the BS so that the known interference from other users is avoided.

Various dirty paper coding techniques for the MU-MIMO downlink are discussed in [48]. A well-known approach is to use QR decomposition of the channel matrix, given by $\mathbf{H} = \mathbf{LQ}$, where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{Q}$ is a unitary matrix. $\mathbf{Q}^H$ is then used for transmit precoding which results in the effective channel $\mathbf{L}$. Therefore, the first user does not see any interference from the other users and no further processing of its signal is required at the BS. However, each of the subsequent users sees interference from the preceding users and dirty paper coding is applied to eliminate this known interference.

Another technique called *vector precoding* jointly precodes the users' signals rather than applying dirty paper coding to the users' signals individually. The vector precoding technique is shown in **Figure 51**. The desired signal vector **d** is offset by a vector *l* of integer values and this operation is followed by channel inversion, resulting in the transmitted signal **x**, given by

$$\mathbf{x} = \mathbf{H}^{-1}(\mathbf{d} + \tau l) \quad (70)$$

where the vector *l* is chosen to minimize the power of **x**, i.e.

Figure 51. Vector precoding [48].
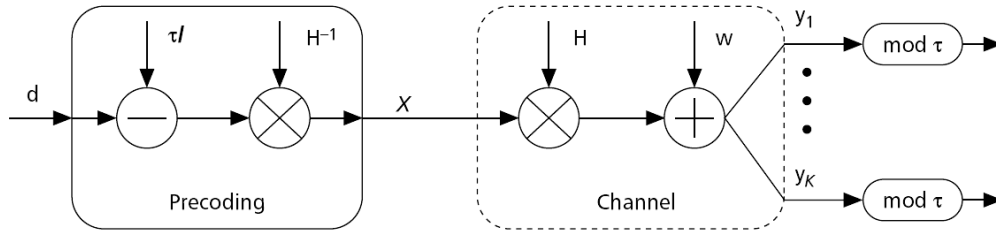
$$l = \arg \min_{l'} \left\| \mathbf{H}^{-1} \left(\mathbf{d} + \tau l' \right) \right\| \qquad (71)$$

The signal received at the *k*-th user is expressed as

$$y_k = d_k + \tau l_k + w_k \qquad (72)$$

where w_k represents the Gaussian noise. A modulo operation is then applied to remove the offset l_k, as given by

$$\begin{aligned}(y_k) \bmod \tau &= (d_k + \tau l_k + w_k) \bmod \tau \\ &= (d_k + w_k) \bmod \tau\end{aligned} \qquad (73)$$

Regularized vector precoding is a modification of vector precoding which uses regularized (MMSE) channel inversion in place of simple channel inversion. The transmitted signal vector is then given by

$$\mathbf{x} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \alpha \mathbf{I} \right)^{-1} (\mathbf{d} + \tau l) \qquad (74)$$

where the vector l is chosen to minimize the norm of **x** and $\alpha = K / P_T$. Dirty paper coding techniques based on vector precoding approach the *sum capacity* of the MU-MIMO downlink channel, which is defined as the maximum system throughput achieved by maximizing the sum of the information rates of all the users [48]. **Figure 52** shows the performance comparison of vari-
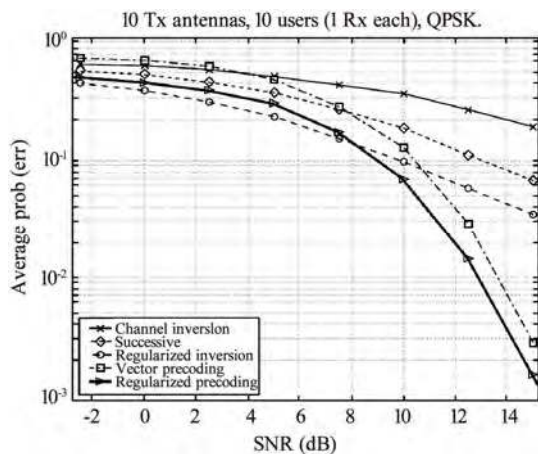

Figure 52. Performance comparison of various channel inversion and dirty paper coding techniques [48].

ous channel inversion and dirty paper coding techniques for uncoded MU-MIMO downlink transmission with $M_T = 10$ BS transmit antennas and $M_R = 10$ single-antenna UEs, using QPSK modulation. The vector precoding techniques clearly outperform the others at high SNR. However, regularized channel inversion performs even better than regularized vector precoding in the low SNR region. A possible reason for the performance loss of regularized vector precoding at low SNR is the use of a finite cubical lattice in this algorithm [48]. Use of different lattice strategies may result in improved performance.

4.2.5. Tomlinson-Harashima Precoding

Tomlinson-Harashima precoding (THP) is a nonlinear precoding technique originally developed for SISO systems for temporal pre-equalization of ISI and is equivalent to moving the decision feedback part of the decision feedback equalizer (DFE) to the transmitter [59]. However, THP can also be applied to MU-MIMO downlink systems for MUI mitigation in the spatial domain.

Two MU-MIMO downlink transmission schemes utilizing THP are described in [57]. The first one called SO THP combines SO and THP to improve performance by eliminating residual MUI. SO THP involves successive BD, reordering of the users and finally, THP. **Figure 53** shows the block diagram of the SO THP system (taken from [57] with a slight notational change). Here **P** represents the combined precoding matrix for all users generated by SO, given by Equation 64. **H** is the channel matrix and **D** represents the combined demodulation (receive filtering) matrix. The lower triangular feedback matrix **B** is generated in the last step and is used for THP. In order to generate matrix **B,** the users are first arranged in the reverse order of precoding and then the lower diagonal
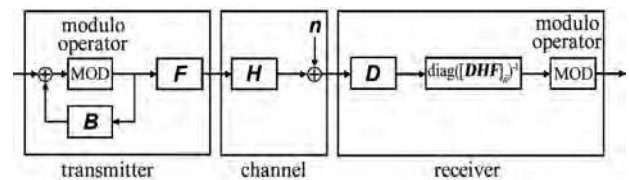

Figure 53. SO THP system block diagram [57].

equivalent combined channel matrix (which includes precoding and demodulation) is calculated, with singular values on the main diagonal. The elements in each row of this matrix are then divided by the corresponding singular values to obtain the feedback matrix **B**. The order in which THP precoding is applied to the users' data streams is opposite to the order in which their precoding matrices are generated. Therefore, THP precoding starts with the data stream of the first user whose precoding matrix **P**$_1$ was generated last.

Use of THP results in increased transmit power and for this reason, a modulo operator is introduced at the transmitter and the receiver so that the constellation points are kept within certain boundaries. At the receiver, each data stream is divided by the corresponding singular value before applying the modulo operator, which ensures that the constellation boundaries remain the same as at the transmitter. Detailed description of SO THP is provided in [60].

The other scheme called MMSE THP [61] combines MMSE precoding and THP to eliminate the MUI below the main diagonal of the equivalent combined channel matrix. MMSE THP is an iterative precoding technique. The users are first arranged according to some optimal ordering criterion and the precoding matrix **P** is calculated column by column starting from the last user K. The i-th column of **P** corresponding to the i-th user is obtained using the coresponding i rows (first i rows for user K) of the channel matrix **H** according to the MMSE criterion, given by

$$\mathbf{P} = \beta \left(\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_{M_T} \right)^{-1} \mathbf{H}^H \quad (75)$$

where

$$\beta = \sqrt{\frac{P_T}{\mathrm{tr}\left(\mathbf{P}\mathbf{x}\mathbf{x}^H \mathbf{P}^H \right)}}, \qquad \alpha = \frac{M_R \sigma_n^2}{P_T}$$

P_T represents the total transmit power, **x** is the data vector to be transmitted and σ_n^2 represents the variance of the zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise. THP is then applied to eliminate the MUI seen by the i-th user from the previous $i-1$ users. **Figure 54** compares the 10% outage capacity of SO THP and MMSE THP schemes for a MU-MIMO downlink system consisting of 4 single-antenna users and 4 BS transmit antennas, denoted as $\{1, 1, 1, 1\} \times 4$ antenna configuration. Results for ZF channel inversion and a $\{2, 2\} \times 4$ TDMA system are also provided.

4.2.6. Successive MMSE Precoding

The Successive MMSE (SMMSE) precoding scheme proposed in [57] addresses the problem of performance degradation associated with MMSE precoding when closely spaced receive antennas are used, like in case of

multi-antenna UEs. SMMSE involves successively calculating the columns of the combined precoding matrix **P**, where each column represents a beamforming vector corresponding to a particular receive antenna.

Consider the system model of Subsection 3.7 where each of the K users is equipped with M_{R_i} receive antennas for $i = 1, \cdots, K$ and $\mathbf{P}_i$ represents the precoding matrix for the i-th user consisting of M_{R_i} columns, each corresponding to a receive antenna. For the j-th receive antenna of the i-th user, the matrix $\bar{\mathbf{H}}_i^{(j)}$ is defined as

$$\bar{\mathbf{H}}_i^{(j)} = \begin{bmatrix} \mathbf{h}_{i,j} & \mathbf{H}_1^T & \cdots & \mathbf{H}_{i-1}^T & \mathbf{H}_{i+1}^T & \cdots & \mathbf{H}_K^T \end{bmatrix}^T \quad (76)$$

where $\mathbf{h}_{i,j}^T$ represents the j-th row of the i-th user's channel matrix $\mathbf{H}_i$. The corresponding column of $\mathbf{P}_i$ is then calculated using the MMSE criterion and is equal to the first column of the matrix

$$\mathbf{P}_{i,j} = \beta \left(\left(\bar{\mathbf{H}}_i^{(j)} \right)^H \bar{\mathbf{H}}_i^{(j)} + \alpha \mathbf{I}_{M_T} \right)^{-1} \left(\bar{\mathbf{H}}_i^{(j)} \right)^H \quad (77)$$

All columns of $\mathbf{P}_i$ are calculated in this manner and this process is repeated for all users to obtain the combined precoding matrix **P**. After precoding, the equivalent combined channel matrix is given by $\mathbf{HP} \in \mathbb{C}^{M_R \times M_R}$ which is block diagonal for high SNR values, resulting in a set of K SU-MIMO channels. Therefore, any SU-MIMO technique e.g. eigen-beamforming for capacity maximization or DET for maximum diversity and array gain can be applied to the i-th user's equivalent channel matrix $\mathbf{H}_i \mathbf{P}_i$. SMMSE precoding has slightly higher complexity than BD [57]. A reduced complexity version called per-user SMMSE (PU-SMMSE) is proposed in [62].
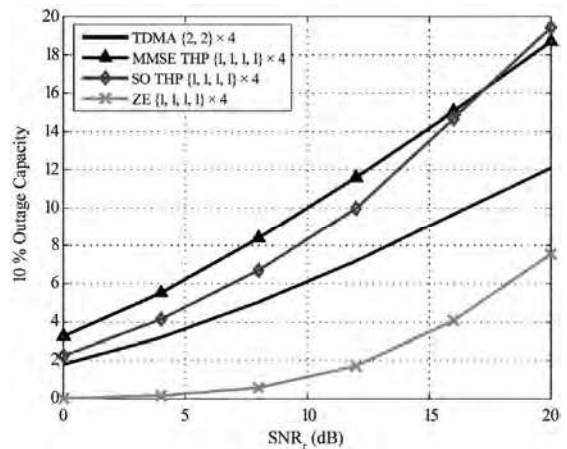
Figure 54. 10% outage capacity of SO THP and MMSE THP for $\{1, 1, 1, 1\} \times 4$ configuration [57].

Figure 55 shows the BER performance comparison of SMMSE, SO THP and BD for {2, 2, 2} × 6 and MMSE THP for {1, 1, 1, 1, 1, 1} × 6 antenna configuration, in a spatially white flat fading channel. These results are based on diversity maximization for the individual users and water-filling is used for power allocation. SMMSE provides the best performance for the case of multi- antenna users while BD surpasses SO THP. **Figure 56** shows the BER performance of SMMSE, SO THP and BD for {1, 1, 2, 2} × 6 and MMSE THP for {1, 1, 1, 1, 1, 1} × 6 configuration. Here SO THP outperforms the others. However, SMMSE performs better than MMSE THP and even SO THP at low SNR.

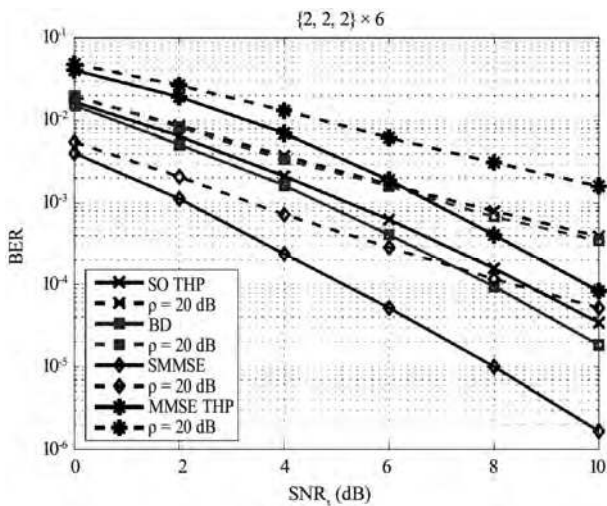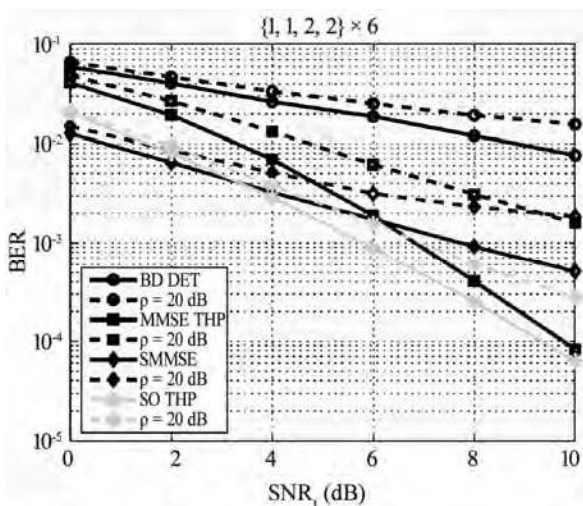**Figure 55. BER performance of SMMSE, SO THP, BD for {2,2,2} × 6 and MMSE THP for {1,1,1,1,1,1} × 6 configuration [57].**

Figure 56. BER performance of SMMSE, SO THP, BD for {1,1,2,2} × 6 and MMSE THP for {1,1,1,1,1,1} × 6 configuration [57].

4.2.7. Iterative Linear MMSE Precoding

Two iterative linear MMSE precoding schemes are discussed in [55] for users with multiple antennas. Consider the system model of **Figure 50** where **P** is the combined precoding matrix at the BS and **V** is the block-diagonal combined decoding matrix consisting of the decoding matrices $\mathbf{V}_i$ of all the users. In case of linear MMSE precoding which minimizes the MSE between $\hat{\mathbf{a}}$ and **a**, $\mathbf{V}_i$ is the linear MMSE receiver for user i and can be estimated locally at the corresponding UE. The first scheme called *direct optimization*, iteratively computes the MMSE solution using a numerical method. The SMMSE solution can be used as an initial guess for the free variables $\mathbf{V}_i$ with $\lambda = 1$ for the free variable $\lambda \in \mathbb{R}$. An iterative process is then used which can lead to a true MMSE solution but not in all cases. The BD solution can also be used as an initial guess but that would result in slower convergence. The other scheme exploits the uplink/downlink duality [63] to obtain the true MMSE solution using an iterative algorithm. The resulting objective function is convex in this case. Detailed description as well as a practically implementable algorithm for this *duality-based* scheme is presented in [64].

The uncoded BER performance of BD, SMMSE, direct optimization and the duality-based scheme is compared in **Figure 57** for {1, 2, 3} × 6 antenna configuration. The bit error rates are averaged over all users. DET is applied for BD and SMMSE while single stream (SS) transmission (consisting of a single data stream per user) is used for direct optimization and the duality-based scheme based on the algorithm from [64]. Independent Rayleigh fading channel perturbed by complex Gaussian noise is considered and QPSK modulation is used for transmission. Both iterative linear MMSE schemes outperform the other two by a large margin. The duality-based scheme shows slightly better performance than
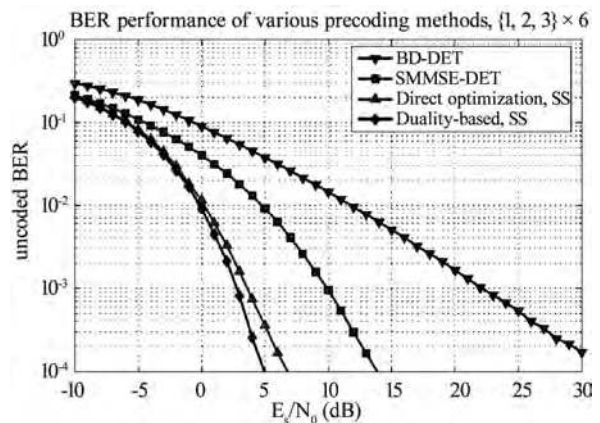

Figure 57. Uncoded BER performance of BD, SMMSE, direct optimization and duality-based scheme [55].

direct optimization. This difference is due to the non-convex objective function used for direct optimization which occasionally causes the optimization routine to undesired minima. BD provides the worst performance because of the zero-forcing constraint.

Figure 58 shows the coded BER performance of these precoding schemes. A rate 1/2 turbo code is used for error correction. OFDM based transmission is considered where the precoding is applied on a per-subcarrier basis. The ITU Vehicular A channel model is used. Direct optimization and the duality-based scheme provide almost identical performance in this case, far better than BD and SMMSE.

4.2.8. Partial CSI Feedback

Transmit precoding for downlink MU-MIMO transmission requires CSI feedback from the users. However, feedback information consisting entirely of the current state of the channel may not be accurate enough in case of rapidly varying channels. Downlink transmission schemes that utilize partial CSI consisting of long-term channel statistics along with some instantaneous channel information like SNR, SINR etc. provide a solution to this problem while reducing the feedback overhead. An interesting MU-MIMO downlink transmission scheme based solely on instantaneous channel norm feedback is proposed in [65]. MU-MIMO configuration with multiple base station (BS) antennas and a single antenna at each UE is considered. The proposed scheme can provide high multiuser diversity gain by optimizing resource allocation at the BS while simply utilizing the instantaneous channel norm feedback from the UEs.

Figure 59 shows the operation of the proposed system at the transmitter (BS). The BS initially transmits orthogonal pilot signals on all transmit antennas which are used by each UE to estimate the received signal energy i.e. the squared norm of the channel vector given by
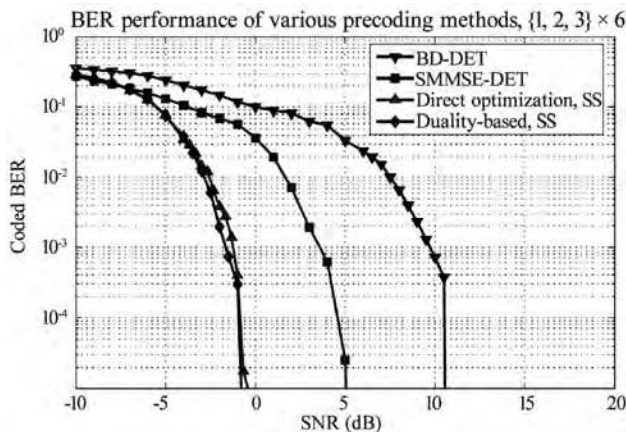

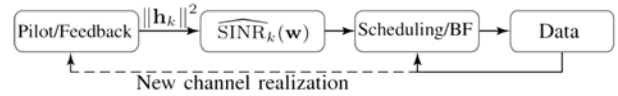**Figure 58. Coded BER performance of BD, SMMSE, direct optimization and duality-based scheme [55].**

Figure 59. System operation at the transmitter [65].

$$\rho_k \triangleq \left\| \mathbf{h}_k \right\|^2 \qquad (78)$$

where $\mathbf{h}_k$ represents the channel vector for the k-th UE from the $\mathcal{I}$ UEs scheduled for transmission. ρ_k is categorized as channel gain information (CGI) in [65]. This quantity is then fed back to the BS. The BS estimates the long-term channel statistics including the channel mean and the channel covariance matrix, defined as

$$\hat{\mathbf{h}}_k \triangleq \mathrm{E}\left\{ \mathbf{h}_k \mid \rho_k \right\} \qquad (79)$$

$$\hat{\mathbf{Q}}_k \triangleq \mathrm{E}\left\{ \mathbf{h}_k \mathbf{h}_k^H \mid \rho_k \right\} \qquad (80)$$

This slow varying statistical information is referred to as channel distribution information (CDI). The CGI feedback along with the CDI is used to estimate the SINR and the optimized beamforming weight vectors for each of the UEs scheduled for transmission. The SINR for user k is estimated as

$$\widehat{\mathrm{SINR}}_k = \frac{\mathbf{w}_k^H \hat{\mathbf{Q}}_k \mathbf{w}_k}{\sum_{i \in \mathcal{I} \setminus \{k\}} \mathbf{w}_i^H \hat{\mathbf{Q}}_k \mathbf{w}_i + \sigma_k^2} \qquad (81)$$

where $\mathbf{w}_k$ is the corresponding beamforming vector, $\hat{\Gamma}_{ki} = \mathbf{w}_i^H \hat{\mathbf{Q}}_k \mathbf{w}_i$ is the MMSE estimate of the signal-to-interference power ratio (SIR) $\Gamma_{ki} = \mathbf{w}_i^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{w}_i$ and σ_k is the AWGN power.

The actual data transmission to the scheduled users then begins. Another set of users can later be scheduled to achieve better fairness and this process goes on until the CGI becomes outdated. At this point, BS transmits the pilot signals again and the whole process is repeated.

4.2.9. Multiuser Scheduling

The BS can only support a limited no. of simultaneous users for MU-MIMO downlink transmission with acceptable performance. The performance degrades in rank-deficient scenarios and also when the users are spatially correlated. In case of multi-antenna users, the no. of users that can be supported simultaneously becomes even lesser. In practical situations, the BS would generally serve a larger no. of users than it can simultaneously support. Therefore, an efficient scheduling algorithm is required to select the group of users that will be spatially multiplexed by the BS at a certain time and frequency. The scheduling algorithm should avoid grouping spatially correlated users and maximize system performance while maintaining fairness toward all users. Fairness

ensures that all users are served including those with weak channels. Otherwise, the BS will transmit to the strong users only and the weaker ones will be ignored [59]. A fair scheduling scheme called the strongest-weakest-normalized-subchannel-first (SWNSF) scheduling proposed in [66] enhances the coverage and capacity of MU-MIMO systems while requiring only a limited amount of feedback.
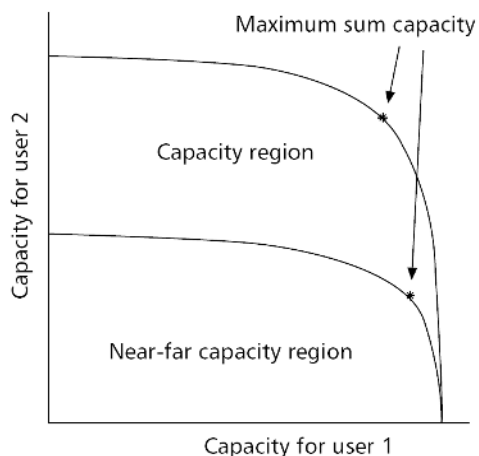
4.3. MU-MIMO Capacity

The maximum capacity of a MU-MIMO system is expressed in terms of the sum capacity (or sum-rate capacity) of the broadcast channel. As mentioned earlier, the sum capacity represents the maximum achievable system throughput and is defined as the maximum sum of the downlink information rates of all the users [48]. **Figure 60** illustrates the capacity region and the sum capacity of a MU-MIMO channel with two users. Clearly, achieving the sum capacity requires some tradeoff between the capacities of the individual users, depending on the shape of the capacity region.

It has been shown in [67] that the sum capacity of a Gaussian MIMO broadcast channel with an arbitrary no. of BS transmit antennas and multi-antenna users, is the saddle-point of a minimax problem and (assuming real-valued signals) is given by [49,67]

$$C_{BC} = \min_{\mathbf{R}_{nn}>0,[\mathbf{R}_{nn}]_{k,k}=\sigma_n^2} \max_{\mathrm{Tr}(\mathbf{R}_{xx})=P} \frac{1}{2}\log\frac{\det\left(\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{R}_{nn}\right)}{\det\left(\mathbf{R}_{nn}\right)}$$

(82)

$$C_{MAC}\left(\mathbf{P};\mathbf{H}^H\right) \triangleq \bigcup_{\{\mathbf{Q}_i \geq 0, \mathrm{Tr}(\mathbf{Q}_i) \leq P_i \forall i\}} \left\{ \begin{array}{l} (R_1,\cdots,R_K): \\ \sum_{i\in S} R_i \leq \frac{1}{2}\log\left|\mathbf{I}+\sum_{i\in S}\mathbf{H}_i^H\mathbf{Q}_i\mathbf{H}_i\right| \quad \forall S \subseteq \{1,\cdots,K\} \end{array} \right\}$$

(83)

Figure 60. The capacity region and sum capacity of a two-user MU-MIMO channel [48].

where **H** is the channel matrix, $\mathbf{R}_{xx}$ is the transmit signal covariance matrix, $\mathbf{R}_{nn}$ is covariance matrix of the additive ZMCSCG noise with variance σ_n^2 and P represents the total transmit power. The sum capacity is obtained by iteratively computing the best transmit covariance matrix $\mathbf{R}_{xx}$ for a given noise covariance and then computing the least favorable noise covariance matrix $\mathbf{R}_{nn}$ for the given transmit covariance. It is also shown that decision-feedback precoding or dirty paper coding is the optimal precoding strategy capable of achieving the sum capacity. In [68], it has been proved that the capacity region of the Gaussian MIMO-BC with single-antenna users is equivalent to the dirty paper coding (DPC) rate region under a certain total transmit power constraint. The DPC achievable rate region for the case of multi-antenna users is formulated in [67] and is also discussed in [69]. Further research work is required on the capacity of fading MIMO broadcast channels.

The capacity of MIMO-MAC channels constitutes a relatively simple problem. The capacity of any MAC is given as the convex closure of the union of rate regions corresponding to every product input distribution $p(u_1)\cdots p(u_K)$ satisfying certain user-by-user power constraints, where $\mathbf{u}_k \in \mathbb{C}^{N\times 1}$ is the k-th user's transmitted signal. However, the convex hull operation is not required for the Gaussian MIMO-MAC and only the Gaussian inputs need to be considered. This can be written as [69]

where $\mathbf{P} = (P_1,\cdots,P_K)$ represents the set of transmit powers corresponding to each of the K users, $|\cdot|$ denotes the determinant and R_i, $\mathbf{Q}_i$ and $\mathbf{H}_i$ represent the rate vector, spatial covariance matrix and the channel matrix respectively for the i-th user. A general expression for the capacity regions of fading MAC channels is also given in [69].

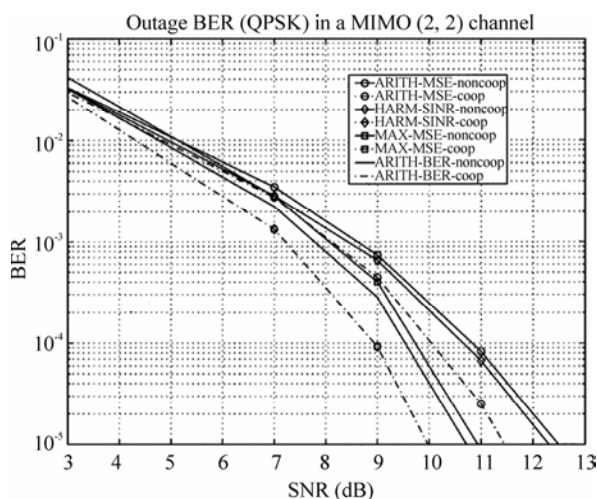## 5. Convex Optimization

Convex optimization methods provide a powerful set of tools for solving optimization problems expressed in convex form. However, most engineering problems are not convex when directly formulated and need to be reformulated in a convex form in order to apply convex optimization. Two methods are commonly used for this reformulation. The first method is to use a change of

variables to obtain an equivalent convex form. The other is to remove some of the constraints so that the problem becomes convex, in such a way that the optimal solution also satisfies the removed constraints. Any problem, once expressed in convex form, can be optimally solved either in closed form using the optimality conditions derived from Lagrange duality theory e.g. Karush-Kuhn-Tucker (KKT) conditions or numerically using iterative algorithms like the interior-point, cutting-plane and ellipsoid methods [70].

In recent years, convex optimization has gained significant importance in optimal joint transceiver design (transmit-receive beamforming) of MIMO systems based on linear precoding and equalization. Various design methods for linear multicarrier SU-MIMO transmit-receive beamformers, based on convex optimization are given in [71]. Two novel low-complexity multilevel water-filling solutions for the MAX-MSE and HARM-SINR criteria are also proposed. The MAX-MSE method minimizes the maximum of the MSEs corresponding to the different substreams whereas HARM-SINR maximizes the harmonic mean of the SINRs. The ARITH-BER method which minimizes the arithmetic mean of the BERs, provides the best average BER performance and is considered as a benchmark in [71]. It is shown that when cooperation among different subcarriers is allowed to improve performance, an exact optimal closed-form solution is obtained in terms of minimizing the average BER which unifies all three optimization criteria mentioned above. In other words, MAX-MSE, HARM-SINR and ARITH-BER provide the same optimal solution for carrier-cooperative schemes.

Figure 61 shows the average BER performance (at 5% outage probability) of ARITH-MSE, HARM-SINR, MAX-MSE and ARITH-BER for 2×2 MIMO configu-

Figure 61. BER performance of ARITH-MSE, HARM-SINR, MAX-MSE and ARITH-BER for 2×2 MIMO configuration using a single substream [71].

ration. The HIPERLAN/2 standard based on OFDM is used for the simulations with frequency selective fading in an indoor NLOS scenario. QPSK modulation is employed and perfect CSI is assumed at the transmitter and the receiver. In the absence of subcarrier cooperation, ARITH-BER provides the best performance followed by MAX-MSE and HARM-SINR respectively. With subcarrier cooperation ARITH-BER, MAX-MSE and HARM-SINR have identical performance which is optimal in the minimum average BER sense. A joint transceiver optimization scheme based on multiplicative Schur-convexity is proposed in [72] for THP precoded MIMO-OFDM systems. This scheme provides better BER performance than the aforementioned linear precoding schemes when the objective function is multiplicatively Schur-convex like in case of ARITH-BER, MAX-MSE or HARM-SINR, and becomes equivalent to the optimal UCD scheme proposed in [37].

Convex optimization is also applicable to downlink beamforming in MU-MIMO systems as mentioned in [70] for the case of single-antenna users. The duality-based iterative MMSE precoding scheme proposed in [64] supports multi-antenna users and uses an iterative algorithm to solve the KKT optimality conditions.

6. Conclusions and Future Research

Open-loop MIMO techniques provide a low-complexity solution for MIMO diversity and SM. STC or SFC e.g. STBC, orthogonal STBC (OSTBC), STTC, SFBC etc. can be used for diversity maximization. These schemes are also well-suited for transmission over high-speed mobile channels where link reliability is the primary concern rather than throughput maximization. Open-loop SM can be implemented by means of a simple V-BLAST system but at the cost of low BER performance. JDM MIMO systems like the CDA-SM-OFDM system which combines CDD and SM, provide better performance. Further research may be carried out to optimize the achievable diversity and multiplexing gain for enhancing system throughput while considering the impact of transmit and receive antenna correlations. The iterative turbo-MIMO systems are also capable of achieving high capacity but are more complex to implement. Turbo-MIMO systems may be improved further by using improved turbo codes and signal constellation shaping [73], improved code interleavers, and stratified processing [74].

Closed-loop MIMO systems like the SVD-based linear transceivers, are capable of achieving the SU-MIMO capacity by transmitting over the channel eigenmodes with optimal water-filling power allocation, provided perfect CSI is available at the transmitter and the receiver. Alternatively, DET can be employed to achieve the maximum diversity and array gain. Closed-loop STC like the closed-loop STBC schemes proposed in [75] which

support more than two transmit antennas, also provide diversity maximization. The GMD-MIMO scheme decomposes the MIMO channel into identical subchannels and attempts to combine MIMO diversity and SM in an optimal manner. MAX-MSE, HARM-SINR and ARITH-BER methods based on convex optimization provide optimal average BER performance for multistream transmission when subcarrier cooperation is allowed. However, the GMD and convex optimization approaches assume the availability of full CSI at the transmitter and the receiver which is not always the case e.g. in rapidly changing mobile channels. Therefore, the performance of these schemes with imperfect feedback needs to be evaluated for practical implementation. Use of convex optimization methods for designing optimal linear transceivers utilizing partial CSI also need to be investigated.

For the MU-MIMO uplink, the LAST-MUD scheme based on V-BLAST detection provides good performance for CDMA systems. The SMMSE MUD provides acceptable performance for MIMO-OFDM systems. The higher complexity Turbo-MUD scheme achieves better performance and can jointly detect the transmit antennas of each multi-antenna user. Use of different turbo detection strategies, joint detection of all users and extension to multicarrier systems may be investigated in future. The GA-assisted MUDs, particularly the TTCM-assisted IGA-MUD provide near-ML detection performance for SDMA-OFDM systems and also perform reasonably well in certain rank-deficient scenarios. The complexity is high but increases slowly with the no. of users as compared to the ML-MUD. GA-assisted MUDs can also incorporate joint channel estimation and symbol detection. Future research work may include extending the GA-assisted MUD schemes to support multi-antenna users and development of more efficient GAs to reduce system complexity. Use of artificial intelligence (AI) techniques like the radial basis function (RBF) based artificial neural networks (ANNs) should also be explored for multiuser detection.

Coming to the MU-MIMO downlink, SMMSE precoding performs reasonably well with manageable complexity. The nonlinear dirty paper coding techniques are capable of achieving the sum capacity of Gaussian multiuser channels with single-antenna users. The iterative linear MMSE techniques like direct optimization and the duality-based scheme which uses convex optimization, provide excellent uncoded and coded BER performance for single stream transmission. Further research is required to develop MU-MIMO downlink transmission schemes capable of supporting SM for individual multi-antenna mobile users. Transmission schemes based on partial CSI which require minimal CSI feedback from the users represent the most suitable choice for practical implementation. Convex optimization tools might be

useful in designing joint transmit-receive beamforming systems less prone to errors resulting from imperfect CSI. Research efforts are also needed for developing efficient multiuser scheduling schemes that maintain fairness to the users while minimizing the loss of system capacity. Determining the sum capacity and capacity regions for fading MU-MIMO downlink channels with multi-antenna users is also an area for future research.

Accurate channel estimation is of prime importance in MIMO communications. Channel estimation errors may result in severe performance degradation. Therefore, research efforts are continuing for further improvements in this domain. A broadband wireless transmission technique called orthogonal frequency- and code-division multiplexing (OFCDM) [76] has recently gained prominence as a better alternative to OFDM. OFCDM readily supports MIMO techniques and extensive research would be required to realize the full potential of MIMO-OFCDM systems.

7. References

[1] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," IEEE Journal on Selected Areas in Communications, Vol. 21, No. 3, pp. 281–302, April 2003.

[2] Y. Jiang, J. Li, and W. 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.

[3] M. Jiang and L. Hanzo, "Multiuser MIMO-OFDM for next-generation wireless systems," Proceedings of the IEEE, Vol. 95, No. 7, pp. 1430–1469, July 2007.

[4] K. W. Park, E. S. Choi, K. H. Chang, and Y. S. Cho, "An MIMO-OFDM technique for high-speed mobile channels," in Proceedings of Vehicular Technology Conference, Vol. 2, pp. 980–983, April 2003.

[5] M. S. Gast, "802.11 wireless networks: The definitive guide," 2nd edition, O'Reilly, April 2005.

[6] Wi-Fi Alliance, "WiFi CERTIFIED™ 802.11n draft 2.0: Longer-range, faster-throughput,multimedia grade WiFi® networks," 2007. http://wi-fi.org/whitepaper_80211n_draft2_technical.php.

[7] FierceBroadbandWireless, "IEEE approves EWC 802.11n as first draft," 24 January 2006. http://www. fiercebroadbandwireless.com/story/ieee-approves-ewc-802-11n-as-first-draft/2006-01-25.

[8] Y. Sun, M. Karkooti, and J. Cavallaro, "High throughput, parallel, scalable LDPC encoder/decoder architecture for OFDM systems," in Proceedings of IEEE Workshop on Circuits and Systems, pp. 225–228, October 2006.

[9] H. N. Niu and C. Ngo, "Diversity and multiplexing switching in 802.11n MIMO systems," in Proceedings of 40th Asilomar Conference on Signals, Systems and Com-

puters (ACSSC '06), pp. 1644–1648, 29 October–1 November 2006.

[10] P. Kim and K. M. Chugg, "Capacity for suboptimal receivers for coded multiple-input multiple-output systems," Vol. 6, No. 9, pp. 3306–3314, September 2007.

[11] S. Abraham, A. Meylan, S. Nanda, "802.11n MAC design and system performance," in Proceedings of IEEE International Conference on Communications (ICC 2005), Vol. 5, pp. 2957–2961, 16–20 May 2005.

[12] Cisco Aironet 1250 Series Access Point Q&A. http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps6973/ps8382/prod_qas0900aecd806b7c82.html.

[13] S. Kim, S.-J. Lee, and S. Choi, "The impact of IEEE 802.11 MAC strategies on multi-hop wireless mesh networks," Wireless Mesh Networks, in Proceedings of 2nd IEEE Workshop on Wireless Mesh Networks (WiMesh 2006), pp. 38–47, 25–28 September 2006.

[14] IEEE Std 802.16e™-2005 and IEEE Std 802.16™-2004/Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004), "IEEE standard for local and metropolitan area networks Part 16: Air Interface for fixed and mobile broadband wireless access systems amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands and Corrigendum 1," 28 February 2006.

[15] IEEE Std 802.16™-2004, "IEEE standard for local and metropolitan area networks Part 16: Air interface for fixed broadband wireless access systems," 1 October 2004.

[16] J. G. Andrews, A. Ghosh, and R. Muhamed, "Fundamentals of WiMAX: Understanding broad-band wireless networking," Prentice Hall, 2007.

[17] I. Kambourov, "MIMO aspects in 802.16e WiMAX OFDMA," WiMAX Tutorial, Siemens PSE MCS RA 2, 22 November 2006.

[18] Nokia Siemens Networks, "Advanced antenna systems for WiMAX." http://www.nokiasiemensnetworks.com.

[19] S. Nanda, R. Walton, J. Ketchum, M. Wallace, and S. Howard, "A high-performance MIMO OFDM wireless LAN," IEEE Communications Magazine, Vol. 43, No. 2, pp. 101–109, February 2005.

[20] Agilent Technologies, "Mobile WiMAX 802.16 Wave 2 Features." http://www.home.agilent.com/agilent/editorial.jspx?action=download&cc=US&lc=eng&ckey=1213544&nid=-536902344.536910932.00&id=1213544.

[21] IEEE C802.16m-07/069, "Draft IEEE 802.16m evaluation methodology document," IEEE 802.16 Broadband Wireless Access Working Group, 2007.

[22] WiMAX Forum®, "Deployment of Mobile WiMAX™ Networks by Operators with Existing 2G & 3G Networks," 2008. http://www.wimaxforum.org/technology/downloads/deployment_of_mobile_wimax.pdf.

[23] C. Ribeiro, "Bringing wireless access to the automobile: A comparison of Wi-Fi, WiMAX, MBWA, and 3G," 21st Computer Science Seminar, 2005.

[24] B. M. Bakmaz, Z. S. Bojković, D. A. Milovanović, and M. R. Bakmaz, "Mobile broadband networking based on IEEE 802.20 standard," in Proceedings of 8th International Conference Telecommunications in Modern Satellite, Cable and Broadcasting Services (TELSIKS 2007), pp. 243–246, 26–28 September 2007.

[25] IEEE Std 802.20™-2008, "IEEE standard for local and metropolitan area networks Part 20: Air interface for mobile broadband wireless access systems supporting vehicular mobility–physical and media access control layer specification," 29 August 2008.

[26] 3GPP TS 36.201 V8.1.0 (2007-11), "LTE physical layer – general description (Release 8)," 3GPP TSG RAN, 2007.

[27] J. Zyren, "Overview of the 3GPP long term evolution physical layer," White Paper, Freescale Semiconductor, Inc., 2007.

[28] 3GPP TR 25.913 V7.3.0 (2006-03), "Requirements for evolved UTRA (E-UTRA) and evolved UTRAN (E-UTRAN) (Release 7)," 3GPP TSG RAN, 2006.

[29] 3GPP TS 36.300 V8.4.0 (2008-03), "Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); Overall description; Stage 2 (Release 8)," 3GPP TSG RAN, 2008.

[30] 3GPP TS 36.211 V8.1.0 (2007-11), "Evolved universal terrestrial radio access (E-UTRA); Physical channels and modulation (Release 8)," 3GPP TSG RAN, 2007.

[31] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in Proceedings of 1998 URSI International Symposium Signals, Systems, and Electronics, pp. 295–300, 29 September–2 October 1998.

[32] S. T. Chung, A. Lozano, and H. C. Huang, "Approaching eigenmode BLAST channel capacity using V-BLAST with rate and power feedback," in Proceedings of IEEE VTC 2001 Fall, Vol. 2, pp. 915–919, 7–11 October 2001.

[33] Q. P. Cai, A. Wilzeck, C. Schindler, S. Paul, and T. Kaiser, "An exemplary comparison of per antenna rate control based MIMO-HSDPA receivers," in Proceedings of 13th European Signal Processing Conference (EUSIPCO 2005), 4–8 September, 2005.

[34] R. Gowrishankar, M. F. Demirkol, and Z. Q. Yun, "Adaptive modulation for MIMO systems and throughput evaluation with realistic channel model," in Proceedings of 2005 International Conference on Wireless Networks, Communications and Mobile Computing, Vol. 2, pp. 851–856, 13–16 June 2005.

[35] M. I. Rahman, S. S. Das, E. de Carvalho, and R. Prasad, "Spatial multiplexing in OFDM systems with cyclic delay diversity," in Proceedings of IEEE Vehicular Technology Conference, pp. 1491–1495, 22–25 April 2007.

[36] H. Busche, A. Vanaev, and H. Rohling, "SVD based MIMO precoding and equalization schemes for realistic channel estimation procedures," Frequenz Journal of RF-Engineering and Telecommunications, Vol. 61, No.

7–8, pp. 146–151, July–August 2007.

[37] Y. Jiang and J. Li, "Uniform channel decomposition for MIMO communications," in Proceedings of 38th Asilomar Conference on Signals, System, and Computers, 7–10 November 2004.

[38] S. Haykin, M. Sellathurai, Y. de Jong, and T. Willink, "Turbo-MIMO for wireless communications," IEEE Communications Magazine, Vol. 42, No. 10, pp. 48–53, October 2004.

[39] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: Theory and experiments," IEEE Transactions on Signal Processing, Vol. 50, No. 10, pp. 2538–2546, October 2002.

[40] D. J. Love, R. W. Heath Jr., 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.

[41] Y. Yuda, K. Hiramatsu, M. Hoshino, and K. Homma, "A study on link adaptation scheme with multiple code words for spectral efficiency improvement on OFDM-MIMO systems," IEICE Transactions on Fundamentals, Vol. E90A, No. 11, pp. 2413–2422, November 2007.

[42] Z. G. Zhou, H. Y. Yi, H. Y. Guo, and J. T. Zhou, "A partial feedback scheme for MIMO systems," in Proceedings of WiCom 2007, pp. 361–364, 21–25 September 2007.

[43] A. Heidari, F. Lahouti, and A. K. Khandani, "Enhancing closed-loop wireless systems through efficient feedback reconstruction," IEEE Transactions on Vehicular Technology, Vol. 56, No. 5, pp. 2941–2953, September 2007.

[44] H. R. Bahrami and T. Le-Ngoc, "MIMO precoding structures for frequency-flat and frequency-selective fading channels," in Proceedings of 1st International Conference on Communications and Electronics (ICCE '06), pp. 193–197, 10–11 October 2006.

[45] M. Tsutsui and H. Seki, "Throughput performance of downlink MIMO transmission with multi-beam selection using a novel codebook," in Proceedings of IEEE VTC 2007-Spring, pp. 476–480, 22–25 April 2007.

[46] K. W. Park and Y. S Cho, "An MIMO-OFDM technique for high-speed mobile channels," IEEE Communications Letters, Vol. 9, No. 7, pp. 604–606, July 2005.

[47] A. A. Hutter, S. Mekrazi, B. N. Getu, and F. Platbrood, "Alamouti-based space-frequency coding for OFDM," Wireless Personal Communications: An International Journal, Vol. 35, No. 1–2, pp. 173–185, October 2005.

[48] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," IEEE Communications Magazine, Vol. 42, No. 10, pp. 60–67, October 2004.

[49] A. Paulraj, R. Nabar, and D. Gore, "Introduction to space-time wireless communications," Cambridge, UK, Cambridge University Press, 2003.

[50] S. Sfar, R. D. Murch, and K. B. Letaief, "Layered space-time multiuser detection over wireless uplink systems," IEEE Transactions on Wireless Communications, Vol. 2, No. 4, July 2003.

[51] V. Stankovic and M. Haardt, "Improved diversity on the uplink of multi-user MIMO systems," in Proceedings of European Conference on Wireless Technology '05, pp. 113–116, 3–4 October 2005.

[52] I. Santamaria, V. Elvira, J. Via, D. Ramirez, J. Perez, J. Ibanez, R. Eickoff, and F. Ellinger, "Optimal MIMO transmission schemes with adaptive antenna combining in the RF path," in Proceedings of 16th European Signal Processing Conference (EUSIPCO 2008), 25–29 August 2008.

[53] N. Veselinovic, T. Matsumoto, and M. Juntti, "Iterative MIMO turbo multiuser detection and equalization for STTrC-coded systems with unknown interference," EURASIP Journal on Wireless Communications and Networking, Vol. 2004, No. 2, pp. 309–321, 2004.

[54] Q. Spencer and M. Haardt, "Capacity and downlink transmission algorithms for a multi-user MIMO channel," in Proceedings of 36th Asilomar Conference on Signals, Systems, and Computers, pp. 1384–1388, November 2002.

[55] B. Bandemer, M. Haardt, and S. Visuri, "Linear MMSE multi-user MIMO downlink precoding for users with multiple antennas," in Proceedings of IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06), pp. 1–5, 11–14 September 2006.

[56] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," IEEE Transactions on Signal Processing, Vol. 52, No. 2, pp. 461–471, February 2004.

[57] V. Stankovic and M. Haardt, "Multi-user MIMO downlink precoding for users with multiple antennas," in Proceedings of 12th Wireless World Research Forum (WWRF), Toronto, ON, Canada, November 2004.

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

[59] V. Stankovic, M. Haardt, and G. D. Galdo, "Efficient multi-user MIMO downlink precoding and scheduling," in Proceedings of 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, pp. 237–240, 13–15 December 2005.

[60] V. Stankovic and M. Haardt, "Successive optimization Tomlinson-Harashima precoding (SO THP) for multi-user MIMO systems," in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Philadelphia, PA, USA, March 2005.

[61] M. Joham, J. Brehmer, and W. Utschick, "MMSE approaches to multiuser spatio-temporal Tomlinson-Harashima precoding," in Proceedings of 5th International ITG Conference on Source and Channel Coding (ITG SCC'04), pp. 387–394, January 2004.

[62] M. Lee and S. K. Oh, "A per-user successive MMSE precoding technique in multiuser MIMO systems," in Proceedings of IEEE VTC 2007-Spring, pp. 2374–2378, 22–25 April 2007.

[63] M. Schubert, S. Shi, 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, pp. 1424–1428, October 2005.

[64] A. Mezghani, M. Joham, R. Hunger, and W. Utschick, "Transceiver design for multi-user MIMO systems," in Proceedings of ITG/IEEE Workshop on Smart Antennas (WSA 2006), March 2006.

[65] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," IEEE Transactions on Signal Processing, Vol. 56, No. 3, March 2008.

[66] C.-J. Chen and L.-C. Wang, "Enhancing coverage and capacity for multiuser MIMO systems by utilizing scheduling," IEEE Transactions on Wireless Communications, Vol. 5, No. 5, May 2006.

[67] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," IEEE Transactions on Information Theory, Vol. 50, No. 9, September 2004.

[68] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," in Proceedings of ISIT 2004, 27 June–2 July 2004.

[69] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," IEEE Journal on Selected Areas in Communications, Vol. 21, No. 5, pp. 684–702, June 2003.

[70] D. P. Palomar, A. Pascual-Iserte, J. M. Cioffi, and M. A. Lagunas, "Convex optimization theory applied to joint transmitter-receiver design in MIMO channels," Space-Time Processing for MIMO Communications, edited by A. B. Gershman and N. D. Sidiropoulos, Chichester, England, Wiley, 2005.

[71] 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, September 2003.

[72] A. A. D'Amico, "Tomlinson-Harashima precoding in MIMO systems: A unified approach to transceiver optimization based on multiplicative Schur-convexity," IEEE Transactions on Signal Processing, Vol. 56, No. 8, August 2008.

[73] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," IEEE Transactions on Communications, Vol. 51, No. 3, pp. 389–399, March 2003.

[74] M. Sellathurai and G. Foschini, "A stratified diagonal layered space-time architecture: Information theoretic and signal processing aspects," IEEE Transactions on Signal Processing, Vol. 51, No. 11, pp. 2943–2954, November 2003.

[75] S. Lambotharan and C. Toker, "Closed-loop space time block coding techniques for OFDM broadband wireless access systems," IEEE Transactions on Consumer Electronics, Vol. 51, No. 3, pp. 765–769, August 2005.

[76] Y. Zhou, T.-S. Ng, J. Wang, K. Higuchi, and M. Sawahashi, "OFCDM: A promising broadband wireless access technique," IEEE Communications Magazine, Vol. 46, No. 3, March 2008.