Multi-antenna
Transceiver Techniques
for 3G and Beyond

Ari Hottinen
Nokia Research Center, Finland

Olav Tirkkonen
Nokia Research Center, Finland

Risto Wichman
Helsinki University of Technology, Finland
This page intentionally left blank
Multi-antenna Transceiver Techniques for 3G and Beyond
This page intentionally left blank
Multi-antenna Transceiver Techniques for 3G and Beyond

Ari Hottinen
Nokia Research Center, Finland

Olav Tirkkonen
Nokia Research Center, Finland

Risto Wichman
Helsinki University of Technology, Finland
# Contents

Preface .......... xi

Acronyms .......... xiii

Part I Introduction

1 Background .......... 3

1.1 Modular System Design .......... 5

1.2 Diversity Techniques in 3G Systems .......... 9

1.2.1 WCDMA Rel99 and Rel4 .......... 9

1.2.2 WCDMA: Recent Releases .......... 13

1.2.3 cdma2000 .......... 15

1.3 GSM/EDGE .......... 15

1.4 Multi-antenna Modems for 3G and Beyond .......... 16

1.4.1 Motivation .......... 16

1.4.2 Examples of Recent Multi-antenna Transmission Methods .......... 17

1.5 Summary .......... 18

2 Diversity Gain, SNR Gain and Rate Increase .......... 21

2.1 Channel Models .......... 21

2.1.1 Multipath Channels .......... 25

2.1.2 Spatial Channels .......... 25
## 2.1.3 MIMO Channel Models

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
</tr>
</tbody>
</table>

## 2.1.4 Polarization Diversity

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
</tr>
</tbody>
</table>

## 2.2 Performance Limits of Transmit Diversity

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
</tr>
</tbody>
</table>

## 2.3 Theoretical MIMO Channel Capacity

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.1 No Channel State Information at Transmitter</td>
<td>34</td>
</tr>
<tr>
<td>2.3.2 Channel State Information at Transmitter</td>
<td>36</td>
</tr>
</tbody>
</table>

## 2.4 MIMO Capacity in Correlated Channels

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
</tr>
</tbody>
</table>

## 2.5 Performance Measures for Closed-loop Transmit Diversity

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
</tr>
</tbody>
</table>

## 2.6 Summary

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
</tr>
</tbody>
</table>

### Part II Open-loop Methods

## 3 Open-loop Concepts: Background

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Delay Diversity</td>
<td>49</td>
</tr>
<tr>
<td>3.2 Implicit Diversity via Phase Modulation</td>
<td>50</td>
</tr>
<tr>
<td>3.3 Code and Time Division Transmit Diversity</td>
<td>53</td>
</tr>
<tr>
<td>3.4 Diversity Transform</td>
<td>53</td>
</tr>
<tr>
<td>3.5 Space—Time Coding</td>
<td>57</td>
</tr>
<tr>
<td>3.5.1 Space—Time Trellis Codes</td>
<td>58</td>
</tr>
<tr>
<td>3.5.2 Space—Time Code Design Criteria</td>
<td>59</td>
</tr>
<tr>
<td>3.5.3 Space—Time Trellis Versus Space—Time Block Coding</td>
<td>61</td>
</tr>
</tbody>
</table>

## 3.6 Space—Time Block Codes

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6.1 Two Tx Antennas: STTD</td>
<td>62</td>
</tr>
<tr>
<td>3.6.2 More Than Two Tx Antennas</td>
<td>63</td>
</tr>
<tr>
<td>3.6.3 Space—Time Block Coding Terminology</td>
<td>66</td>
</tr>
</tbody>
</table>

## 3.7 Non-linear Matrix Modulation

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
</tr>
</tbody>
</table>

## 3.8 Summary

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
</tr>
</tbody>
</table>

## 4 Matrix Modulation: Low SNR Aspects

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Linear Matrix Modulation</td>
<td>69</td>
</tr>
<tr>
<td>4.1.1 Basis Matrices</td>
<td>70</td>
</tr>
<tr>
<td>4.1.2 Diversity and Self-interference</td>
<td>71</td>
</tr>
<tr>
<td>4.2 Examples</td>
<td>74</td>
</tr>
</tbody>
</table>

## 4.3 Heuristic Design Rules at Low SNR

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.1 Frobenius Orthogonality</td>
<td>71</td>
</tr>
<tr>
<td>4.3.2 Minimal Self-interference</td>
<td>72</td>
</tr>
<tr>
<td>4.3.3 Symbol Homogeneity</td>
<td>73</td>
</tr>
<tr>
<td>4.3.4 Maximal Symbolwise Diversity</td>
<td>74</td>
</tr>
<tr>
<td>4.3.5 Maximizing Mutual Information</td>
<td>75</td>
</tr>
</tbody>
</table>

## 4.4 Matched Filtering and Maximum Likelihood Metric

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
</tr>
</tbody>
</table>
CONTENTS

4.4.1 Equivalent Channel Matrix 82
4.4.2 Maximum Likelihood Detection Metric 84
4.4.3 Design Criteria and ML Metric 85

4.5 Mutual Information 86
4.5.1 Information and Interference 86
4.5.2 Expanding Information 87

4.6 Expansion around Diagonal Dominance 92
4.6.1 Diagonal Dominance in Information Measures 93
4.6.2 Diagonal Dominance in Performance Measures 94

4.7 Performance of Examples 95

4.8 Summary 97

5 Increasing Symbol Rate: Quasi-orthogonal Layers 99

5.1 Orthogonal Designs 99
5.1.1 Performance Optimum for Linear Space–Time Codes 100
5.1.2 Consequences of Unitarity and Linearity 100
5.1.3 Construction of Orthogonal Designs 101
5.1.4 Orthogonal Designs and Information 103

5.2 Complexity Issues: Choosing Symbol Rate and Target Tx Diversity 104
5.2.1 Reaching Capacity 104
5.2.2 Linear Detection 105
5.2.3 Choosing the Symbol Rate 107
5.2.4 Choosing Target Tx Diversity Degree 107

5.3 Multimodulation Schemes 108
5.4 Matrix Modulation with Quasi-orthogonal Layers 110
5.4.1 Clifford Basis for Matrix Modulation 110
5.4.2 Idle Directions of a Modulation Matrix 111
5.4.3 Layered Schemes for $N_t = 2, N_r = 2$ 112
5.4.4 Minimal Self-interference 3+1 Layered 4×4 Schemes 114
5.4.5 Minimal Self-interference 2+2 Layered 4×4 Schemes 118
5.4.6 Quasi-orthogonal Schemes for $N_t > 4, N_r = 1$ 119

5.5 Summary 121

6 Receiver Algorithms 123
6.1 Channel Estimation Issues 123
6.2 Maximum Likelihood Detection 124
6.3 Quasi-orthogonality Assisted Maximum Likelihood Detection 127
6.4 Linear Receivers 128
6.5 Iterative Receivers 129
6.6 Joint Decoding and Detection .................................................. 131
6.7 Example: Linear Detection for ABBA ........................................ 132
6.8 Performance ........................................................................... 133
6.9 Summary ................................................................................ 135

7 Matrix Modulation: High SNR Aspects ....................................... 137
7.1 Symmetries of Information and Performance ......................... 140
  7.1.1 Orthogonal Real Symbol Symmetry .................................... 140
  7.1.2 Unitary Left and Right Symmetries .................................... 141
  7.1.3 Examples ........................................................................... 142
7.2 Optimizing Performance with Orthogonal Symbol Rotations .... 143
  7.2.1 Symbol Rotations that Preserve Performance .................... 144
  7.2.2 Symbol Rotations that Change Diversity Degree ................. 146
  7.2.3 Performance-changing, Diversity-preserving Symbol Rotations 147
7.3 Explicit Performance Optima for ABBA .................................. 148
  7.3.1 Scalar Constellation Rotations .......................................... 149
  7.3.2 Matrix Rotations ............................................................... 154
  7.3.3 Performance ....................................................................... 157
7.4 Improved Performance by Extending Block ......................... 159
  7.4.1 Extending ABBA ................................................................. 159
  7.4.2 Extending 3+1 Layered ....................................................... 160
7.5 Comparison of Layered Schemes for Four Tx Antennas .......... 162
  7.5.1 Design Metrics ................................................................. 163
  7.5.2 Performance ................................................................. 164
  7.5.3 Detection Complexity ...................................................... 166
7.6 Weighted and Multimodulation Non-orthogonal Matrix Modulation .......................................................... 167
  7.6.1 Weighted ABBA .............................................................. 167
  7.6.2 Multimodulation ABBA .................................................. 169
7.7 Summary ................................................................................ 170

8 Robust and Practical Open-loop Designs ................................. 171
8.1 Randomized Matrix Modulations ........................................... 172
  8.1.1 Randomization for Non-orthogonal Schemes ...................... 173
  8.1.2 Randomization for Orthogonal Space-Time Block Codes .... 176
8.2 Space-Time Block Code with Rotated Constellations .............. 178
  8.2.1 Transformed STBC-OTD Construction ............................ 181
  8.2.2 Effect on Equivalent Channel Correlation Matrix .............. 182
  8.2.3 Improved Performance in Correlated Channels ................. 183
8.2.4 Combined Diversity Transform and STBC 185
8.3 Performance Evaluation 186
  8.3.1 Performance without Channel Coding 186
  8.3.2 Performance with an Outer Code 187
8.4 Summary 189

9 High-rate Designs for MIMO Systems 191
  9.1 Sets of Frobenius Orthogonal Unitary Matrices 192
    9.1.1 Clifford Basis 192
    9.1.2 Weyl Basis 193
    9.1.3 Hadamard Basis 194
  9.2 Optimizing Rate 2 MIMO-Modulation for $N_t = T = 2$ 196
  9.3 Four Transmit Antennas, Rate 2 197
    9.3.1 Double STTD 198
    9.3.2 Double ABBA 199
    9.3.3 Rate 2 Scheme with Three-symbol Quasi-orthogonality 202
  9.4 Four Transmit Antennas, Rate 3 202
  9.5 Four Transmit Antennas, Rate 4 203
  9.6 The Information Provided by the Schemes 204
  9.7 Summary 205

Part III Closed-loop Methods

10 Closed-loop Methods: Selected Multi-antenna Extensions 209
  10.1 Closed-loop Transmit Diversity in WCDMA 209
    10.1.1 Calculating the Feedback Weight 210
    10.1.2 Quantization and Feedback Signalling 211
    10.1.3 Enhancements 216
  10.2 More than Two Transmit Antennas 219
    10.2.1 Extensions Using Fast Feedback Signalling 220
    10.2.2 Linear Preprocessing Using Long-term (Structured) Feedback 223
    10.2.3 Feedback Signalling and Array Parameterization 225
  10.3 Performance 227
  10.4 Summary 228

11 Analysis of Closed-loop Concepts 233
  11.1 Generalized Feedback Signalling Design 233
  11.2 Analysis of SNR Gain of the Co-phase Algorithm 236
11.3 Analysis of SNR Gain of the Order and Co-phase Algorithm
  11.3.1 Reducing the Amount of Feedback Signalling
  11.3.2 Analysis of Two Transmit Antennas
11.4 SNR Gain in Multipath Rayleigh Fading Channels
11.5 Errors in Feedback Signalling
11.6 Feedback Latency
11.7 Bit-Error Probability
11.8 Transmit Weight Generation for the Order and Co-phase Algorithm
11.9 SNR Gain in Correlated Rayleigh Fading Channels
  11.9.1 Quantization of Long-term Feedback
11.10 Summary
12 Hybrid Closed-loop and Open-loop Methods
  12.1 A Comparison of Open-loop and Closed-loop Systems
  12.2 Transceiver Concepts
    12.2.1 Correlated Channels, Precoding and Matrix Modulation
    12.2.2 Long-term Beam-forming with MISO Transmission
    12.2.3 Long-term Beam-forming with Non-orthogonal Matrix Modulation
  12.3 Performance
  12.4 Space–Time Retransmission
    12.4.1 Retransmission with Symbol Rate One Schemes
    12.4.2 Retransmission with Symbol Rate Two Schemes
    12.4.3 Numerical Example
  12.5 Adaptive Space–Time Modulation Arrangements
  12.6 Summary

Appendix A: Symmetries, Invariants and Inequalities
  A.1 Symmetries and Transformations
  A.2 Unitary Invariants and Determinant Inequalities

Appendix B: Matrix Representations of Clifford Algebras
  B.1 Dimensions of Matrix Representations
  B.2 Clifford Basis of Generic Matrices

References

Index
Preface

The target of this book is to present the core ideas behind a very up-to-date research area involving modulation design for multi-input multi-output (MIMO) wireless channels. Our discussion is aimed at presenting the key principles of different mathematical and engineering approaches that have recently emerged in a number of current and upcoming standards. We restrain ourselves from delving into the physical aspects related to the design of practical antenna elements for mobile or fixed wireless communication units. Rather, we choose to explore and develop multi-antenna transceiver techniques from the signal processing perspective. Such an approach is commonly used when proposing and developing new coding or modulation concepts for wireless systems.

Many of the concepts described herein are aimed at improving data rates, signal quality, capacity or system flexibility. To reach this goal, we adopt matrix-valued modulation alphabets, defined over two orthogonal dimensions, usually referred to as space and time. The space-dimension is realized by using multiple transmit and receive antennas, and involve multi-antenna transceiver structures. Such multi-antenna techniques are generally considered as the most promising avenue for significantly increasing the bandwidth efficiency of wireless data transmission systems. In MIMO systems, multiple antennas are deployed both at the transmitter and the receiver. In ideal situations, this allows signalling over several parallel channels between the transmitter and receiver. These channels can be separated using signal processing means, provided that the channels are sufficiently different. In MISO (multiple-input single-output) systems, the receiver has only one antenna, and the multiple transmit antennas are used for transmit diversity.

This book presents the key aspects of multiple antenna transceiver techniques for evolving 3G systems and beyond. MIMO and MISO (transmit diversity) techniques are explained in a common setting. A special emphasis is put on combining theoretical understanding with engineering applicability.
In particular, the book covers linear processing transmit diversity methods with and without side information at the transmitter, including a description of the current transmit diversity concepts in the WCDMA and cdma2000 standards, as well as promising MIMO concepts, crucial for future high data-rate systems. Furthermore, examples of high throughput, low complexity matrix modulation schemes will be provided, when signalling without side information (open loop concepts). The theory of linear matrix modulations will be developed, and optimal non-orthogonal high throughput schemes will be constructed, both for MIMO and MISO systems.

Performance may be further improved by feedback from receiver to transmitter. The corresponding closed-loop modes in the current 3GPP specifications will be discussed, along with their extensions for more than two transmit antennas. In addition, feedback signalling for MIMO channels will be addressed, as well as optimal quantization methods of the feedback messages. Finally, hybrid schemes are constructed, where the amount of overhead due to feedback is reduced by combining open-loop transmission with closed-loop signalling.

We would like to express our gratitude to a number of colleagues who have helped in preparing this work. We thank Drs Jyri Hämäläinen, Rinat Kashaev and Jussi Vesma and Mr Mikko Kokkonen for fruitful collaboration related to the subject matter of this book. Numerous discussions with colleagues at Nokia Research Center are also acknowledged. Drs Nikolai Nedefov and Kari Kalliojärvi provided a number of constructive comments that enabled us to improve the readability of the text. Financial support from Nokia Foundation is also gratefully acknowledged. A large part of the results documented here have been developed at Nokia Research Center in recent years with support from Dr Jorma Lilleberg at Nokia Mobile Phones. Finally, we appreciate the seemingly unlimited patience of our respective home troops.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMC</td>
<td>Adaptive Modulation and Coding</td>
</tr>
<tr>
<td>AOA</td>
<td>Angle Of Arrival</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic Repeat reQuest</td>
</tr>
<tr>
<td>AS</td>
<td>Azimuth Spread</td>
</tr>
<tr>
<td>ASTMA</td>
<td>Adaptive Space–Time Modulation Arrangement</td>
</tr>
<tr>
<td>BEP</td>
<td>Bit-Error Probability</td>
</tr>
<tr>
<td>BER</td>
<td>Bit-Error Rate</td>
</tr>
<tr>
<td>BICM</td>
<td>Bit Interleaved Coded Modulation</td>
</tr>
<tr>
<td>BLAST</td>
<td>Bell Laboratories Layered Space-Time architecture</td>
</tr>
<tr>
<td>BPSK</td>
<td>Binary PSK</td>
</tr>
<tr>
<td>BS</td>
<td>Base Station</td>
</tr>
<tr>
<td>CDMA</td>
<td>Code Division Multiple Access</td>
</tr>
<tr>
<td>CDTD</td>
<td>Code Division Transmit Diversity</td>
</tr>
<tr>
<td>CL</td>
<td>Closed-Loop</td>
</tr>
<tr>
<td>CPICH</td>
<td>Common Pilot Channel</td>
</tr>
<tr>
<td>CQI</td>
<td>Channel Quality Indicator</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>DOT</td>
<td>Direction Of Transmission</td>
</tr>
<tr>
<td>DSTTD</td>
<td>Double STTD</td>
</tr>
<tr>
<td>D-BLAST</td>
<td>Diagonal BLAST</td>
</tr>
</tbody>
</table>
Acronyms

EVD  EigenValue Decomposition
FB  Feedback
FBI  Feedback Back Indicator
FEC  Forward Error Coding
FER  Frame Error Rate
FSM  Feedback Signalling Message
FCS  Fast Cell Selection
FO  Frobenius Orthogonality
FDD  Frequency Division Duplex
HSDPA  High-Speed Downlink Packet Access
HS-DSCH  High-Speed Downlink Shared CHannel
i.i.d.  independent and identically distributed
IC  Interference Cancellation
ID  IDentification
ISI  Inter-Symbol Interference
LAN  Local Area Network
LLR  Log-Likelihood Ratio
LMMSE  Linear MMSE
LOS  Line Of Sight
MRC  Maximal Ratio Combining
MF  Matched Filter
MIMO  Multiple-Input Multiple-Output
MISO  Multiple-Input Single-Output
ML  Maximum Likelihood
MMSE  Minimal Mean-Square Estimate
MS  Mobile Station
MSD  Maximal Symbolwise Diversity
Node-B  Base station
OFDM  Orthogonal Frequency Division Multiplexing
OD  Orthogonal Design
OL  Open Loop
OSIC  Ordered SIC
OTD  Orthogonal Transmit Diversity
PAM  Pulse Amplitude Modulation
PAS  Power Azimuth Spread
PHOP  Phase Hopping
<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIC</td>
<td>Parallel Interference Cancellation</td>
</tr>
<tr>
<td>PSK</td>
<td>Phase-Shift Keying</td>
</tr>
<tr>
<td>PSTD</td>
<td>Phase Sweep Transmit Diversity</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature Amplitude Modulation</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase Shift Keying</td>
</tr>
<tr>
<td>QOML</td>
<td>Quasi-Orthogonality assisted ML</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RH</td>
<td>Radon–Hurwitz</td>
</tr>
<tr>
<td>RHO</td>
<td>RH Orthogonal</td>
</tr>
<tr>
<td>Rx</td>
<td>Receiver</td>
</tr>
<tr>
<td>SIC</td>
<td>Successive Interference Cancellation</td>
</tr>
<tr>
<td>SIMO</td>
<td>Single-Input Multiple-Output</td>
</tr>
<tr>
<td>SISO</td>
<td>Single-Input Single-Output</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SH</td>
<td>Symbol Homogeneity</td>
</tr>
<tr>
<td>SSDT</td>
<td>Site Selection Diversity Transmission</td>
</tr>
<tr>
<td>STBC</td>
<td>Space–Time Block Code</td>
</tr>
<tr>
<td>STC</td>
<td>Space–Time Code</td>
</tr>
<tr>
<td>STTC</td>
<td>Space–Time Trellis Code</td>
</tr>
<tr>
<td>STD</td>
<td>Selective Transmit Diversity</td>
</tr>
<tr>
<td>STTD</td>
<td>Space–Time Transmit Diversity</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>TDD</td>
<td>Time Division Duplexing</td>
</tr>
<tr>
<td>TDTD</td>
<td>Time Division Transmit Diversity</td>
</tr>
<tr>
<td>TrSTBC-OTD</td>
<td>Transformed STBC-OTD</td>
</tr>
<tr>
<td>TSTD</td>
<td>Time Switched Transmit Diversity</td>
</tr>
<tr>
<td>TTI</td>
<td>Transport Time Interval</td>
</tr>
<tr>
<td>Tx</td>
<td>Transmitter</td>
</tr>
<tr>
<td>UB</td>
<td>Upper Bound or Union Bound</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
<tr>
<td>ULA</td>
<td>Uniform Linear Array</td>
</tr>
<tr>
<td>UTRA</td>
<td>Universal Terrestrial Radio Access</td>
</tr>
<tr>
<td>V–BLAST</td>
<td>Vertical BLAST</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband CDMA</td>
</tr>
<tr>
<td>XTRM</td>
<td>eXTRemuM</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>2G</td>
<td>Second Generation</td>
</tr>
<tr>
<td>3G</td>
<td>Third Generation</td>
</tr>
<tr>
<td>3GPP</td>
<td>3G Partnership Project</td>
</tr>
<tr>
<td>4G</td>
<td>Fourth Generation</td>
</tr>
</tbody>
</table>
Part I

Introduction
The research and standardization of 3rd generation (3G) wireless systems has been ongoing for about a decade. Initially standardization work was carried out in national standardization bodies. Different national bodies, major corporations and several research institutes initiated the 3rd Generation Partnership Project (3GPP), which has been active since the end of 1998 (see http://www.3gpp.org). The first WCDMA based 3G solutions developed by 3GPP are Release '99 and Release 4 [1,2]. These constitute the basis for the first commercial WCDMA systems. The most recent update, Release 5, is optimized for high-speed downlink packet access [3].

3GPP is continuously enhancing the WCDMA specifications to further enhance the performance of 3G systems. It is anticipated that some future physical layer standard release will contain further enhancements related to multi-antenna transmission techniques. In particular, future releases are likely to include support to novel transmit diversity concepts and high-rate transmission schemes that explicitly support multiple transmit antennas. A similar project, 3GPP2, is currently ongoing in the USA in an attempt to develop and define a cdma2000-based solution for 3G [4–6] (see http://www.3gpp2.org). Many of the recent technical solutions in the cdma2000 family of systems are similar to those in the WCDMA system, and certain solutions have been developed hand-in-hand. For example, both systems include support for multi-antenna transmission.

Compared to 2G systems such as GSM or IS-95 [4], 3G systems provide enhanced services and significantly higher data rates. In addition, the overall capacity of 3G systems has significantly increased beyond 2G systems through the adoption of latest technological achievements. Many of these technological achievements were discovered in the 1990s after 2G systems were already operational. Examples of such techniques include major leaps in coding theory via the invention of turbo codes [7,8]. Turbo codes have been demonstrated to approach the theoretical channel capacity limit, as derived by Shannon in the 1940s [9]. Eventually they found a way to both WCDMA and cdma2000 systems. Another important development has been the introduction of novel multi-antenna transmission techniques [10].
Multi-antenna transmission techniques provide transmit diversity to 3G systems and enable a significant increase in downlink capacity. Transmit diversity is not an entirely new concept, although significant breakthroughs have been made in recent years. Simple space-diversity techniques are already applied in 2G systems, such as GSM or IS-95. As an example, diversity reception using multiple receive antennas is a mature technology which is often applied in 2G base stations to improve uplink coverage. However, due to implementation costs and space constraints, receive diversity methods are not as applicable for mobile handsets. For this reason, the first release of the 3G wideband CDMA standard [11] applies transmit diversity schemes at base stations to improve downlink reliability. These schemes are specifically designed for two co-channel antennas [10, 12, 13]. Recently, such schemes have been suggested for more than two Tx antennas [14]. These transmit diversity solutions mitigate the need to deploy multiple antenna in mobile handsets solely for the purpose to increase diversity. They are in many respects even simpler than well-known downlink beam-forming concepts, in which directional beam patterns are formed towards the desired user. The availability of such simple approaches for downlink capacity enhancement is pleasing, since many of the proposed services, like wireless web browsing, are likely to be downlink-intensive.

With multiple antennas at the base station, and one antenna at the mobile, the uplink is a SIMO (single-input multiple-output) radio channel, whereas the downlink is MISO (multiple-input single-output). In a MIMO (multiple-input multiple-output) system, one has multiple antennas both at transmitter and receiver. In fading channels, these create respectively Tx- and Rx-diversity. The attractive characteristic of a MIMO channel is that it may be used to increase the data rate by transmitting multiple streams simultaneously using different spatial channels. Loosely speaking, Rx-diversity is used to separate these multiple streams from each other, while Tx-diversity may be used to improve performance. Thus, in a sense, a MIMO multi-antenna channel increases the effective bandwidth of a wireless channel. This challenges the conventional thinking which suggests that extremely high data rates either require extremely wide frequency bandwidth and/or extremely high transmit power. On the other hand, even if multi-antenna modems may avoid some of these problems, they tend to require rather heavy signal processing in an attempt to gain access to this projected spatial bandwidth. Hence, the complexity of the terminals and network elements will increase, when compared to traditional modems offering similar data rates with a larger bandwidth. Despite the implementation cost, Moore’s Law suggests that at some point in the future multi-antenna techniques may also enable high data rate services for those that do not have bandwidth in abundance. On the other hand, if extremely high data rates are not required, the increased fading resistance (diversity) inherent in many multi-antenna transceiver concepts can be used to increase system capacity or coverage.

Although several transmit diversity and general multiple-input multiple-output (MIMO) transmission techniques have been known for some twenty years [15], the theoretical capacity results developed in [16, 17] revitalized the research area. Essentially, it was shown in [16, 17] that under certain conditions the capacity increases linearly with \( \min(N_t, N_r) \), where \( N_t \) is the number of deployed transmit antennas and \( N_r \) is the number of receive antennas. Communication theorists and engineers have since developed coding and modulation concepts that realize a large portion of this gain. In light of these recent results, multi-antenna transmission and reception techniques are currently seen as the most promising avenue for significantly increasing the capacity and spectral efficiency of wireless systems. Novel bandwidth efficient multi-antenna modulation concepts can be used to enable power
efficient signalling at data rates beyond 10 Mbps using the WCDMA 3.84 MHz Chip rate occupying 5 MHz bandwidth. Recently, MIMO concepts have entered the standardization discussion of future releases of 3G standards.

This book deals with the impact of these multi-antenna techniques on the design of 3G and future mobile communication systems. Applications to WCDMA downlink have been our primary motivation, but the considered methods are not restricted to downlink direction nor WCDMA system evolution. Most of the developed solutions are general and applicable to generic 3G and 4G systems, as well as for example wireless LANs as long as multiple transceivers are deployed. The approach will mainly be one of baseband signal processing. The number of transmit and receive antennas is assumed to be equal to the number of Tx and Rx radio frequency (RF) front ends. The problem of multiple physical antennas coupled to a smaller number of RF front ends, interesting as it is, is not considered in this book.

Transmission Resources: The methods discussed in this book divide transmission resources into two complementary categories. Often these will be called “space” and “time”. In the spatial dimension, the discrete unit is referred to as an “antenna”, whereas in the temporal dimension, the discrete unit is referred to as “symbol period” or “symbol epoch”. The essential difference between these two dimensions is that the “time” dimension is substantially orthogonal, whereas the “space” direction is typically non-orthogonal—symbols transmitted simultaneously from two co-channel antennas typically interfere.

Substantially Orthogonal: Time, Frequency, Code: Instead of (or in parallel to) time division multiplexing, the substantially orthogonal “time” dimension may indicate frequency division multiplexing or code division multiplexing. To keep the near orthogonality of the “time” dimension, inter-symbol interference in multipath channels should be mitigated. This means proper equalization when using time-division multiplexing, or proper guard intervals when using, for example, orthogonal frequency division multiplexing (OFDM).

Spatially Separate, or Polarization: The “space” dimension may indicate antennas operating in spatially separate locations and/or in different polarizations. Due to different local scattering environments, sufficiently separated antenna elements provide almost independent fading channels. What sufficient means, depends on the environment. In rural macro cells, separations of many wavelengths may be required to de-correlate antennas, whereas in indoor environments a half-wavelength separation may be sufficient. For polarization, the cross-polarization coupling ratio determines whether polarizations provide diversity, or whether they provide near-orthogonal parallel channels.

1.1 MODULAR SYSTEM DESIGN

It has gradually become clear that one of the most efficient ways to answer the demand of ever increasing spectrum efficiency is to deploy multiple antennas at the transmitter and receiver end of a radio link. In this book we advocate a modular design solution in which the number of antennas is increased so that the impact to other system parts is minimized. The (single user) MIMO/MISO transmission chain considered in this book is as follows.

1. The source emits information bits $b$ at a source rate $R_{source}$. From the source, the information bits go to an encoding and interleaving circuit. Typically, it includes a
binary encoder, with code rate $R_c$ and an interleaver with interleaver depth $N_1$. The output is a stream (or a vector) of coded bits $c$.

2. The coded bits go to a modulator, which maps $M$ bits into complex modulation vector $x$.

3. The stream of symbols goes into a space–time modulator, which takes sequences of $R_s T$ symbols and maps them to a $T \times N_b$ matrix $X$, preparing the symbols for transmission over $T$ substantially orthogonal transmission resources (time, subcarrier, spreading codes) and $N_b$ beams. The symbol rate of the space–time modulator is $R_s$ symbols per transmission resource. The simplest form of a space–time modulator is a serial-to-parallel multiplexer, which constructs a vector modulation.

4. The output of the space–time modulator is conveyed to the beam-forming circuit. This constructs $N_b$ beams out of $N_t$ spatial transmission resources (antennas, polarizations). The action of the beam-former can be described by a $N_b \times N_t$ matrix $W$.

5. Finally, the signals to be transmitted on the beams are upconverted to the radio frequency, and transmitted.

A diagram of the transmission chain can be found in Figure 1.1. This transmission chain is general enough to cover most MISO/MIMO transmission schemes discussed in the literature, with or without channel information at the transmitter:

- For feedback modes [2, 12], the beam-forming matrix $W$ constructs a single beam, which is used to weight the transmission using $N_t$ antennas. The beam-forming matrix is determined using feedback information conveyed by the receiver to the transmitter, or estimated at the transmitter from reverse direction signalling. The space–time modulator is trivial, $X$ collapses to a $1 \times 1$ matrix for each modulation symbol.

- For space–time block codes [10, 18], and more generally for linear MISO/MIMO modulation, the space–time modulator maps the modulation symbols $x$ to a space–time code matrix $X$ in a linear fashion. The space–time block length is $T$. The beam-forming matrix is trivial, $W = I_{N_t}$, when used as a pure open-loop concept.

- For space–time trellis codes [19], $T = 1$, the space–time modulator is a vector modulator, and the beam-forming matrix is trivial. Thus $XW$ outputs sequences of $R_s = N_t$ symbols from the $N_t$ transmit antennas during each symbol period. In the coding

---

**Fig. 1.1:** Multi-antenna transmission chain.
circuit, the interleaver is trivial, and the encoder and modulator may be thought of as a joint trellis-coded modulator.

- **Multistream MIMO systems**, or spatial multiplexing, in the spirit of [15, 16, 20, 21], differ from space–time trellis codes in that the coding and interleaving circuit includes de-multiplexing and multiplexing units with the effect that the information stream is split into multiple parallel streams (or layers), which are independently encoded and transmitted simultaneously from the multiple antennas. The number of these streams is $R_s$, the symbol rate. Depending on the multiplexing units, one has diagonally (DBLAST) [20] or vertically (VBLAST) [15, 21] encoded vector modulation, the latter either with horizontal or vertical coding.

- For **unitary space–time modulations** [22], the modulator is trivial, and it outputs just a sequence of bits. The space–time modulator takes this sequence and maps it in a non-linear fashion to the matrix $X$.

- **Randomization techniques** apply a (pseudo)randomly chosen beam-forming matrix $W$. Some have a trivial space–time modulator $X$, and apply phase or antenna hopping [23, 24], which can be described by choosing different $1 \times N_t$ beam-forming vectors $W$. Some have a non-trivial space–time modulator $X$, and the beam-forming matrix performs multibeam-forming by applying antenna permutations or multi-antenna hopping.

- **Hybrid open/closed-loop schemes** [25–29] resemble randomizing techniques with non-trivial $X$, except that $W$ is chosen based on feedback.

As indicated above, various parts of the transmission chain can be merged and jointly optimized. When merging the encoder and the modulator, one may discuss trellis-coded modulation. Similarly, merging the symbol modulator and the space–time modulator, one may discuss unitary (non-linear) space–time modulations. The main paradigmatic split is between jointly or separate design of channel coding and space–time modulation. This is the difference between space–time trellis coding and the modular approach in 3G standards [1], where space–time block codes are applied.

In this book, the modular approach of the 3G standards is adopted. In this approach, the parts responsible for exploiting the spatial structure of the channel are separate modules, $X$ and $W$, that are designed separately from binary channel codes. These two modules are joined with the coding module by an interleaver, the design of which may take into account possible periodicities produced by $X$, $W$ and the modulator, when systematic bits of the channel coders are mapped.

The reasons for favouring a modular approach are

- **Flexibility.** Several rates can be treated with the same space–time modulator. The principle of digital convergence suggests that a multitude of different services converge to be operated by the same device. In the case of mobile communication, speech and different kinds of streaming/packet data services operate at widely differing data rates and QoS requirements. In a modular system, different services with different rates may apply the same transmit diversity scheme without losing too much in performance. The rate-matching algorithms adopted in WCDMA and cdma2000 systems enable power-efficient transmission regardless of the packet size delivered to the encoding.
chain. The space–time modulation methods should not impact on the performance of the rate matching procedure.

- **Robustness against change of channel conditions.** This is the underlying reason why bit-interleaved coded modulation is favourable in mobile communication. A mobile communication link may operate under a wide variety of channel conditions, from completely specular line-of-sight to rich scattering environments with severe multi-path and deep fades, with corresponding fluctuations in SNR. Bit-interleaved coded modulation (BICM) schemes have a tendency to be more robust against changes in channel conditions [30].

- **Implementation simplicity.** Hardware implementation complexity should not be underestimated. In a modular system, the circuitry is divided into smaller pieces that can be designed separately.

For these reasons, the bulk of this book is dedicated to the modules $X$ and $W$, which form the core of exploiting the spatial structure provided by a MISO/MIMO channel. Clearly, joint optimization of beam-forming matrix $W$ and the symbol matrix $X$ for a given multiple-access or modulation method is a challenging problem. In practice, when multiple users or services share the same channel, the corresponding matrices should be jointly defined to mitigate interference between different users and to maximize system capacity. Recent multiuser scheduling and power allocation solutions involving one or multiple transmit antennas [31–33] lead to orthogonal multiple-access signaling concepts. This allows to mitigate the signal processing aspects of a multiuser communication system, while increasing the complexity of control and access protocols. In decentralized systems, e.g. in ad-hoc networks, the control problems are even more challenging and new paradigms may be needed [34–36]. For ease of exposition, we only address in this book the signal processing aspects, and consider a signal model incorporating one point-to-point wireless link.

**Signal Model:** The baseband signal model corresponding to these modules is formulated concisely as follows. The coding, interleaving and multiplexing unit is left out from the signal model. The space–time modulator matrix $X$ transmits $R_S T$ complex modulation symbols over $N_b$ beams during a block of $T$ symbol epochs. The number of parallel streams $R_s$ is defined as the (average) number of complex symbols transmitted per symbol epoch, i.e. the symbol rate. In a space–time modulator with block of length $T$, altogether $R_S T$ complex symbols are thus transmitted. The beam-forming unit prepares the $N_b$ beams for transmission from $N_t$ antennas.

Much of this book will suppress delay spread considerations in favour of concentrating on the spatial structure of the channel. In a one-path channel the MIMO/MISO signal model covers one transmission block:

$$
Y_{T \times N_r} = X_{T \times N_b} \cdot W_{N_b \times N_t} \cdot H_{N_t \times N_r} + \text{noise}_{T \times N_r}
$$

(1.1)

Here $N_r$ is the number of receive antennas, $Y$ is the $T \times N_r$ matrix of received signals, $X$ is the $T \times N_b$ transmission matrix (the space–time modulation) and $W$ is the $N_b \times N_t$ beam-forming matrix. The channel $H$ is a matrix where each column is a channel vector.
from the multiple transmit antennas to one receive antenna,

$$H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & h_{22} & \cdots & h_{2N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_t1} & h_{N_t2} & \cdots & h_{N_tN_r} \end{bmatrix},$$

(1.2)
as depicted in Figure 1.2. In a MISO system, $H$ is a column vector $h$. When a multi-path channel model is considered, the channel matrix $H$ is extended to cover the multipath components, and $X$ is extended to cover multiple transmission blocks.

**BICM in WCDMA:** The WCDMA [1,3,11] system supports three transmit diversity concepts on dedicated traffic channels. The open loop scheme is the $2 \times 2$ space–time block code proposed by [10], known as STTD (Space–Time Transmit Diversity) in WCDMA. The closed loop schemes, known as Mode 1 and 2, apply a 2- or 4-bit quantization for the feedback weight, respectively, to parameterize the matrix $W$. There are four channel coding options: rate 1/3 and 1/2 convolutional and turbo codes (for speech and data services, respectively). Between the channel code and interleaver, there is a rate-matching circuit, which performs additional puncturing/padding so that transmitted data packets are made to fit into a radio frame.

From the point of view of coding theory, the space–time coding scheme in WCDMA can thus be interpreted as bit-interleaved coded space–time modulation, in the spirit of [30]. This is a simple and efficient way of exploiting two transmit antennas. In [37], it was proven that STTD reaches channel capacity when $N_t = 2, N_r = 1$, if the channels are independent and identically distributed (i.i.d.) Rayleigh block fading, with a channel coherence time longer than the delay of $T = 2$ symbol periods required for transmitting STTD. Inherent capacity is turned into reliable signalling by efficient concatenated binary codes. Similarly, the 2- and 3-bit quantization concepts of Modes 1 and 2 provide the optimal increase in SNR [38] for the number of feedback bits used.

**1.2 DIVERSITY TECHNIQUES IN 3G SYSTEMS**

To increase downlink and uplink capacity, the Universal Terrestrial Radio Access (UTRA) WCDMA system, cdma2000 and GSM evolutions incorporate various diversity techniques. These mitigate the effects of disadvantaged channel conditions, by enabling access to multiple (independent) channel realizations.

**1.2.1 WCDMA Rel99 and Rel4**

Diversity benefit can be obtained by a number of different technical solutions in 3G systems. WCDMA Release '99 and Release 4 support the following diversity techniques:

- multipath diversity
- time diversity (Automatic Repeat ReQuest)
- Rx diversity, using multiple receive antennas
Fig. 1.2: MIMO model with $N_t$ transmit and $N_r$ receive antennas.

- Tx diversity, with one open and two closed loop solutions
- macro diversity (soft handover)
- Site Selection Diversity Transmission (SSDT)

1.2.1.1 Multipath Diversity  A wideband channel enables the receiver to resolve a large number of multipath components. This increases diversity in channels with sufficient delay spread and boosts performance when compared to narrowband CDMA or TDMA systems. With multipath diversity the independent signal copies are sampled in time domain, with the assumption that signals arriving at different delays do not fully correlate. Diversity can be captured using the RAKE or some other linear detector or channel equalizer. Multipath diversity is not available in all environments. For example, a single tap channel is typical in indoor environments.

1.2.1.2 Macro Diversity  Macro diversity can be used both in uplink and downlink to combine signals transmitted to or received from multiple base stations (Node Bs). In uplink signal copies are sampled from spatially separate sensors or antennas with independent fading. In downlink, multiple copies of the same signal are transmitted from spatially separate source locations, again to result in independent fading.

1.2.1.3 SSDT  The specification also includes an additional downlink macro diversity option, known as Site Selection Diversity Transmission (SSDT). In SSDT the UE maintains a list of active set cells, and determines the “primary” cell. All other cells are labelled as “non-primary”. Each cell is assigned a temporary identification (ID) and UE periodically informs a primary cell ID to the base stations using an uplink signalling field. The dedicated channel in non-primary cells turn off the transmission power. The primary cell ID can be
signalled 1-5 times per 10 ms frame, with different signalling formats. SSDT is activated by higher-layer signalling. In addition, the cell ID assignment is all carried out by higher layer signalling. Site selection can thus be carried out without network intervention, differentiating it from conventional hard handover.

The main objective of SSDT is to reduce interference due to multiple transmissions in a soft handover mode. It is essentially an antenna selection concept combined with efficient power allocation, and together these increase both diversity and power efficiency.

### 1.2.1.4 Time Diversity

UTRA Release '99 supports Type I Automatic Repeat Request (ARQ) protocol, where erroneous frames are discarded and the frame is repeated later on. If the frame is repeated after a sufficiently long time interval (beyond channel coherence time), ARQ provides time diversity. Time diversity can also be exploited via the combined use of interleaving and forward error correction (FEC) codes.

### 1.2.1.5 Receive Antenna Diversity

Multiple receive antennas can be implemented in both Node Bs and UE to capture spatial receive diversity (Rx diversity). Rx diversity can easily be utilized in the base station to improve uplink capacity or coverage. Often, however, due to cost and space considerations manufacturers tend to avoid implementing multiple antennas in the smallest handheld terminals. Nevertheless, Rx diversity is one of the most efficient diversity techniques. Moreover, in addition to the diversity benefit, the received aggregated signal power is theoretically \(N_r\) fold, when compared to single-antenna reception.

### 1.2.1.6 Transmit Diversity

A significant effort has been devoted in 3GPP to develop efficient transmit diversity solutions to enhance downlink capacity. Transmit diversity methods also provide space diversity for terminals with only one receive antenna, and in that sense retain the complexity at the base station. Typically, in a 3G base station, the transmitting antenna elements are relatively close to each other. In this case the delay profile is essentially the same for each transmitting element. The closed loop Tx diversity solutions developed for the FDD mode support two transmit antennas. Both open-loop and closed-loop Tx diversity solutions are specified for UTRA FDD and TDD modes.

**Open-loop Mode:** The first open-loop concepts proposed in 3G standardization were based on Code Division Transmit Diversity (Orthogonal Transmit Diversity [39]) and Time Switched Transmit Diversity [40]. Time Switched Transmit Diversity (TSTD) is applied in the WCDMA standard for certain common channels. In TSTD the transmitted signal hops across two transmit antennas, according to [23]. TSTD can be considered as a special case of the Time Division Transmit Diversity (TDTD) concept described in Chapter 3. Eventually a more efficient Space-Time Transmit Diversity (STTD) solution, based on the space–time block code developed by Alamouti [10], was adopted for Release '99 [41].

The Alamouti code used in STTD is

\[
X_{Ala}(x_1, x_2) = \begin{bmatrix}
  x_1 & -x_2^* \\
  x_2 & x_1^*
\end{bmatrix},
\]

where column 1 is transmitted from antenna 1 and column 2 from antenna 2. The symbols are QPSK modulated in Rel. 99 and Rel. 4. The transmitter structure (omitting spreading and scrambling) is shown in Figure 1.3. In the TDD mode of the WCDMA system a variant called Block STTD (B-STTD) is used. The principle of B-STTD is the same as that of STTD.
but the encoding operation is performed over symbol sequences. The motivation for B-STTD is to simplify receiver processing when applying multiuser or multichannel detection. In the TDD mode the spreading factor is reduced and therefore such advanced baseband receiver algorithms are often required.

**Closed-loop Modes:** It was noticed early in the 3G WCDMA standardization that even crude feedback signalling can be extremely useful in improving the downlink performance. The first feedback mode introduced to 3G systems was based on selective transmit diversity (STD), where only one additional feedback bit is used to select the desired transmit antenna [14, 42].

Closed-loop methods that provide beam-forming gains, as opposed to antenna selection gains, are more efficient. These gains are obtained through coherent signal combining or co-phasing in different transmit antennas, provided that the antennas share a common delay profile. In analogy with STD, co-phasing coefficients, matched to the instantaneous downlink channel, can be signalled from the UE to the BS using a fast feedback channel.

The WCDMA Release '99 and Release 4 specifications include two closed-loop transmit diversity concepts. In both closed-loop transmit diversity modes co-phasing information, contained in a fast feedback signal (of rate 1500 bps), is used to select one of 4 or 16 possible beam weights. These two modes approximate coherent transmission or channel-matched beam-forming. In both modes the terminal selects a transmission antenna route or a beam. However, they use different channel quantization and feedback signalling strategies. The related transmitter architecture is depicted in Figure 1.4.

The transmit weight is selected using the following approach. First, a given terminal obtains channel estimates for the $l$-path channels $h_1 \in \mathbb{C}^t$ and $h_2 \in \mathbb{C}^t$ for antenna 1 and antenna 2, respectively. These are estimated using antenna-specific orthogonal common channel pilot signals (CPICH). Then each terminal determines how the desired transmit weights (beam coefficients) for the dedicated channel should be modified in order to maximize the signal-to-noise ratio (or to minimize the transmit power). Thus, the problem

$$w = \text{arg max}_w (w_1 h_1 + w_2 h_2) \dagger (w_1 h_1 + w_2 h_2)$$

is solved by the UE. In an attempt to reduce feedback signalling we make the *a priori* assumption that only one complex weight is signalled.