Reliable and Spectrally Efficient Encoding and Modulation Techniques



Ferhat Yarkin

Linacre College

Department of Engineering Science, University of Oxford

Supervised by Professor Justin P. Coon

A thesis submitted to the University of Oxford for the degree of Doctor of Philosophy in Engineering Science

Trinity 2022

Abstract

Thesis Title: Reliable and Spectrally Efficient Encoding and Modulation Techniques
Name: Ferhat Yarkin
College: Linacre College
Degree: Doctor of Philosophy
Term and Year of Submission: Trinity 2022

Due to remarkable growth in mobile data traffic, intelligent devices, and connectivity, the need for spectrally/energy efficient and reliable communication systems increases day by day. Moreover, to meet the fast-growing demand for data communication and massive connectivity, next-generation networks will require to deploy more advanced modulation techniques than the existing ones. In this respect, index modulation (IM) and its combinatorial encoding structure have gained considerable popularity thanks to their attractive advantages, including low-complexity implementation, better error performance, and improved energy/spectral efficiency over conventional modulation and multiplexing schemes. Although IM is a potential candidate for the next-generation networks due to its beneficial properties, one needs to design more reliable and spectrally efficient modulation techniques to satisfy the requirements of the next-generation networks. Therefore, this thesis investigates reliable and spectrally efficient encoding and modulation techniques, proposing novel designs based on combinatorial and well-known encoding mechanisms. Our designs uncover novel applications of combinatorial tools such as set partitioning and composition in encoding. The resulting concepts are general modulation methods, which comprehend the classical and IM techniques as particular cases. They also exhibit a better error performance and higher achievable rate than the IM and classical modulation techniques, making them a better candidate for future wireless networks. Apart from the combinatorial framework, we exploit a simple maximum distance separable code to show the actual potential of the IM and related schemes as well as design a modulation concept that achieves better error and achievable rate performance than the classical and combinatorial methods while exhibiting a low-complexity structure.

Acknowledgements

First and foremost, I would like to express my sincerest thanks to my supervisor, Prof. Justin P. Coon, for allowing me to conduct research in his group, encouraging me with his insightful advice, and providing his guidance, support, and extensive knowledge. I cannot thank him enough for canalizing me toward a curiosity-driven academic mindset that helped me to carry out this research.

I would also like to thank Dr. Mihai-Alin Badiu for the help, support, and friendship that he kindly offered. Throughout the years, I truly enjoyed discussing research problems and having conversations with him. I would also like to thank Dr. Jan Vonk and Dr. Ye Liu for their valuable contributions to the discussions on my research.

I would also like to thank the Engineering and Physical Sciences Research Council (EPSRC) and Bristol Innovation & Research Laboratory of Toshiba Research Europe Ltd. for generously funding my DPhil.

I would also like to thank Prof. Dominic O'Brien and Prof. Steve Elston for being on my first and third years' assessments committee and for providing constructive feedback on my work.

My sincere thanks also go to my office and group colleagues for several enlightening conversations. It was a pleasure working with them. A special thanks to the lovely people I met at Linacre College for their friendship and for making Oxford feel like home. It was great fun having interesting discussions with them during our regular coffee, lunch, and dinner meetings.

I heartily thank my parents and brother for the innumerable things they have done for me. Though I missed them a lot throughout the years, I always felt their

unconditional love, support, encouragement, and warmth with me.

Lastly, I thank my wife, Murside, for believing in me, supporting me, and for always being there for me. You mean the world to me.

List of Publications

Journal Publication

1. **F. Yarkin** and J. P. Coon, "Modulation based on a simple MDS code: Achieving better error performance than index modulation and related schemes," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 118-131, Jan. 2022.

2. **F. Yarkin** and J. P. Coon, "Index and composition modulation," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 911-915, Mar. 2021.

3. **F. Yarkin** and J. P. Coon, "Set partition modulation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7557-7570, Nov. 2020.

4. **F. Yarkin** and J. P. Coon, "Q-ary multi-mode OFDM with index modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 7, pp. 1110-1114, Jul. 2020.

Conference Publication

1. **F. Yarkin** and J. P. Coon, "Comparing minimum codeword distances and error performance for index modulation and maximum distance separable coded modulation," *2022 EuCNC/6G Summit*, pp. 106-111, 2022.

2. **F. Yarkin** and J. P. Coon, "Simple Gray coding and LLR calculation for MDS modulation systems," *2022 IEEE WCNC*, pp. 268-273, 2022.

3. **F. Yarkin** and J. P. Coon, "Composition modulation," *2020 IEEE 20th ICCT*, pp. 131-135, 2020, **best paper award**.

4. **F. Yarkin** and J. P. Coon, "Set partition modulation," *2019 IEEE 30th PIMRC*, pp. 1-7, 2019.

Acronyms

AWGN	Additive white Gaussian noise
BER	Bit error rate
bpcu	bits per channel use
bps	bits per subcarrier
CI-IM	Coordinate interleaved index modulation
CM	Composition modulation
CSI	Channel state information
DM-IM	Dual-mode index modulation
FSPM	Full set partition modulation
GDM-IM	Generalized dual-mode index modulation
ICM	Index and composition modulation
IM	Index modulation
IQ-IM	In-phase and quadrature index modulation
LC-ML	Low-complexity maximum likelihood
LLR	Log-likelihood ratio
MDS	Maximum distance separable
MDS-APM	Maximum distance separable amplitude and phase modulation
MDS-IQM	Maximum distance separable in-phase and quadrature modulation
MDS-PSK	Maximum distance separable phase shift keying
MDS-QAM	Maximum distance separable quadrature amplitude modulation
MED	Minimum Euclidean distance
MIMO	Multiple-input multiple-output
MISO	Multiple-input single-output
ML	Maximum likelihood
MM-IQ-IM	Multi-mode in-phase and quadrature index modulation
MM-IM	Multi-mode index modulation
OFDM	Orthogonal frequency division multiplexing
OFSPM	Ordered full set partition modulation
OSPM	Ordered set partition modulation
PAM	Pulse-amplitude modulation
PAPR	Peak-to-average power ratio
PDF	Probability density function
PEP	Pairwise error probability
PM	Permutation modulation
PSK	Phase shift keying
QAM	Quadrature amplitude modulation
RV	Random variable
SE	Spectral efficiency
SISO	Single-input single-output
SM	Spatial modulation
SNR	Signal-to-noise ratio
SPM	Set partition modulation
SSK	Space shift keying
WCM	Weak composition modulation

Contents

Abstract i

Acknowledgements ii

List of Publications iv

Acronyms vi

List of Figures xi

List of Tables xiv

List of Mathematical Notations xv

1 Introduction 1
- 1.1 Index Modulation: A Promising Modulation Concept . . . 3
- 1.2 Thesis Organization and Contribution . . . 5

2 Background and Literature Review 9
- 2.1 Overview of Digital Communication . . . 9
  - 2.1.1 Basic Structure . . . 10
  - 2.1.2 Baseband Modulation and Demodulation . . . 12
  - 2.1.3 Channel Models . . . 15
- 2.2 Channel Capacity . . . 16
- 2.3 Constellation Design . . . 18
  - 2.3.1 Design Criteria . . . 19
    - 2.3.1.1 Error Performance . . . 19
    - 2.3.1.2 High Mutual Information . . . 22
    - 2.3.1.3 Low-Complexity Implementation . . . 23
    - 2.3.1.4 Peak-to-Average Power Ratio . . . 24
  - 2.3.2 One and Two Dimensional Constellations . . . 24
  - 2.3.3 Multidimensional Constellations . . . 26
- 2.4 Combinatorial Modulation Techniques . . . 27
  - 2.4.1 Permutation Modulation . . . 27
  - 2.4.2 Index Modulation . . . 29
    - 2.4.2.1 Index Modulation and Related Schemes . . . 29

2.4.2.2 Applications of Index Modulation . . . 34
2.5 Performance of Combinatorial Modulation . . . 37
2.5.1 Achievable Data Rate . . . 40
2.5.1.1 Spatial Modulation . . . 40
2.5.1.2 OFDM with Index Modulation . . . 45
2.5.2 Bit Error Rate . . . 49
2.5.3 Minimum Codeword Distances . . . 53
2.5.4 Detection Complexity . . . 56
2.6 Concluding Remarks . . . 57

3 Encoding Using Set Partitions 59
3.1 Basic Definitions . . . 61
3.2 Set Partition Modulation . . . 62
3.2.1 SPM . . . 62
3.2.2 Ordered SPM . . . 63
3.2.3 Full SPM . . . 63
3.2.4 Ordered Full SPM . . . 63
3.3 Practical Model for Parallel Channels . . . 65
3.3.1 Transmitter . . . 65
3.3.2 Receiver . . . 69
3.4 Codebook Selection . . . 72
3.4.1 Brute-force Search for k-clique Problem . . . 74
3.4.2 Vertex Exclusion for the Maximum Clique Problem . . . 76
3.4.3 Complexity Comparison . . . 78
3.5 Performance Analysis . . . 80
3.5.1 Data Rate . . . 80
3.5.1.1 SPM . . . 80
3.5.1.2 OSPM . . . 83
3.5.1.3 FSPM . . . 84
3.5.1.4 OFSPM . . . 84
3.5.2 Bit Error Rate . . . 85
3.6 Numerical Results . . . 87
3.6.1 Data Rate . . . 88
3.6.2 Bit Error Rate . . . 92
3.7 Summary . . . 97

4 Encoding Using Compositions 99
4.1 Basic Definitions . . . 101
4.2 Composition Modulation . . . 102
4.2.1 Weak Composition Modulation . . . 102
4.2.2 Composition Modulation . . . 103
4.2.3 Index and Composition Modulation . . . 103
4.3 Practical Model For Parallel Channels . . . 105
4.3.1 WCM . . . 106
4.3.1.1 Transmitter . . . 106

4.3.1.2 Receiver . . . 109
4.3.2 CM . . . 109
4.3.2.1 Transmitter . . . 109
4.3.2.2 Receiver . . . 111
4.3.3 ICM . . . 111
4.3.3.1 Transmitter . . . 112
4.3.3.2 Receiver . . . 116
4.4 Codebook Selection . . . 120
4.5 Performance Analysis . . . 122
4.5.1 Bit Error Rate . . . 122
4.5.2 Spectral Efficiency . . . 123
4.5.3 Mutual Information . . . 124
4.5.3.1 Exact Mutual Information . . . 125
4.5.3.2 A Lower-Bound . . . 127
4.5.3.3 An Upper-Bound . . . 128
4.5.3.4 Closed-Form Asymptotics . . . 130
4.6 Numerical Results . . . 131
4.6.1 Bit Error Rate . . . 132
4.6.2 Spectral Efficiency . . . 137
4.6.3 Mutual Information . . . 138
4.7 Summary . . . 140
4.A Appendices . . . 141
4.A.1 Proof of Proposition 1 . . . 141
4.A.2 Proof of Proposition 3 . . . 142
4.A.3 Proof of Proposition 5 . . . 143

5 Encoding Using an MDS Code 144
5.1 Maximum Distance Separable Coded Modulation . . . 146
5.1.1 Maximum Distance Separable Amplitude and Phase Modulation 148
5.1.2 Maximum Distance Separable In-phase and Quadrature Modulation . . . 151
5.2 Practical Model For Parallel Channels . . . 153
5.2.1 Transmitter . . . 154
5.2.1.1 MDS-PSK/QAM . . . 155
5.2.1.2 MDS-APM . . . 157
5.2.1.3 MDS-IQM . . . 160
5.2.2 Receiver . . . 163
5.2.2.1 MDS-PSK/QAM . . . 163
5.2.2.2 MDS-APM . . . 165
5.2.2.3 MDS-IQM . . . 166
5.2.2.4 Detection Complexity . . . 167
5.3 Implementation Issues . . . 170
5.3.1 Gray Coding . . . 171
5.3.2 Log-Likelihood Ratio Calculation . . . 173
5.4 Performance Analysis . . . 178

5.4.1 Spectral Efficiency . . . 178
5.4.2 Average Bit Error Rate . . . 179
5.4.3 Minimum Codeword Distances . . . 181
5.4.3.1 MDS-PSK . . . 181
5.4.3.2 MDS-QAM . . . 182
5.4.3.3 MDS-APM . . . 183
5.4.3.4 MDS-IQM . . . 185
5.5 Numerical Results . . . 186
5.5.1 Uncoded Bit Error Rate . . . 187
5.5.1.1 The Optimum ML Performance . . . 187
5.5.1.2 The Low-complexity Detector Performance . . . 190
5.5.2 Coded Bit Error Rate . . . 194
5.5.3 Achievable Rate . . . 197
5.5.4 Minimum Codeword Distance . . . 198
5.6 Summary . . . 205

6 Conclusion **207**
6.1 Thesis Summary . . . 207
6.2 Future Directions . . . 209
6.2.1 Performance Enhancement . . . 210
6.2.2 Noncoherent Transmission . . . 211
6.2.3 Adaptive Modulation and Bit Loading . . . 212
6.2.4 Variable-Length Encoding and Distribution Matching . . . 213
6.3 Conclusions . . . 214

Bibliography **216**

List of Figures

2.1 Basic structure of digital communication. 10
2.2 An example of two disjoint 8-QAM constellations. 31
2.3 An example of four disjoint PAM constellations given in [1]. 33
2.4 Transceiver structure of the SM concept. 37
2.5 The block diagram of the OFDM-IM transmitter in [2]. 38
2.6 Ergotic capacity comparison of SM and STBC for $N_t \in \{2, 4, 8\}$ and $N_r = 1$. Here, the capacity curves are reproduced from [3]. 41
2.7 Mutual information comparison of MISO-BF, TAS, and SM for $N_t \in \{2, 4\}$ and $M \in \{2, 4\}$. Here, the mutual information curves are reproduced from [4]. 44
2.8 Achievable rate comparison of OFDM-IM and conventional OFDM when $\eta = 1$ bps and $\eta = 2$ bps. Here, the rate curves are reproduced from [5]. 46
2.9 Achievable rate comparison of the OFDM-IM and conventional OFDM techniques when $\eta = 4$ bps and $\eta = 5$ bps. 47
2.10 Comparison of the average mutual information of OFDM-IM and the OFDM capacity when $K = 1, \ldots, 8$. Here, the case with $K = 8$ corresponds to the OFDM capacity. 49
2.11 BER comparison of the SM and SSK techniques with the SIMO and V-BLAST techniques for $N_r = 4$ and $\eta \in \{3, 4, 6\}$. Here, "V-BLAST (N_t, M)" stands for the V-BLAST system that employs M-PSK modulation over N_t transmit antennas, whereas "SIMO (M-PSK/QAM)" is the SIMO concept transmitting M-PSK or M-QAM signals over single transmit antenna. "SM (N_t, M-PSK/QAM)" and SSK (N_t) are the SM concept employing M-PSK/QAM over N_t transmit antennas and the SSK scheme having N_t transmit antennas, respectively. 50
2.12 BER comparison of the OFDM-IM techniques with the conventional OFDM technique. 51
2.13 BER comparison of the MM-OFDM-IM and OFDM-IM techniques with the conventional OFDM technique. 52
2.14 Comparison of the minimum codeword distances for MM-OFDM-IQ-IM, MM-OFDM-IM, OFDM-IM, and OFDM. 56

3.1 An example of disjoint PSK and QAM constellations. 66
3.2 Transmitter structure of the OFDM-SPM scheme. 68
3.3 Vertex exclusion example for SPM when $N = 4$ and $K = 3$. 77

3.4 Data rate comparison of the SPM variants with DM-IM, GDM-IM, and MM-IM. 89
3.5 Achievable rate comparison of SPM $(4, 2, 2)$ and OSPM $(4, 2, 2)$ with MM-IM $(2, 2)$, DM-IM $(4, 2)$, IM $(4, 2, 4)$, and BSPK. 91
3.6 Achievable rate comparison of OFSPM $(4, 4)$ and FSPM $(4, 4)$ with MM-IM $(4, 4)$, IM $(4, 2, 8)$, and 8-PSK. 91
3.7 BER comparison of SPM $(4, 2, 2)$, OSPM $(4, 2, 2)$, and OSPM $(4, 2, 2)$, Alg. 3 with MM-IM $(2, 2)$, DM-IM $(4, 2)$, IM $(4, 2, 4)$, and BPSK. . . 93
3.8 BER comparison of FSPM $(4, 2)$, OFSPM $(4, 2)$, and OFSPM $(4, 2)$, Alg. 3 with MM-IM $(4, 2)$, IM $(4, 3, 4)$, and QPSK. 94
3.9 BER comparison of FSPM $(4, 4)$, OFSPM $(4, 4)$, and OFSPM $(4, 4)$, Alg. 3 with MM-IM $(4, 4)$, IM $(4, 3, 8)$, and 8-PSK. 95
3.10 BER comparison of the proposed LC-ML and the optimum ML detectors for the SPM techniques. 97

4.1 The constellation points that are used to form the WCM, CM, and ICM symbols. 105
4.2 Constellation diagram and associated codewords of WCM when $N = I = 3$ and $\lambda = 1$. 108
4.3 The constellation points that are used to construct the CM codewords when $N = 3, I = 5$, and $M = 4$. 111
4.4 The constellation points and codewords that form the ICM codebook when $N = I = 4$, $K = 3$, and $M = 4$. 114
4.5 Transmitter structure of the OFDM-ICM scheme. 115
4.6 BER comparison of WCM $(4, 4, 1)$, Alg. 4 and CM $(4, 7, 2)$ with IM $(4, 3, 4)$ and QPSK. 133
4.7 BER comparison of WCM $(4, 6, 1)$, Alg. 4, CM $(4, 6, 4)$, and CM $(4, 12, 2)$ with IM $(4, 3, 8)$. 134
4.8 BER comparison of ICM with WCM, CM, IM, and M-PSK when $N = 4$, $K \in \{2, 3\}$, $I \in \{3, 4, 5, 8\}$, $M \in \{2, 4\}$, and $\lambda = 1$. 134
4.9 BER comparison of ICM with IM, WCM, CM, QAM, and APSK when $N = 4$, $K = 3$, $I \in \{4, 7, 8\}$, $M \in \{4, 8\}$, and $\lambda = 2$. 135
4.10 BER comparison of the proposed LC-ML detector with the optimum ML detector for ICM. 136
4.11 Comparison of the theoretical, numerical, and asymptotic SE results for ICM. Here, the curves with extensions "Theory" in the legend are obtained by substituting (4.15) into (4.6), whereas those with extensions "Numerical" in the legend are obtained by substituting all possible values of K into (4.6) and choosing the maximum SE. On the other hand, the curves labeled "Asymptotic" in the legend are obtained by substituting (4.14) into (4.18). 137
4.12 Mutual information vs. SNR for $N = 4$, $I = 9$, $\delta = 0.2$, and $\delta = 0.7$ where $|h(n)|^2 = \delta^n$ for $n = 1, 2, \ldots, N$. 138
4.13 Comparison of mutual information for CM, parallel Gaussian channels, and IM when $N = 4$. 139

5.1 Constellation diagram of MDS-APM when $Z = J = 2$ and $N = 3$. . . . 149
5.2 Constellation points that are used to generate the MDS-IQM codewords when $R = T = 2$ and $N = 3$. 153
5.3 An example of disjoint PSK and QAM constellations. 157
5.4 Transmitter structure of the OFDM-MDS-PSK/QAM scheme. 158
5.5 Constellation diagram of MDS-APM when $Z = 4$, $J = 2$, and $M = 2$. . 160
5.6 Four disjoint PAM constellations produced from the 8-PAM constellation [1]. 161
5.7 Complexity comparison of MDS-APM with MM-IM, IM, and convetional modulation in terms of decoding complexity per bit. 171
5.8 Constellation diagram of MDS-APM when $Z = 4$, $J = 2$, and $M = 2$. . 183
5.9 BER comparison of MDS-APM (2, 2, 2, 1) and MDS-IQM (2, 2, 2, 1) with ICM (4, 2, 5), IM (2, 1, BPSK), and BPSK. 188
5.10 BER comparison of MDS-PSK $(4, 4, 2)$ and MDS-PSK $(4, 8, 1)$ with OFSPM $(4, \text{BSPK})$, MM-IM $(4, \text{BSPK})$, IM $(4, 3, \text{QPSK})$, and QPSK. 189
5.11 BER comparison of MDS-IQM (4, 4, 4, 1), MDS-APM (4, 2, 8, 1), and MDS-APM (4, 2, 4, 2) with OFSPM (4, 4-QAM), MM-IM (4, 4-PSK), IM (4, 3, 8-PSK), and 8-QAM. 189
5.12 BER comparison of MDS-QAM (N, Q, M) and MDS-PSK (N, Q, M) with OFSPM $(4, 4\text{-QAM})$, OFSPM $(4, \text{QPSK})$, MM-IM $(4, \text{QPSK})$, IM $(4, 3, 8\text{-PSK})$ and 8-PSK for $N = 4, Q \in \{8, 16\}, M \in \{1, 2\}$. 191
5.13 BER comparison of MDS-APM (4, 2, 8, 2) and MDS-IQM (4, 8, 6, 1) with MM-IQ-IM (4, 2), MM-IM (4, 8-QAM), ICM (4, 3, 10, 8), IM (16, 15, 16-QAM), and 16-QAM. 192
5.14 BER comparison of MDS-APM (4, 4, 4, 4) and MDS-IQM (4, 4, 4, 2) with MDS-QAM (8, 4, 8), MDS-PSK (4, 16, 4), MM-IM (4, 16-QAM), IM (4, 3, 64-QAM), and 32-QAM. 193
5.15 BER comparison of the LC-ML detector with the optimum ML detector for the MDS-PSK, MDS-IQM, and MDS-APM systems. 194
5.16 BER comparison of the MDS-IQM schemes employing the mapping in Section 5.1.1 and the Gray-like mapping. 195
5.17 BER comparison of the uncoded and coded MDS-IQM schemes. . . . 195
5.18 Coded BER comparison of the MDS-IQM and combinatorial techniques. 197
5.19 Achievable rate comparison of the proposed methods with the combinatorial and conventional concepts. 198
5.20 BER comparison of the MDS-IQM and MDS-QAM systems with the QAM, IM, and IQ-IM techniques when $N = 2$ and $\eta \in \{2, 4\}$. 200
5.21 BER comparison of the MDS modulation techniques with the QAM, IM, IQ-IM, ICM, and CI-IM techniques. 203
5.22 Minimum codeword distance vs. η for the MDS modulation, QAM, and IM techniques. 204

List of Tables

2.1 A look-up table implementation example for $N = 4$ and $K = 2$. 39
2.2 Comparison of detection complexities for OFDM-IM, MM-OFDM-IM, and OFDM. 57

3.1 An example of codeword generation for the SPM variants when $N = 3$ and $K = 2$. 64
3.2 The run-time complexities (in seconds) of the vertex selection algorithms along with the numbers of achieved vertices (in brackets) for the OSPM concept when $N \in \{4, 6, 8\}$ and $K = 2$. 79
3.3 The run-time complexities (in seconds) of the vertex selection algorithms along with the numbers of achieved vertices (in brackets) for the OFSPM concept when $N \in \{3, 4, 5, 6\}$. 80
3.4 Descriptions of the modulation concepts considered in this section. . . 88

4.1 Codebook generation example for ICM when $N = I = 3$ and $K = 2$. . . 105
4.2 Codeword generation and look-up table implementation example for WCM when $N = I = 3$ and $\lambda = 1$. 107
4.3 Look-up table implementation for ICM when $N = I = 4$ and $K = 3$. . 113

5.1 Generating amplitude and phase vectors for MDS-APM when $Z = J = 2$ and $N = 3$. 149
5.2 Bits-to-index symbols mapping for MDS-APM when $Z = 3$ and $N = 3$. 151
5.3 Generating in-phase and quadrature components for MDS-IQM when $R = T = 2$ and $N = 3$. 152
5.4 Detection complexity comparisons for MDS-IQM, IM, MM-IM, and conventional modulation. 167
5.5 An example of Gray mapping for the MDS code when $N = 3$, $Q = 4$, and $l = 2$. 172
5.6 Complexity order comparison of LLR calculations per bit for MDS-IQM, IM, and MM-IM. 177
5.7 Minimum codeword distance comparison of the MDS modulation techniques with the QAM, IM, and IQ-IM techniques when $N = 2$. 199
5.8 Minimum codeword distance comparison of the MDS techniques with the QAM, IM, IQ-IM, ICM, and CI-IM techniques when $N = 4$. . . . 201
5.9 Minimum codeword distance comparison of the MDS techniques with the QAM, IM, IQ-IM, ICM, and CI-IM techniques when $N = 8$. . . . 202

List of Mathematical Notations

$\mathbf{h}$	Channel vector
$\mathbf{w}$	Noise vector
L	Modulation diversity
d_1	Minimum L-product distance
d_2	Minimum Euclidean distance
P_b	Average bit error rate
b	Number of input bits
G	Undirected graph
P	Number of communication blocks
f	Number of bits in a block
N	Codeword length
N_0	Noise power
η	Spectral efficiency
$\mathbb{D}(\cdot\|\|\cdot)$	Relative entropy operator
$\mathbb{C}$	Set of complex numbers
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_{\geq 0}$	Set of non-negative real numbers
$\mathbb{N}$	Set of natural numbers
$\mathbb{Z}^+$	Set of positive integers
$\mathbf{I}$	Identity matrix
$\mathrm{diag}(\cdot)$	Diagonal matrix operator
$\mathrm{E}[\cdot]$	Expectation operator
$[\cdot]^T$	Transpose operator
$Q(\cdot)$	Gaussian Q-function
$\Pr(\cdot)$	Probability operator
$\mathbb{H}(\cdot)$	Entropy operator
$\mathbb{h}(\cdot)$	Differential entropy operator
$\mathrm{I}(\cdot;\cdot)$	Mutual information operator
$\|\| \cdot \|\|$	Norm operator
$\ln(\cdot)$	Natural logarithm
$\log_2(\cdot)$	Logarithm function with base two
$\lfloor \cdot \rfloor$	Floor operator
$\lceil \cdot \rceil$	Ceiling operator
$\det(\cdot)$	Determinant operator

Chapter 1

Introduction

The mathematical and practical establishment of today's communication networks dates back to the time when Claude Shannon defined the fundamental problem of communication theory as "that of reproducing at one point either exactly or approximately a message selected at another point". In his seminal paper [6], Shannon invented *information theory*, thus he showed the way of measuring information in communication networks and how the information could be reliably transmitted over imperfect channels through *digital communications*.

The introduction of digital communications paved the way for modern communication networks and enabled developments in today's wireless communications. However, until the 1980s, we could not conduct wireless communications by using our mobile phones or personal computers. With the introduction of first-generation (1G) mobile networks in 1979, long-distance wireless voice transfer became available. The era of digital communications started in the 1990s with the launch of second-generation (2G) mobile networks. 2G networks enabled its users to send/receive digital voice and data, thus improving coverage and reliability as well as decreasing latency and cost compared to the 1G networks. In 2001, third-generation (3G) mobile networks started bringing new applications such as web browsing, video downloading, send-

ing/receiving e-mails, and multimedia messages. It has also begun to provide much faster data rates than 2G. In 2009, fourth-generation (4G) mobile networks reached even faster data rates that enabled video conferencing, high-definition (HD) video transfer, and online gaming available for mobile users. The fifth-generation (5G) networks, on the other hand, will mainly support three services, including enhanced mobile broadband, ultra-reliable low-latency communication, and massive machine-type communication, by using different technologies such as massive multiple-input multiple-output (MIMO), millimeter-wave (mmWave) communications, ultra-dense networks, and so on. 5G's key enabling technologies are expected to satisfy the user demands for a considerably high data rate, ultra-reliable transmission, and very low latency.

The next-generation networks will require an even more data rate, higher error performance, and lower latency than the 5G networks. That is evident from the proliferation of a variety of new applications, including extended-reality services, telemedicine, haptics, flying vehicles, brain-computer interfaces, and connected autonomous systems [7]. Moreover, mobile data traffic, data traffic per subscriber, and the number of connected devices are expected to increase dramatically by 2030 [8]. Due to the accumulation of new applications and remarkable growth in mobile data services, intelligent devices as well as connectivity, the need for spectrally/energy efficient and reliable communication systems is growing fast. In this context, to satisfy these needs, the next-generation networks will require to deploy advanced modulation techniques.

1.1 Index Modulation: A Promising Modulation Concept

Modulation is one of the main stages of digital communications that play a vital role in the realization of transmission standards and the mobile networks mentioned above. Modulation embarks the task of encoding the source message into analog waveforms for transmission through the channel. As the transmission medium has disruptive effects on the transmitted signals, the choice of waveforms or related constellation points is a non-trivial problem that plays a critical role in the system performance. To overcome the disruptive effects of wireless channels, most networks exploit well-known one or two-dimensional constellations such as pulse amplitude modulation (PAM), phase shift keying (PSK), and quadrature amplitude modulation (QAM). Although these concepts exhibit a low implementation complexity, their performance is limited. One can achieve better performance by exploiting higher dimensions through vector modulation concepts since these concepts encode the source message into a broader space where the distances among constellation points could be much higher. Moreover, vector modulation concepts offer higher flexibility in system design since one can adjust the points in each dimension and cull out the ones with undesirable properties to attain good performance. However, these advantages come at the cost of increased transceiver complexity since the search space of the optimum detector becomes considerably expanded, and there may not exist straightforward ways of implementing bits-to-symbols/symbols-to-bits mappings when employing these concepts.

Index modulation (IM), which is a subclass of permutation modulation (PM) in [9], emerged as a promising vector modulation method thanks to its combinatorial structure. Such a structure enables IM to achieve a better error performance and higher achievable rate than the classical techniques while exhibiting a low implementation complexity [2, 10, 11]. In [11], a space domain IM concept called spatial modula-

tion has been proposed and shown to perform better than conventional modulation concepts such as QAM and PSK in terms of error performance. The authors of [2] have introduced a frequency domain application of IM called orthogonal frequency division multiplexing with IM (OFDM-IM) and show that OFDM-IM exhibits a low implementation complexity as well as a high SNR efficiency when compared to conventional OFDM. The potential of IM has also been documented for other domains such as time and code domains [12].

Despite the advantages of the IM techniques over conventional modulation techniques, it is not apparent if the combinatorial structure can achieve more and attain better performance. The actual potential of IM is also not evident from the literature since most of the IM works focus on showing the effectiveness of the IM techniques against the conventional modulation techniques such as QAM and PSK or the OFDM concepts employing these techniques. In this context, one can utilize different tools of enumerative combinatorics and well-known encoding methods to reveal if there exist more reliable and spectrally efficient modulation techniques.

On the other hand, it is not always easy to solve practical problems that come with IM. For instance, the studies in the literature show that Gray mapping between index symbols and information bits is not straightforward. Although one can find low-complexity Gray coding algorithms for IM in the literature [13], the IM systems using these algorithms still need to employ the optimum detection at the receiver that exhibits excessive complexity when the codebook size is large. Besides, to calculate log-likelihood ratio (LLR) values related to information bits of the IM techniques, we need to consider all IM codewords due to the dependence between the codeword elements. The complexity of the LLR calculations will still be high when the codebook size is large since there is no direct way to calculate element-wise low-complexity LLR values for the IM systems. Such issues may limit the applications of the IM schemes for the 5G and beyond 5G systems.

Against this background, the primary goals of the thesis are

- To design reliable and spectrally efficient encoding and modulation techniques that have the potential for application in the next-generation networks.
- To exploit the beneficial properties of the combinatorial structure and discover novel applications of enumerative combinatorics in encoding.
- To achieve better performance than the IM and classical concepts while preserving a similar or lower complexity structure.
- To show the true potential of the existing mechanisms and solve their practical problems through our novel modulation designs.

1.2 Thesis Organization and Contribution

In this thesis, we propose and analyze encoding and modulation techniques based on enumerative combinatorics and well-known encoding mechanisms. The thesis is comprised of three research chapters, which are devoted to presenting our novel encoding and modulation concepts. The research chapters are preceded by a background and literature review chapter on digital communication and recent modulation techniques. Moreover, a chapter that provides our conclusions and ideas on potential future works follows the research chapters. To better represent the content of the thesis and our technical contributions, we give detailed outlines of each chapter in the following paragraphs. Note that we define and use most mathematical notations independently in each chapter for clarity. Also, we use MATHEMATICA and MATLAB to verify theoretical expressions and conduct Monte Carlo simulations.

In **Chapter 2**, we present the necessary background on digital communications, channel models, and constellation design to be able to draw the complete picture of the existing and proposed encoding techniques. We then provide a comprehensive

literature review on combinatorial modulation concepts, highlighting the advantages and disadvantages of employing these concepts.

In **Chapter 3**, we propose a novel modulation scheme called set partition modulation (SPM). In this scheme, set partitioning and ordered subsets in the set partitions are used to form codewords. We define different SPM variants and depict a practical model for using SPM with parallel channels. For this model, we specify transmitter and receiver operations and use different constellations to distinguish between different subsets in a set partition. We also propose a sub-optimal detection algorithm to overcome the excessive complexity of the optimum detection at the receiver. The proposed algorithm is shown to reduce the receiver complexity remarkably and attain near-optimum performance. To achieve good distance properties and better error performance for the SPM codewords, we define a codebook selection problem and formulate such a problem as a clique problem in graph theory. In this regard, we propose two codebook selection algorithms based on brute-force search and vertex exclusion. The complexity of these algorithms is investigated, and the latter is shown to be a fast and efficient solution for the codebook selection problem. We analyze the error and achievable rate performance of the proposed schemes and provide asymptotic results for the performance. It is shown that the proposed SPM variants are general methods, which encompass other techniques as special cases. We also show that the SPM methods can exhibit a better error performance and higher achievable rate than the IM and classical modulation techniques. This research has been published in *IEEE Transactions on Wireless Communications* [14] and *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)* [15].

In **Chapter 4**, we uncover a different tool of enumerative combinatorics in the modulation design and propose novel encoding techniques based on integer partitioning or composition. We provide basic definitions related to composition along with the

details of the proposed concepts that we call *weak composition modulation* (WCM), *composition modulation* (CM), and *index and composition modulation* (ICM). We use weak and strict compositions of an integer to form codewords of WCM and CM, respectively. Besides, in ICM, we utilize indices of active/deactive codeword elements and compositions of an integer to encode information. We depict a practical model employing the proposed methods in parallel channels. In this model, we perform the composition encoding over the energies of parallel channels that carry M-PSK symbols. We design an efficient low-complexity detector for the proposed techniques. Moreover, we analyze bit error rate (BER), spectral efficiency (SE), and mutual information for the proposed techniques and derive respective closed-form expressions. Our numerical and analytical results show that the proposed concepts can outperform the IM and classical modulation techniques. This research has been published in *IEEE Communications Letters* [16] and *2020 IEEE 20th International Conference on Communication Technology (ICCT)* [17].

Chapter 5 proposes two novel modulation concepts based on a simple maximum distance separable (MDS) code and shows that these concepts can achieve better error performance than IM and related schemes. In the first concept, we use amplitude and phase levels to form a simple MDS code, whereas, in the second one, in-phase and quadrature components of codeword elements are used to construct the MDS code. We depict practical schemes for using the proposed concepts with parallel channels. We exploit disjoint constellations in these concepts and introduce different variants of MDS-coded modulation. The chapter presents that the proposed techniques exhibit a structure enabling efficient low-complexity detection and simple bits-to-symbols/symbols-to-bits mappings. They have simple solutions for practical problems such as Gray coding between information bits and constellation points; and low-complexity LLR calculations for channel-encoded information bits. We analyze the performance in terms of the minimum codeword distance and BER and prove

through extensive numerical results that the proposed techniques outperform IM and related concepts. More importantly, contrary to the vast majority of IM studies that focus on showing the superiority of the IM techniques against conventional modulation techniques, the chapter shows that modulation concepts based on a well-known MDS code can achieve better error performance and higher codeword distances than the IM and related schemes while exhibiting a structure as simple as these schemes. This research has been published in *IEEE Transactions on Communication* [18], *IEEE Wireless Communications Letters* [19], *2022 IEEE Wireless Communications and Networking Conference (WCNC)* [20], and *2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)* [21].

In **Chapter 6**, we summarize the motivations and novel contributions of the thesis. Moreover, we point out the potential future work for the proposed methods and emphasize the challenging problems related to performance enhancement, noncoherent and adaptive transmission, and variable-length encoding. We also provide our perspectives to address these problems. Finally, the chapter concludes the thesis.

Chapter 2

Background and Literature Review

2.1 Overview of Digital Communication

The notion of communication theory is based on conveying information-bearing signals (messages) from a source (transmitter) to one or more destinations (receivers) as reliably as possible. In principle, the source is distant, and its message which can take several forms such as audio, video, image, text, and so on is unknown to the destination. Thus, the main problem of communication is to reproduce an exact or approximate replica of this message at the destination [6]. Since the transmitter and receiver are apart from each other, the transmission is conducted through a physical medium called a *channel*. Due to the noise, interference, and other dispersive effects in the transmission medium, the amount of information that can be transmitted reliably is limited. Therefore, capturing the source's message accurately at the destination is a non-trivial problem that may require sophisticated algorithms and equipment.

Digital communication performs the message transmission in *digital form*. Here, *digital form* implies nothing but surjective mappings[1] between the analog waveforms and binary digits, i.e., bits. The main benefits of these mappings come from the ability to manipulate the message waveforms as well as corresponding bit sequences,

[1]We use the terms 'coding', 'encoding', and 'mapping' interchangeably throughout this thesis.

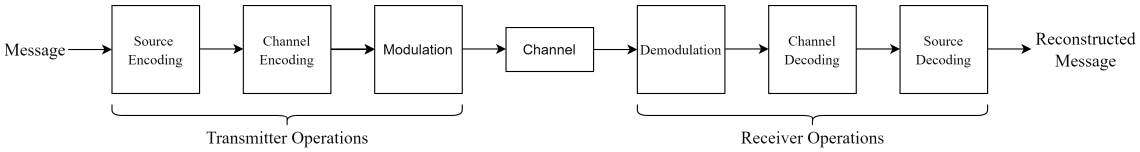


Figure 2.1. Basic structure of digital communication.

and transform them into forms that are more suitable for transmission through a noisy channel. Thus, the resulting networks are far more resilient to noisy channel conditions than analog communication networks where the original message waveforms are directly carried on the amplitude or phase of a carrier signal during the transmission. In the following subsection, we will touch on the main components of a digital communication system and elaborate more on how we perform the encoding at the transmitter and decoding at the receiver.

2.1.1 Basic Structure

Fig. 2.1 illustrates the basic structure of digital communication. As seen from the figure, a message signal enters the transmitter.[2] Transmitter performs some operations on the message signal for efficient transmission and intends to send it to the receiver. Moreover, receiver operations are based on reversing these operations and recovering the message signal.

The basic transmitter operations can be listed as the source encoding, channel encoding, and modulation. Source encoding is the process in which a message signal is encoded into bit strings. A source encoder takes discrete samples of continuous-time message signals, i.e., *sampling*. In practice, these samples can take an infinite number of values. To be able to map them to finite-length bit strings, we need to quantise them to predetermined values. Once we have the quantised version of the message signal, we can easily map the related quantised values to the bit strings.

[2]Although the message signal is generally a continuous-time signal, it can also be a discrete-time signal such as sensor data measured at specific time instants.

This is where *data compression* plays an important role to increase the efficacy of the transmission.

Unlike the source encoder, the channel encoder introduces systematic redundancy among the bit strings, or in other words, maps them to longer bit strings. This process aims to increase the distance between bit strings and detect/correct errors at the receiver. Although a digital communication system does not necessarily employ channel encoding, it is of paramount importance to achieve data rates close to the highest data rates that a system can transmit information reliably in a noisy channel [6].

After channel encoding, we need to encode the resulting bit strings into analog waveforms for transmission through the channel. *Modulation* undertakes this encoding operation. As the channel has disruptive effects on the transmitted signals, the choice of the waveforms is non-trivial, and it plays a significant role in the system performance. Designing time-domain waveforms is especially hard since it requires solving complex integrals and performing time-domain analysis. Instead, one can use a signal space representation and solve the equivalent problem in vector space. We will focus on the signal space representation in the following subsection. For now, let us assume we have the signal waveforms for transmission. These waveforms are baseband signals since their frequency spectra have components at near-zero frequencies. Therefore, the encoding procedure defined here is called baseband modulation. Ideally, a modulator lifts the baseband signals to much higher frequencies. This process is called passband modulation or up-conversion. Since the resulting passband signals can be down-converted to the baseband frequencies at the receiver, it is sufficient to focus on baseband signals when designing modulation techniques. *Therefore, this thesis demonstrates reliable and spectrally efficient modulation techniques at the baseband level.*

The transmitter sends the up-converted signal through the channel. In the mean-

time, the signal may be subject to disruptive effects depending on the nature of the transmission medium. Additionally, it is always subject to noise at the receiving end due to the unavertable existence of thermal noise. As a result, the received signal at the receiver will be a corrupted version of the transmitted signal. At this point, the receiver down-converts the received signal and compares it with available waveforms,[3] then decides the transmitted signal based on pairwise distances among the received and available waveforms. Here, it must suffice to say that the receiver employs reverse operations to recover the source message. Therefore, the receiver operations are demodulation, channel decoding, and source decoding.

2.1.2 Baseband Modulation and Demodulation

The process of baseband modulation is based on mapping b binary information bits to one of M waveforms, $x_m(t)$, where $m = 1, \ldots, M$ and $M = 2^b$. The Gram-Schmidt orthogonalization procedure ensures that any function $x_m(t)$ can be represented as a linear combination of orthonormal basis functions, thus we can write

$$x_m(t) = \sum_{i=1}^{N} x_{m,i}\phi_i(t) \tag{2.1}$$

for some function $x_m(t)$ and some set of orthonormal basis functions $\{\phi_1(t), \ldots, \phi_N(t)\}$ [22]. Here, $x_{m,1}, \ldots, x_{m,N}$ are the points in the N-dimensional complex signal space. There is a bijective mapping between the waveforms, $x_m(t)$, $m = 1, \ldots, M$, and their vector representations, $\mathbf{x}_m = (x_{m,1}, x_{m,2}, \ldots, x_{m,N})^T \in \mathbb{C}^{N\times 1}$ where $\mathbf{x}_m \in \mathcal{X}$, $\mathcal{X}$ is a vector space spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_M$, and $\mathbb{C}^{N\times 1}$ is the set of complex numbers of the size $N \times 1$. Thus, once we know $\mathbf{x}_m$ and the set of orthonormal basis functions, we can construct the waveform $x_m(t)$ by using (2.1). Such a representation of a function in (2.1) is called a signal space representation. Therefore, if $\phi_1(t), \ldots, \phi_N(t)$ are given,

[3] Prior to the transmission, the receiver has knowledge of available waveforms at the transmitter.

one can neglect the design of the waveforms $x_1(t), \ldots, x_M(t)$ and focus on designing a signal constellation that is comprised of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_M$ as constellation points in an N-dimensional complex space. In this way, the resulting constellation will be suitable for any orthogonal transmission scheme in any domain. After this point, we assume that the modulated symbols are up-converted before transmission at the transmitter, and the received symbols are down-converted before detection at the receiver.

Based on the signal space representation in (2.1) and assuming $\mathbf{x}_m$ is selected for transmission, the received signal vector can be written as

$$\mathbf{y} = \mathbf{x}_m + \mathbf{w} \tag{2.2}$$

where $\mathbf{w}$ is an $N \times 1$ complex noise vector. Observe from (2.2) that the transmitted signal is subject to noise at the receiver. We refer this specific model as an *additive white Gaussian noise (AWGN)* channel since the elements of $\mathbf{w}$ are Gaussian distributed, i.e., $w(n) \sim \mathcal{CN}(0, N_0)$ where $\mathbf{w} = [w(1), \ldots, w(N)]^T$, $n = 1, \ldots, N$, and N_0 is the noise power. In general, we use the AWGN channel to model the wired, satellite, and deep-space communication links among others. Here, the probability density function (PDF) of the noise vector is given by

$$f(\mathbf{w}) = \left(\frac{1}{\sqrt{\pi N_0}}\right)^N e^{-\frac{||\mathbf{w}||^2}{N_0}}. \tag{2.3}$$

We need another model when the transmission medium causes multipath propagation, in other words, the transmitted signal reaches the receiver by one or more paths since the objects in the medium cause reflection, refraction, and scattering of the signals. Therefore, the received signal will be an attenuated and phase-shifted version of the transmitted signal. In a narrowband system, the signal duration is much higher than that of the delay spread caused by multi-path transmission, i.e., the bandwidth

of the transmitted signal is less than that of the channel coherent bandwidth. Here, the bandwidth of the transmitted signal is the frequency range that the transmitted signal occupies, whereas the channel coherent bandwidth is the frequency range that the channel's frequency response remains almost constant. Thus, the narrowband model assuming constant attenuation and phase shift over a signal duration can be given as

$$\mathbf{y} = \mathbf{X}_m \mathbf{h} + \mathbf{w} \tag{2.4}$$

where $\mathbf{X}_m = \text{diag}(\mathbf{x}_m)$ and $\mathbf{h} = [h(1), \ldots, h(N)]^T$ is the $N \times 1$ complex channel coefficient vector that is used to model the attenuation and phase shift caused by the channel. Here, $\text{diag}(\mathbf{x}_m)$ gives a diagonal matrix with the elements of vector $\mathbf{x}_m$ on the main diagonal, and $h(n)$ is the complex channel coefficient related to the nth codeword element.

To determine the transmitted signal, the receiver can employ the optimum decision rule called the *maximum a posteriori probability (MAP)* rule as follows [22]

$$\hat{\mathbf{x}}_m = \max_{\mathbf{x}_m \in \mathcal{X}} \Pr(\mathbf{x}_m|\mathbf{y}) = \max_{\mathbf{x}_m \in \mathcal{X}} \frac{\Pr(\mathbf{x}_m) f(\mathbf{y}|\mathbf{x}_m)}{f(\mathbf{y})} \tag{2.5}$$

where $\Pr(\mathbf{x}_m|\mathbf{y})$ is the probability that $\mathbf{x}_m$ is transmitted given $\mathbf{y}$, and $\Pr(\mathbf{x}_m)$ is the probability that $\mathbf{x}_m$ is selected for transmission. Here, the second equation follows from the Bayes' theorem. Moreover, $f(\mathbf{y})$ and $f(\mathbf{y}|\mathbf{x}_m)$ are the PDFs of $\mathbf{y}$ and $\mathbf{y}$ given $\mathbf{x}_m$, respectively. Since $f(\mathbf{y})$ is not dependent on $\mathbf{x}_m$, we can drop it in the decision rule. Assuming equally likely signals, i.e., $\Pr(\mathbf{x}_m) = \frac{1}{M}, \forall m$, and substituting $f(\mathbf{y}|\mathbf{x}_m) = \left(\frac{1}{\sqrt{\pi N_0}}\right)^N e^{-\frac{||\mathbf{y}-\mathbf{x}_m||^2}{N_0}}$ into Eq. (2.5), one can obtain the following decision rule [22]

$$\hat{\mathbf{x}}_m = \max_{\mathbf{x}_m \in \mathcal{X}} f(\mathbf{y}|\mathbf{x}_m) = \min_{\mathbf{x}_m \in \mathcal{X}} ||\mathbf{y} - \mathbf{x}_m||^2. \tag{2.6}$$

Notice that this is based on deciding the transmitted signal that is the nearest to the received signal. This detection method is called the *maximum likelihood (ML) detector*, and it can be given for a wireless channel by

$$\hat{\mathbf{x}}_m = \max_{\mathbf{x}_m \in \mathcal{X}} f(\mathbf{y}|\mathbf{x}_m, \mathbf{h}) = \min_{\mathbf{x}_m \in \mathcal{X}} ||\mathbf{y} - \mathbf{X}_m \mathbf{h}||^2. \tag{2.7}$$

Although the ML detection is optimum when the transmitted signals are equally likely, it can cause high detection complexity when the size of the constellation is large. Therefore, the existence of low-complexity detection algorithms is crucial to making a constellation practical. In this regard, this thesis work focuses on designing practical constellations in which low-complexity detection can be achieved at the cost of slight performance degradation.

2.1.3 Channel Models

We need statistical distributions to model the randomness in the channels. We already mentioned that the noise at the receiver is a random process, and the channel is called the AWGN channel when the elements of the noise vector are Gaussian distributed.

On the other hand, in a wireless channel, the transmitted waveform is subject to path loss depending on the signal wavelength and distance between the transmitter and receiver. It is easy to calculate the path loss since it is a deterministic function of the wavelength and distance. However, small-scale effects due to the multi-path propagation result in random fluctuations in the transmitted waveform. This phenomenon is called *channel fading*, and one can simply depict its effect on the transmitted signal by a multiplicative channel coefficient vector $\mathbf{h}$ as shown in (2.4). Since such a phenomenon results in random fluctuations in the signal, the elements of the channel coefficient vector are random variables (RVs). When there are a large number of reflected waves arriving at the receiver, the central limit theorem holds,

and real and imaginary parts related to elements of the channel vector become uncorrelated Gaussian random variables. The channel is said to be a *Rayleigh fading channel*, when the real and imaginary parts of the vector elements, $h(n)$, are independent and identically distributed (i.i.d.) Gaussian RVs, i.e., $h^I(n), h^Q(n) \sim \mathcal{N}(0, \sigma^2)$ where $h^I(n) = \mathfrak{Re}\{h(n)\}$ and $h^Q(n) = \mathfrak{Im}\{h(n)\}$. In this case, the magnitude of the elements, $|h(n)|$, will be Rayleigh distributed [22], i.e.,

$$f_{|h(n)|}(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, \qquad r \geq 0. \tag{2.8}$$

It can be shown that the phase of the channel coefficient follows a uniform distribution. The Rayleigh model well suits urban environments where there is no direct line of sight between the transmitter and receiver, and the transmitted signal arrives at the receiver through many different paths. Note that, in the simulations of this thesis, we assume all schemes operate over a Rayleigh fading channel, whose elements are i.i.d. RVs. The use of the Rayleigh model in the simulations is realistic when considering environments with a large number of small reflectors [23].

Alternatively, the channel is said to be a *Rician fading channel* when there is a line-of-sight path between the transmitter and receiver, and the real and imaginary components of the channel coefficients are not zero-mean, i.e., $h^I(n) \sim \mathcal{N}(m_1, \sigma^2)$, $h^Q(n) \sim \mathcal{N}(m_2, \sigma^2)$, and $m_1, m_2 \neq 0$. In this case, $|h(n)|$ follows Rician distribution.

2.2 Channel Capacity

There is a limit on the rate of information that can be reliably transmitted through a noisy channel. This limit is called the channel capacity. The channel capacity is of paramount importance due to Shannon's noisy channel coding theorem. The theorem states that error-free transmission is possible if the rate of the transmission η is less than the channel capacity, C, i.e., $\eta < C$. On the contrary, it is impossible to

ensure error-free transmission when $\eta > C$. Therefore, we will focus on deriving the channel capacity of different techniques to see their actual potential and limits. In this regard, this short section is dedicated to providing a formal definition of the channel capacity. Here, our definition is based on the assumption that the transmitted and received symbols are RVs. We denote these RVs by X and Y, respectively.

Definition 1. Channel Capacity: For a system whose input-output relation is defined by RVs X and Y, one can define the channel capacity under a power constraint on the input distribution, i.e., $\mathrm{E}[X^2] \leq \mathcal{P}$, as the maximum mutual information with respect to the input distribution, $f(x)$, i.e.,

$$C = \max_{f(x), \mathrm{E}[X^2] \leq \mathcal{P}} I(X;Y). \tag{2.9}$$

Here, $I(X;Y)$ is the mutual information, and it is the measure related to the amount of information that an RV, X, has about another RV, Y. The constraint on the power of the input signal comes from the fact that power is an important resource that is limited in practical communication systems.

When X is a continuous RV, the input distribution that yields the channel capacity for the AWGN and fading channels is Gaussian distribution, i.e., $X \sim \mathcal{CN}(0, \sigma_x^2)$ where σ_x^2 is the power of input X. In this case, the AWGN channel capacity can be written as

$$C = \log_2 \left(1 + \frac{\sigma_x^2}{N_0}\right), \tag{2.10}$$

whereas the (ergodic) capacity of the fading channels is given by

$$C = \mathrm{E}_h \left[\log_2 \left(1 + |h|^2 \frac{\sigma_x^2}{N_0}\right)\right] \tag{2.11}$$

where we assume that h is a realization of an ergodic random process and only the

receiver has the knowledge of the fading coefficients. These assumptions hold in the rest of this thesis unless stated otherwise.

2.3 Constellation Design

Since the transmission of Gaussian signals is not possible in practice, constellation design plays an important role in efficient information transmission. Although the channel capacity is achieved in the AWGN and fading channels when the input symbols are Gaussian distributed, these channels have different characteristics when it comes to constellation design. Unlike the AWGN channels, the fading channels have a randomly changing signal-to-noise ratio (SNR) and they may have a time-varying model depending on the propagation environment. Moreover, shadowing, Doppler shifts, interference, and nonlinear characteristics make the fading channels considerably different from the AWGN channel. Therefore, by defining the set of all possible codewords as a *code*, a "good" code in the AWGN channel may not be good enough in a fading channel. This is mostly the case when we have multidimensional constellations. That is why here we limit our discussion on constellation design based on a fading channel affected by flat, slow fading, and AWGN. In this regard, we assume an interleaver of sufficient length that permutes the order of the transmitted codeword elements is used in this model so that each element experiences independent fading. Thus, the model output related to the nth codeword element, when $\mathbf{x}_m$ is sent, is

$$y(n) = h(n)x_{m,n} + w(n), \qquad 1 \leq n \leq N. \tag{2.12}$$

Following the basic structure of a digital communication system given in Section 2.1.1, the constellation design problem can be regarded as a problem of finding the "most suitable" points in an $2N$-dimensional space to map the bit strings.[4]

[4]We call the set, $\mathcal{X}$, containing the constellation points a *constellation* or, alternatively, a *code-*

2.3.1 Design Criteria

Finding the most suitable constellation points is not an easy problem as several desirable characteristics make a constellation appropriate for transmission through a fading channel. Some of these characteristics can be listed as good error performance, high mutual information, low-complexity implementation, compatibility with the existing coding techniques, and low peak-to-average power ratio (PAPR) [24]. We now focus on these characteristics and touch upon practical guidelines on the codebook design.

2.3.1.1 Error Performance

Although maximizing the *minimum Euclidean distance (MED)* among codewords and minimizing the number of codeword pairs having the MED are the main criteria to achieve good asymptotic error performance in an AWGN channel, meeting these objectives does not necessarily provide good error performance for fading channels. To see the main factors affecting the performance in fading channels, we need to specify the effects of erroneous detection of $\mathbf{X}_m$ as $\mathbf{X}_{m'}$ where $\mathbf{X}_m = \text{diag}(\mathbf{x}_m)$, $\mathbf{X}_{m'} = \text{diag}(\mathbf{x}_{m'})$, $m \neq m'$, and $m, m' \in \{1, \ldots, M\}$ [25]. In this regard, let $P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'})$ denote pairwise error probability (PEP) associated with the erroneous detection of $\mathbf{X}_m$ as $\mathbf{X}_{m'}$.[5] When the codebook consists of only two codewords $\mathbf{X}_m$ and $\mathbf{X}_{m'}$, $P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'})$ is enough to quantify the error probability of the codebook. That is not the case when there are $M > 2$ codewords in the codebook. However, we can still use the PEP expression to upper-bound the error probability of a signaling concept given $\mathbf{x}_m$ is sent, $P(e|\mathbf{X}_m)$, as follows [25]

$$P(e|\mathbf{X}_m) \leq P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'}). \tag{2.13}$$

To obtain an expression for $P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'})$, let us assume that channel state

book.

[5]Here, we assume unit energy per codeword element for ease of representation and simplicity.

information (CSI) is available at the receiver and focus on the conditional PEP given the CSI, i.e., $P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'}|\mathbf{h})$. Based on the ML rule in (2.7), one can write the PEP conditioned on $\mathbf{h}$ as

$$P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'}|\mathbf{h}) = \Pr(||\mathbf{y} - \mathbf{X}_m\mathbf{h}||^2 > ||\mathbf{y} - \mathbf{X}_{m'}\mathbf{h}||^2|\mathbf{h}) = \Pr(X > 0|\mathbf{h}) \tag{2.14}$$

where $X = -||(\mathbf{X}_m - \mathbf{X}_{m'})\mathbf{h}||^2 - 2\Re\mathfrak{e}\{\mathbf{w}^* \cdot ((\mathbf{X}_m - \mathbf{X}_{m'})\mathbf{h})\}$ is a Gaussian random variable whose mean and variance are given by $m_X = -||(\mathbf{X}_m - \mathbf{X}_{m'})\mathbf{h}||^2$, and $\sigma_X^2 = 2N_0||(\mathbf{X}_m - \mathbf{X}_{m'})\mathbf{h}||^2$, respectively [26]. Therefore, the conditional PEP expression in (2.14) can be calculated by

$$P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'}|\mathbf{h}) = Q\left(\sqrt{\frac{||(\mathbf{X}_m - \mathbf{X}_{m'})\mathbf{h}||^2}{2N_0}}\right) \tag{2.15}$$

where $Q(.)$ is the Gaussian Q-function, i.e., $Q(r) = \frac{1}{\sqrt{2\pi}}\int_r^\infty e^{-x^2/2}dx$. It is not easy to interpret the performance by looking at (2.15) as we have the Gaussian Q-function which does not yield a closed-form expression. Therefore, by using the well-known Chernoff bound, i.e., $Q(r) \leq \frac{1}{2}\exp(-r^2/2)$, the conditional PEP expression can be upper-bounded by

$$P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'}|\mathbf{h}) \leq \frac{1}{2}\exp\left(-\frac{||(\mathbf{X}_m - \mathbf{X}_{m'})\mathbf{h}||^2}{4N_0}\right) = \frac{1}{2}\prod_{n=1}^{N}\exp\left(-\frac{\delta_n|h(n)|^2}{4N_0}\right) \tag{2.16}$$

where $\delta_n = |x_{m,n} - x_{m',n}|^2$. Assuming $h(n) \sim \mathcal{CN}(0,1)$ is an i.i.d. RV for each n and averaging (2.16) over $|h(n)|$, one can obtain the following upper-bound on the PEP associated with the erroneous detection of $\mathbf{X}_m$ as $\mathbf{X}_{m'}$

$$P(\mathbf{X}_m \rightarrow \mathbf{X}_{m'}) \leq \frac{1}{\prod_{n=1}^{N}(1 + \delta_n/4N_0)}. \tag{2.17}$$

Let us define SNR per codeword element as $\gamma = 1/N_0$. When the SNR is high,

one can ignore the one in the denominator of (2.17) and rewrite the upper-bound as

$$P(\mathbf{X}_m \to \mathbf{X}_{m'}) \leq \frac{4^L}{(\prod_{n\in\Gamma}\delta_n)\gamma^L} = \left(\frac{4}{(\prod_{n\in\Gamma}\delta_n)^{\frac{1}{L}}\gamma}\right)^L \tag{2.18}$$

where Γ is the set of indices related to nonzero values of δ_n, and $L = |\Gamma|$ is the minimum Hamming distance between the codewords of a codebook, i.e., *diversity order*. Here, the type of diversity is *modulation diversity* since we deal with the minimum number of the distinct points between any two codewords [26]. Note also that one can use $(G_c\gamma)^{-G_d}$ to represent the right-hand side of (2.18) [25]. Here, G_c signifies the *coding gain*, whereas the exponent of the SNR term, G_d, is called the *diversity gain*. For the right-hand side of (2.18), we have $G_c = \frac{(\prod_{n\in\Gamma}\delta_n)^{\frac{1}{L}}}{4}$ and $G_d = L$.

Assuming $P(\mathbf{X}_m \to \mathbf{X}_{m'})$ is the worst-case PEP for the codebook, the code performance will be limited by $P(\mathbf{X}_m \to \mathbf{X}_{m'})$ [25]. By looking at (2.18), the diversity order is L, whereas the coding gain relates to the product distance, i.e., $\prod_{n\in\Gamma}\delta_n$. Therefore, the primary criterion is to maximize the minimum Hamming distance, L, to achieve good error performance at high SNR. Then, we need to maximize the product distance to improve the performance further. These criteria are equivalent to the well-known rank-determinant criteria of [27].

When SNR is low, i.e., $\gamma << 1$, one can rewrite the upper-bound in (2.17) by neglecting the contribution of higher order terms of γ as follows [28]

$$P(\mathbf{X}_m \to \mathbf{X}_{m'}) \leq \frac{1}{1+\gamma\sum_{n\in\Gamma}\delta_n}. \tag{2.19}$$

It is clear from this upper-bound that the MED between the codeword pairs should be maximized to achieve a good error performance at low SNR.

At medium SNR, say $\gamma \approx 1$, the upper-bound in (2.17) becomes

$$P(\mathbf{X}_m \to \mathbf{X}_{m'}) \leq \frac{1}{\prod_{n=1}^{N}(1+\delta_n)}. \tag{2.20}$$

In this case, one should maximize the product-distance to achieve good error performance. However, this criterion may not provide optimum performance for the whole medium SNR range since it is based on a certain SNR value. Thus, it is less practical than the criteria at high and low SNR [28].

Moreover, an upper-bound on the average bit error probability of the codebook is given by the well-known union bound [22]

$$P_b \leq \frac{1}{M \log_2 M} \sum_{m=1}^{M} \sum_{m'=1}^{M} P(\mathbf{X}_m \to \mathbf{X}_{m'}) D(\mathbf{X}_m \to \mathbf{X}_{m'}) \tag{2.21}$$

where $D(\mathbf{X}_m \to \mathbf{X}_{m'})$ is the number of bit errors for the corresponding pairwise error event. As seen from the upper-bound, minimizing the PEP expressions alone is not enough to attain the optimum performance. One should also consider the problem of mapping constellation symbols to bit sequences so that the resulting bit error probability is minimized. The optimum solution for this problem is to map the nearest constellation points to the bit strings in a way that the resulting bit strings differ in only one bit, i.e., *Gray code*. Although the implementation of Gray code is usually straightforward for a regular one and two-dimensional constellations, it may require excessive storage or processing power for higher dimensions as it gets harder to assess the distance between the constellation points.

2.3.1.2 High Mutual Information

Another design criterion is related to the potential of a codebook in maximizing mutual information between input and output. Thus, a codebook that can maximize mutual information is preferable. Here, the aim is to maximize the throughput [25].

Note that criteria related to good error performance and high mutual information are different, and the designs related to these criteria may result in substantially different codebooks. However, it is usually desirable for a codebook to achieve good error performance and high mutual information. This is why BER and achievable data rate are the key performance indicators for the modulation concepts in this thesis.

2.3.1.3 Low-Complexity Implementation

An efficient codebook is the one that allows low-complexity implementation at the transmitter and receiver. In this context, the mappings between the bit strings and constellation symbols should be simple enough to implement. Although the use of look-up tables may be feasible when the codebook size is not large, it becomes impractical when we increase the number of constellation points beyond some number [24]. Thus, a desirable feature of a codebook is to have a structure enabling simple mappings between the bit strings and constellation points that prevent the use of look-up tables. It may not be clear how a codebook can have such a structure; however, we believe it will become apparent in the upcoming sections when we mention several modulation concepts with this structure.

On the other hand, it is critical to perform uncomplicated detection at the receiver. Although the ML detection is optimum in error performance, its complexity becomes prohibitive when the size of the constellation is large. Therefore, a constellation should have low-complexity detectors that provide a good trade-off between error performance and low-complexity detection [29].

Also, as discussed in Section 2.1, channel encoding is crucial for a digital communication system to conduct reliable transmission. However, we need to calculate LLR values for each bit of the used modulation techniques to obtain markedly improved error performance in the presence of channel encoding. That is because soft-decision

decoding based on the LLR values has a clear BER advantage against hard-decision decoding for the coded techniques. Therefore, an efficient modulation concept should be compatible with the existing channel encoding techniques in the sense that reliable and low-complexity LLR calculations are possible [30].

2.3.1.4 Peak-to-Average Power Ratio

PAPR is the ratio of the peak power related to the constellation points to the average power of the constellation, and it quantifies the sensitivity of a codebook to nonlinearities as well as other signal-dependent perturbations [24]. PAPR of a multi-dimensional constellation can occasionally be high when each codeword element has high energy. High PAPR can cause out-of-band radiation unless the transmit power amplifier operates in its linear region, which degrades the efficiency of the power amplifier since the conversion from direct current to radio frequency (RF) power is inefficient in this region [31]. Thus, a constellation should have a low PAPR.

2.3.2 One and Two Dimensional Constellations

In one dimension, one can achieve the optimum error performance under an average power constraint when the constellation points are located equidistantly over a dimension or an axis. The modulation concept associated with this strategy is called pulse amplitude modulation (PAM), and the PAM constellation is given by

$$\mathcal{X}_{PAM} = \Big\{ -\frac{d}{2}(M-1), \ldots, -\frac{d}{2}, \frac{d}{2}, \ldots, \frac{d}{2}(M-1) \Big\} \tag{2.22}$$

where d is the distance between the neighboring constellation points.

In two dimensions, one can embed information into both the amplitude and phase of the signal. A well-known modulation concept based on transmitting amplitude and phase information is called quadrature amplitude modulation (QAM). The hexagonal

QAM constellations are optimum in the AWGN channel since the densest lattice in two dimensions is a hexagonal lattice [32]. However, regular QAM constellations are widely preferred against these constellations as they have less complex ML detection where their performance is only marginally worse than that of their hexagonal counterparts. In a regular QAM constellation, in-phase and quadrature components are drawn from the same set of constellation points as in (2.22).

Although the QAM constellations work very well in AWGN channels, they are not optimum in fading channels since they do not fully take advantage of two dimensions. To utilize both dimensions and obtain modulation diversity in a Rayleigh fading channel, one can rotate a regular QAM constellation by an angle so that the in-phase and quadrature components of the constellation points are not equal. With the help of interleaving, such parts encounter independent fading [26]. As a result, the constellation achieves a diversity order of two, and the probability that a constellation point is in a deep fade decreases. In addition to the diversity gain, it is important to maximize the minimum product-distance between the constellation points and choose the rotation angle accordingly. In this way, one can considerably improve the error performance in a fading channel. Here, the performance improvement is attained in an uncoded manner since the constellation rotation does not introduce any redundancy to the bit strings. Moreover, the bandwidth and power of the transmission remain unchanged compared to the unrotated signal set.

Another well-known two-dimensional modulation concept is phase-shift keying (PSK). In PSK, constellation points are located equidistantly on the unit circle, and the information is embedded into the phase of the signal points. In this regard, one can denote the mth point in an M-PSK constellation as $e^{j2\pi(m-1)/M}$ for $m = 1, 2, \ldots, M$. Therefore, the PSK constellation is given by

$$\mathcal{X}_{PSK} = \left\{1, e^{j2\pi/M}, \ldots, e^{j2\pi(M-1)/M}\right\}. \tag{2.23}$$

Although the PSK concept is widely used for wireless local area networks (LAN) and Bluetooth communication due to its desirable properties such as having the same energy across the codewords and low PAPR, the error performance of the QAM constellation is superior to that of the PSK concept. Such a superiority becomes more evident when the constellation size is large since the Euclidean distance between the PSK symbols decreases rapidly as we increase the number of the constellation points.

2.3.3 Multidimensional Constellations

Assuming independent fading affects each element of a multidimensional constellation through component interleaving, one can overcome the effect of Rayleigh fading by achieving high modulation diversity. Moreover, the performance becomes close to that in the AWGN channel when we keep increasing the modulation diversity [26]. As mentioned earlier, this type of diversity is attractive because it improves the error performance without requiring additional power, bandwidth, and the use of channel encoding. However, such improvement comes at the cost of increased demodulation complexity.

There are two well-known methods to construct multidimensional constellations with high modulation diversity. One is based on designing high diversity lattice constellations by using complex algebraic number fields [33]. The other is based on rotating multidimensional constellations that draw their elements from well-known constellations such as QAM and PSK [26]. The latter has some practical advantages since it offers good error performance in the AWGN channel and enables simple Gray mapping between the bit strings and codewords.

2.4 Combinatorial Modulation Techniques

We now turn our attention to a different family of multidimensional constellations, which exploits combinatorial tools such as permutation and combination. Although the related modulation concepts cannot attain high diversity order, they offer attractive trade-offs among performance, spectral efficiency (SE), and complexity.[6] Also, these concepts received considerable attention recently in academia due to their satisfactory error performance, low-complexity decoding, simple bits-to-symbols and symbols-to-bits mappings, and compatibility with existing coding and modulation techniques. In this section, we describe these modulation concepts and focus on showing their potential.

2.4.1 Permutation Modulation

The idea of embedding information in the element permutations of a codeword was first proposed by Slepian in [9]. Slepian's idea, which he called permutation modulation (PM), was hinged upon constructing a codebook by permuting elements of a codeword. PM has some desirable properties such as requiring the same energy across the codewords, low-complexity detection, and inherent modulation diversity of two in the Rayleigh channel [34]. It is also a general modulation concept that encompasses pulse code modulation, pulse position modulation, and orthogonal codes.

In [9], Slepian proposed two different types of PM. In the first type, called Variant I, each codeword is an N-tuple of real numbers, and the elements of an initial codeword are permuted to obtain the other codewords. Let the initial codeword be

$$\mathbf{x}_1 = (\overbrace{\mu_1, \ldots, \mu_1}^{m_1}, \overbrace{\mu_2, \ldots, \mu_2}^{m_2}, \ldots, \overbrace{\mu_K, \ldots, \mu_K}^{m_K}) \tag{2.24}$$

[6] Here, the SE or the bandwidth efficiency refers to the rate of transmission over a given bandwidth, whereas the complexity is related to the transceiver complexity. Since we assume the same bandwidth for all schemes, it suffices to describe the SE over the information rate per codeword element during a transmission interval, i.e., η.

where

$$\mu_1 < \mu_2 < \ldots < \mu_K, \tag{2.25}$$

$\mu_k \in \mathbb{R}$, m_k is the multiplicity of the kth element, μ_k, in the codeword, and $k = 1, \ldots, K$. Based on this form, we have $N = \sum_{k=1}^{K} m_k$. PM obtains the remaining codewords of its codebook by permuting the elements of $\mathbf{x}_1$. Therefore, the codebook size is

$$M_I = \frac{N!}{m_1! m_2! \ldots m_K!}. \tag{2.26}$$

For the second type of PM, called Variant II, the codeword elements are nonnegative real numbers, i.e., $\mu_k \in \mathbb{R}_{\geq 0}$, and it is assumed that

$$0 \leq \mu_1 < \mu_2 < \ldots < \mu_K. \tag{2.27}$$

The initial codeword of this type takes the same form as that of the Variant I code. However, in addition to the element permutations, changing the sign of the codeword elements is used to generate different codewords. Therefore, the size of the Variant II codebook is given by

$$M_{II} = \frac{N! 2^{\lambda}}{m_1! m_2! \ldots m_K!} \tag{2.28}$$

where $\lambda = N - m_1$, if $\mu_1 = 0$ and $\lambda = N$, otherwise.

The combinatorial structure, in other words, the use of permutations, enables PM to have beneficial properties. Although PM is designed primarily for AWGN channels, it achieves a diversity gain of two in Rayleigh fading channels as the minimum Hamming distance between any two codewords is two. This can be observed from

(2.24) as any permutations of $\mathbf{x}_1$ will differ in at least two elements. For example, let us consider the initial codeword $\mathbf{x}_1 = (\mu_1, \mu_2, \mu_3)$ for the PM Variant-I codebook. The codebook consists of the codewords $\mathbf{x}_1 = (\mu_1, \mu_2, \mu_3)$, $\mathbf{x}_2 = (\mu_1, \mu_3, \mu_2)$, $\mathbf{x}_3 = (\mu_2, \mu_1, \mu_3)$, $\mathbf{x}_4 = (\mu_2, \mu_3, \mu_1)$, $\mathbf{x}_5 = (\mu_3, \mu_1, \mu_2)$, and $\mathbf{x}_6 = (\mu_3, \mu_2, \mu_1)$. Notice that the minimum Hamming distance between the codewords is two. Moreover, although PM is a vector modulation technique, element-wise low-complexity detection mechanisms exist in the literature. Also, PM can support simple bits-to-symbols and symbols-to-bits mappings without requiring look-up tables, which bring exhaustive complexity when the codebook size is large.

2.4.2 Index Modulation

2.4.2.1 Index Modulation and Related Schemes

Index modulation (IM) techniques, which can be considered as a subclass of PM, have attracted remarkable interest due to their capabilities for achieving better error performance and improved energy/spectral efficiency compared to conventional systems [35]. In IM, K out of N elements of constellation points are nonzero elements, and the combinations related to the locations of these elements are used to form the N-dimensional constellation points [2]. Since each combination corresponds to a different constellation point, this structure produces a total of $\binom{N}{K}$ constellation points.[7] Moreover, unlike the PM concept in [9], IM produces different constellation points by choosing each nonzero element from an M-ary two-dimensional constellation diagram such as the M-PSK and M-QAM constellations. Therefore, an IM codeword related to the initial combination and the ξth modulation symbol is given by

$$\mathbf{x}_{1,\xi} = (\overbrace{\nu_1, \ldots, \nu_K}^{K}, \overbrace{0, \ldots, 0}^{N-K}) \tag{2.29}$$

[7]By using a simple fixed-length bit mapping scheme, the IM mapper can only map $\lfloor \log_2 \binom{N}{K} \rfloor$ information bits to these combinations. Here, $\lfloor . \rfloor$ is the floor operation.

where $v_k \in \mathcal{M}$, $k = 1, \ldots, K$, and $\mathcal{M}$ is an M-ary constellation, i.e., $|\mathcal{M}| = M$. Since we activate K elements out of N and each activated element can take M different values, we have $\xi \in \left\{1, 2, \ldots, M^K\right\}$. One can form all other codewords of IM through the element combinations of $\mathbf{x}_{1,\xi}$ and the change of M-ary symbols. Assuming M is an integer power of two, the SE of IM is given by

$$\eta = \frac{\lfloor \log_2 \binom{N}{K} \rfloor + K \log_2 M}{N}. \tag{2.30}$$

On the other hand, IM can be employed separately over the in-phase and quadrature parts of the codeword elements. This special version of IM is called in-phase and quadrature index modulation (IQ-IM) [36, 37]. In this technique, K nonzero in-phase and quadrature components of the codeword elements are located independently to form the combinations related to the constellation points. Thus, a total of $2\lfloor \log_2 \binom{N}{K} \rfloor$ bits are mapped to the corresponding constellation points. Moreover, each nonzero component is chosen from an M-PAM constellation. Thus, the SE of IQ-IM is

$$\eta = \frac{2\lfloor \log_2 \binom{N}{K} \rfloor + 2K \log_2 M}{N}. \tag{2.31}$$

Similar to IM, an initial IQ-IM codeword is given by

$$\mathbf{x}_{1,\xi} = \mathbf{x}^I_{1,\xi_1} + j\mathbf{x}^Q_{1,\xi_2} \tag{2.32}$$

where $\mathbf{x}^I_{1,\xi_1} = (\overbrace{v^I_1, \ldots, v^I_K}^{K}, \overbrace{0, \ldots, 0}^{N-K})$, $\mathbf{x}^Q_{1,\xi_2} = (\overbrace{v^Q_1, \ldots, v^Q_K}^{K}, \overbrace{0, \ldots, 0}^{N-K})$. Here, v^I_k and v^Q_k are points in an M-PAM constellation, $\xi_1, \xi_2 \in \left\{1, 2, \ldots, M^K\right\}$, and $\xi \in \left\{1, 2, \ldots, M^{2K}\right\}$.

One can also construct the constellation points in a way that no element is zero [1, 38]. In this case, one can use different constellations to form the combinations or permutations related to the constellation points. Here, the different constellations are

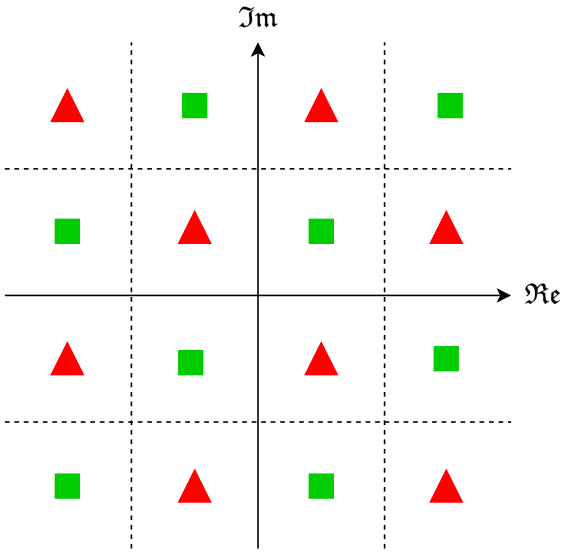


Figure 2.2. An example of two disjoint 8-QAM constellations.

disjoint, or in other words, they have no constellation points in common. Dual-mode IM (DM-IM) employs these constellations over the elements of a constellation point, and it exploits the combinations of elements chosen from these constellations to form different constellation points [38]. Assuming K and $N-K$ out of N elements are chosen from the first and second disjoint constellations, respectively, an example of DM-IM codeword and the SE of the DM-IM concept are given by

$$\mathbf{x}_{1,\xi} = (\overbrace{\nu_{1,1}, \ldots, \nu_{K,1}}^{K}, \overbrace{\nu_{1,2}, \ldots, \nu_{N-K,2}}^{N-K}), \tag{2.33}$$

$$\eta = \frac{\lfloor \log_2 \binom{N}{K} \rfloor + N \log_2 M}{N} \tag{2.34}$$

where $\nu_{k,1} \in \mathcal{M}_1$, $\nu_{k,2} \in \mathcal{M}_2$, $\mathcal{M}_1$ and $\mathcal{M}_2$ are disjoint M-ary constellations.[8] Since we activate all the elements of DM-IM codeword, we have $\xi \in \{1, 2, \ldots, M^N\}$.

To show the points that form the DM-IM constellation, we give an example of two disjoint 8-QAM constellations in Fig. 2.2. Here, the constellation points with the same color and marker are related to the same disjoint constellation. One can observe

[8]For two disjoint constellations $\mathcal{M}_1$ and $\mathcal{M}_2$, $\mathcal{M}_1 \cap \mathcal{M}_2 = \emptyset$.

from Fig. 2.2 that we obtain the two 8-QAM constellations by set partitioning the classical 16-QAM constellation.

Instead of having fixed K and $N-K$, one can generalize the DM-IM concept by assigning different number of codeword elements to the first and second disjoint constellations. The resulting concept is called generalized DM-IM (GDM-IM) [39]. In GDM-IM, the set, $\mathcal{K}$, containing possible numbers of codeword elements having one of a number of distinguishable constellations in each block is given by $\mathcal{K} = \{k_1, k_2, \ldots, k_T\}$ where T is the size of the set and $k_t, k_{t'} \in \{1, 2, \ldots, N\}$ for $k_t \neq k_{t'}$, $t \neq t'$, and $1 \leq t, t' \leq T$. Thus, the SE of GDM-IM is given by

$$\eta = \frac{\lfloor \log_2 \left(\sum_{t=1}^{T} \binom{N}{k_t} \right) \rfloor + N \log_2 M}{N}. \tag{2.35}$$

Moreover, another IM concept called multi-mode IM (MM-IM) chooses each element of a constellation point from a disjoint constellation to form the permutations [1]. Thus, we need N disjoint constellations to form the constellation points in MM-IM. By permuting the elements selected from N disjoint constellations, one can obtain $N!$ different constellation points. Hence, assuming each disjoint constellation is an M-ary constellation, the SE of MM-IM is given by

$$\eta = \frac{\lfloor \log_2 N! \rfloor + N \log_2 M}{N}. \tag{2.36}$$

Moreover, an MM-IM codeword is given by

$$\mathbf{x}_{1,\xi} = (\nu_{1,1}, \nu_{2,2}, \ldots, \nu_{N,N}) \tag{2.37}$$

where $\nu_{n,n} \in \mathcal{M}_n$, $\mathcal{M}_n$ is the nth disjoint constellation, and $|\mathcal{M}_n| = M$. Similar to DM-IM, we have $\xi \in \{1, 2, \ldots, M^N\}$. All other codewords of MM-IM can be obtained by permuting locations of the codeword elements and choosing such elements from

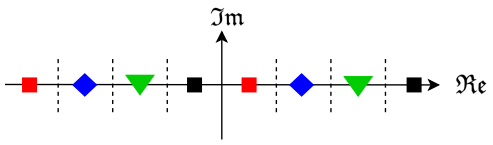


Figure 2.3. An example of four disjoint PAM constellations given in [1].

the related disjoint constellations.

In multi-mode IQ-IM (MM-IQ-IM), on the other hand, in-phase and quadrature parts of the elements are drawn from N disjoint M-PAM constellations. By permuting these parts independently, one can obtain $2N!$ different constellation points for MM-IQ-IM. Thus, the SE of MM-IQ-IM is

$$\eta = \frac{2\lfloor \log_2 N! \rfloor + 2N \log_2 M}{N}. \tag{2.38}$$

To show the points that form the MM-IQ-IM constellation, we give an example of $N = 4$ disjoint PAM constellations in Fig. 2.3 [1]. Here, each disjoint constellation has two points, and the constellation points with the same color and marker belong the same disjoint constellation. As seen from the figure, we obtain the disjoint PAM constellations by set partitioning the classical 8-PAM constellation.

BER curves of the aforementioned IM techniques cannot have a diversity order higher than one due to the use of conventional modulation techniques on top of the combinatorial encoding. However, one can achieve a diversity order for the BER curves of the IM and MM-IM techniques by employing coordinate interleaving. In coordinate interleaved IM (CI-IM) and MM-IM (CI-MM-IM), the points in the M-QAM/PSK constellations are rotated by an angle in a way that the imaginary parts of the points will not be equal. Then, the imaginary parts are interchanged between the $(2k-1)$th and $2k$th activated elements for CI-IM and between the $(2l-1)$th and $2l$th elements for CI-MM-IM where $k = 1, 2, \ldots, K$, $K = 2I_1$, $l = 1, 2, \ldots, N$, $N = 2I_2$, I_1, and I_2 are positive integers. In this way, the real and imaginary parts of the

complex data symbols are transmitted over different active codeword elements. It is important to note that the minimum Hamming distance between the codewords of CI-IM and CI-MM-IM is equal to two.

2.4.2.2 Applications of Index Modulation

It is possible to perform the IM approach through the indices of active/inactive sources. For example, combinations of the (in)active transmit antenna indices form the IM codewords in multi-antenna communication techniques called *space shift keying* (SSK) [40] and *spatial modulation* (SM) [11]. SSK exploits only antenna indices to convey information, whereas SM uses both the active antenna indices as well as conventional M-ary constellations and achieves a higher SE than SSK and better error performance than the conventional modulation techniques [11, 41–43]. Moreover, *quadrature SM* in [44] extends the SM structure through the in-phase and quadrature dimensions. A more general technique called *space-time shift keying* extends SM to include the time dimension and embeds information into the combinations of dispersion matrices [45, 46]. The authors of [47] and [48] consider antenna polarization as another domain of implementation for IM. The networks in these studies carry additional information by selecting polarization states among the available polarization states of transmit antennas based on the IM mechanism. Another modulation concept called *media-based modulation* encodes information bits into the indices of channel realizations formed by the ON-OFF status of RF mirrors at the transmit antennas [49, 50]. This is only possible because the RF mirrors contain PIN diodes and turning these mirrors ON and OFF alters the far-field radiation pattern of a reconfigurable antenna [51]. On the other hand, the authors of [52, 53] proposed the *code IM* technique based on direct sequence-spread spectrum modulation in which the indices of spreading codes are selected according to the IM method. In contrast, the activation patterns of *subcarriers* are used to construct IM codewords in a multicarrier

communication technique called orthogonal frequency-division multiplexing with IM (OFDM-IM) [2, 54, 55]. The multidimensional IM concepts of [56] and [57] convey extra information through the active indices in more than one dimension. In [56], the IM concept is implemented through the indices of antennas, time slots, and RF mirrors, whereas [57] exploits the indices of transmit antennas and subcarriers. Since only a fraction of communication sources are activated in IM, these concepts benefit from the sparse structure of IM and make use of compressed sensing. On the other hand, one can find many applications of IM in vehicular communications [58, 59], molecular communications [60–62], optical communications [63–65], physical layer security [66, 67], and relaying networks [68–70]. Other PM/IM-based applications exist with manifestations in space, time, frequency, polarization, code, and channel domains (see, e.g., [10, 12, 71, 72]).

The application of IM to parallel channels in the frequency domain through the well-known OFDM structure brings various advantages [2, 5, 36, 54, 55, 73–81]. In particular, the authors of [2, 54, 55] show that OFDM-IM is capable of achieving substantially better error performance than conventional OFDM. To further improve the error performance, transmit diversity and trellis-coded modulation were applied to OFDM-IM in [73] and [74], respectively. These studies show how one can achieve a diversity gain for the OFDM-IM scheme by sacrificing the data rate. In [75], OFDM with CI-IM (OFDM-CI-IM) was proposed, and it was shown that the OFDM-CI-IM scheme can achieve an additional diversity gain without sacrificing the data rate. In [36], two different generalized OFDM-IM schemes were proposed to improve the SE and error performance. Moreover, it was revealed by [5,76] that OFDM-IM is capable of achieving a better achievable rate than conventional OFDM for small modulation orders. Useful guidelines to design spectrally efficient OFDM-IM schemes are reported in [77–79]. In [80], a binary tree encoding method was applied to OFDM-IM to cover all of the subcarrier activation patterns and, therefore, improve the data rate beyond

recent benchmarks. For the same purpose, a discrete cosine transform-based solution for OFDM-IM was implemented in [81]; substantial data rate improvements compared to OFDM-IM and conventional OFDM were shown possible.

To encode information in the combinations of the subcarriers, a certain number of subcarriers are nulled in OFDM-IM. Although carrying information on the combinations of active/inactive subcarriers results in better error performance and improved spectral/energy efficiency for small modulation orders, it becomes difficult to achieve SEs comparable to conventional OFDM for high modulation orders. To overcome this problem, the idea of employing disjoint constellations on different subcarriers rather than nulling them has been considered [1, 12, 38, 39, 82–86]. In [38], an OFDM scheme called *dual-mode OFDM with IM* (DM-OFDM-IM) was documented; this method uses two disjoint constellations rather than active/inactive subcarriers to encode information. In [82], a dual-mode index modulation aided OFDM scheme that employs two PSK constellations with different power levels was proposed. The authors of [39] further considered altering the number of subcarriers modulated by the same constellation and provided a more general IM scheme called *Generalized DM-OFDM-IM* (GDM-OFDM-IM). In [83], the authors proposed two different precoding-aided OFDM-IM schemes that use disjoint constellations. The first scheme of [83] can be considered as a generalization of DM-OFDM-IM since such a scheme partitions the subcarriers into groups and uses the same amount of constellations as the number of groups to modulate the subcarriers. In [84], in addition to two disjoint constellations, some subcarriers are allowed to remain unused, which yields a third mode that, as it turns out, enhances the BER performance relative to DM-OFDM-IM and OFDM-IM. In [1], Wen et al. proposed a *multi-mode OFDM-IM* (MM-OFDM-IM) scheme, which uses disjoint constellations on each subcarrier of OFDM blocks to increase the SE as well as to improve BER performance. In [85] and [12], the MM-OFDM-IM scheme is generalized. The authors of [86] extended the MM-OFDM-IM scheme to space and

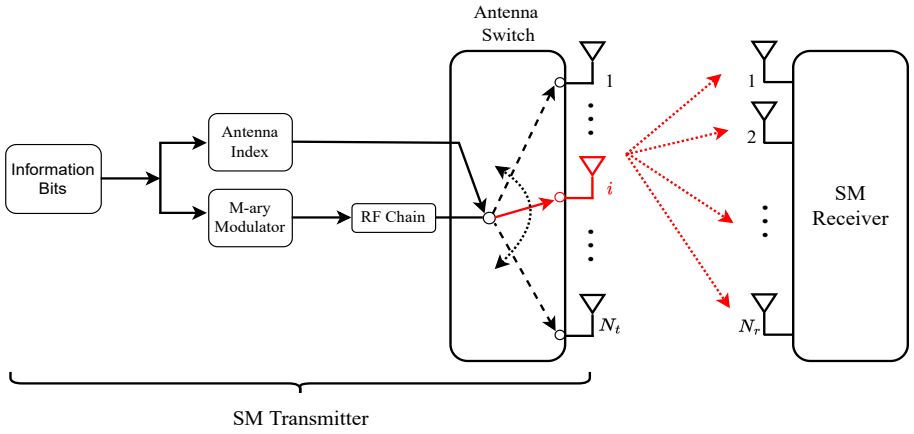


Figure 2.4. Transceiver structure of the SM concept.

time domains by using Latin matrices.

2.5 Performance of Combinatorial Modulation

In this section, we investigate the performance of the combinatorial modulation techniques. For that purpose, we focus on the two most studied applications of IM, namely SM and OFDM-IM, thus we briefly reveal the signal models for the SM and OFDM-IM techniques in this section.

We illustrate the transceiver structure of the SM concept in Fig. 2.4. Here, the SM transmitter maps the information bits to transmit antenna indices and M-ary symbols. An M-ary modulator performs the mapping among the information bits and the M-ary symbols. Depending on the one-to-one mapping between the information bits and antenna indices, only one transmit antenna is activated in a transmission interval, and the other antennas remain turned off. The activated transmit antenna conveys the M-ary signal to the receiver. Therefore, it is sufficient to use an antenna switch and only one RF chain to execute the SM transmission at the transmitter. As a result, SM avoids inter-channel interference, requires no inter-antenna synchronization among the transmit antennas, and reduces the transceiver complexity [11].

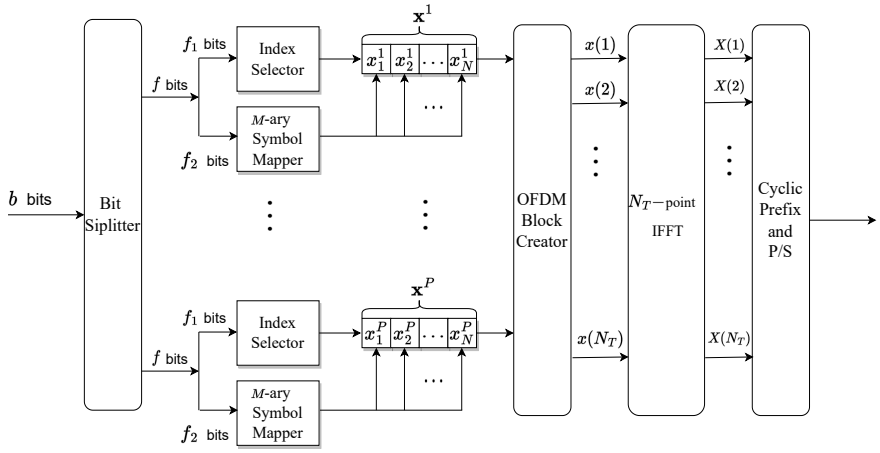


Figure 2.5. The block diagram of the OFDM-IM transmitter in [2].

Now, let us consider an SM system with N_t transmit and N_r receive antennas as shown in Fig 2.4. With $i \in \{1, \ldots, N_t\}$ denoting the active antenna index, the received signal vector in this system is given by

$$\mathbf{y} = \sqrt{\rho}\mathbf{h}_i x_m + \mathbf{w} \tag{2.39}$$

where $\mathbf{h}_i$ is the $N_r \times 1$ complex channel vector between transmit antenna i, $i \in \{1, \ldots, N_t\}$, and the receive antennas. x_m and $\mathbf{w}$ are the M-ary signal and AWGN vector, respectively, where $m = 1, \ldots, M$. Moreover, ρ is the average energy of activated antenna and it can also be denoted as the received SNR when $\mathrm{E}[|x_m|^2] = \mathrm{E}[||\mathbf{h}_i||^2] = N_0 = 1$.

The block diagram of the OFDM-IM transmitter, on the other hand, is shown in Fig. 2.5. The OFDM-IM encoder divides a total of N_T subcarriers into $P = N_T/N$ groups where the size of each group is N. In each group, $K \leq N$ subcarriers are active, while the remaining $N - K$ subcarriers are silent. Subcarrier grouping aims to decrease the implementation complexity since the number of combinations without subcarrier grouping can be excessive. At each time instant, a bit splitter splits b incoming data bits into P blocks where each block contains $f = b/P$ bits and each group of f bits is mapped to an OFDM block of length N. In each block, f bits are

Table 2.1. A look-up table implementation example for $N = 4$ and $K = 2$.

f_1 bits	Subcarrier Combinations	Indices of Active Subcarriers
[0 0]	$[\nu_1, \nu_2, 0, 0]$	$\alpha_1 = [1, 2]$
[0 1]	$[\nu_1, 0, \nu_2, 0]$	$\alpha_2 = [1, 3]$
[1 0]	$[\nu_1, 0, 0, \nu_2]$	$\alpha_3 = [1, 4]$
[1 1]	$[0, \nu_1, \nu_2, 0]$	$\alpha_4 = [2, 3]$
unused	$[0, \nu_1, 0, \nu_2]$	$\alpha_5 = [2, 4]$
unused	$[0, 0, \nu_1, \nu_2]$	$\alpha_6 = [3, 4]$

divided into f_1 and f_2 bits. f_1 bits are mapped to one of $2^{\lfloor \log_2 \binom{N}{K} \rfloor}$ combinations of active subcarriers; in other words, the indices of activated subcarriers are determined according to f_1 bits. The vector that comprises such indices is denoted by α_c where $c = 1, \ldots, \log_2 \binom{N}{K}$. f_2 bits are mapped to amplitude and/or phase modulated symbols on the activated subcarriers. We can perform the mapping between f_1 bits and the combinations/indices by using a look-up table. An example of the look-up table implementation for $N = 4$ and $K = 2$ is given in Table 2.1.[9]

Once the OFDM-IM mapper decides the activated subcarriers and the modulation symbols, the corresponding OFDM-IM symbol vector of the pth block is given by

$$\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p] \tag{2.40}$$

where $x_n^p \in \mathcal{M} \cup \{0\}$ and $p = 1, \ldots, P$. Notice that $\mathbf{x}^p$ has the same structure as $\mathbf{x}_{1,\xi}$ and a subcarrier can either carry an M-ary symbol or be silent based on this structure. After obtaining symbol vectors for all blocks, an OFDM block creator forms the overall symbol vector $\mathbf{x} := [x(1), x(2), ..., x(N_T)]^T = [\mathbf{x}^1, ..., \mathbf{x}^p, ..., \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$ where $\mathbb{C}^{N_T \times 1}$ is the set of complex numbers of the size $N_T \times 1$. After this point, the same operations as conventional OFDM are performed. Therefore, inverse fast Fourier transform, cyclic prefix addition, pulse shaping, and up-conversion are employed. At

[9]Note that some of the combinations in the table are not used as the mapping is based on fixed-length encoding. However, by using a variable-length encoding technique, one can cover all of the combinations and increase the rate of the transmission [80].

the receiver, the received signal is down-converted, filtered, and sampled. Following the cyclic prefix removal, the frequency domain received OFDM-IM vector can be written as

$$\mathbf{y} \coloneqq [y(1), y(2), \ldots, y(N_T)]^T = \sqrt{E_S}\mathbf{X}\mathbf{h} + \mathbf{w} \tag{2.41}$$

where E_S is the symbol energy and $\mathbf{X} = \text{diag}(\mathbf{x})$. Moreover, $\mathbf{h}$ and $\mathbf{w}$ are $N_T \times 1$ channel and noise vectors, respectively.

2.5.1 Achievable Data Rate

2.5.1.1 Spatial Modulation

Yang and Jiao made a solid attempt to uncover the SM capacity in [3], and they studied the capacity of an SM system with N_t transmit and one receive antennas. Considering two independent input signal spaces (symbol space, X_{sy}, and channel space, X_{ch}) and one output signal space, Y, they derived the mutual information between the input and output signals using the chain rule, i.e.,

$$\text{I}(X;Y) = \text{I}(X_{sy}, X_{ch}; Y) = \text{I}(X_{sy}; Y|X_{ch}) + \text{I}(X_{ch}; Y). \tag{2.42}$$

It is well-known that the first term in the right-hand side of (2.42) is maximized when X follows a Gaussian distribution [87]. Hence, assuming Gaussian input symbols and each transmit antenna has the same activation probability of $1/N_t$, this term can be written as $C_1 \coloneqq \max_{f_X(x)} \text{I}(X;Y|X_{ch}) = (1/N_t)\sum_{i=1}^{N_t} \log_2(1+\rho|h_i|^2)$ where $f_X(x) = \frac{1}{\pi\sigma_X^2}\exp\left(-\frac{|x|^2}{\sigma_X^2}\right)$ and σ_X^2 is the symbol variance. Here, the channel fading becomes a scalar as there is only one receive antenna. Moreover, when the modulated symbol is

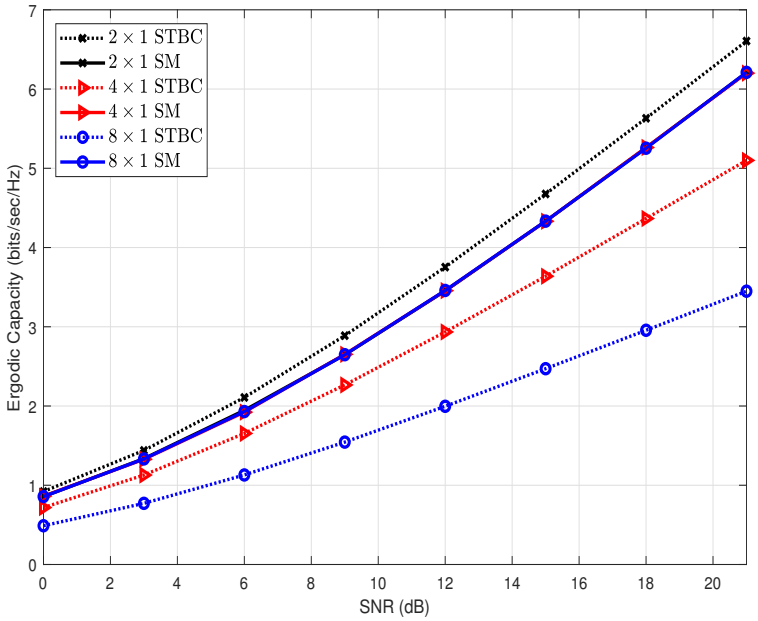


Figure 2.6. Ergotic capacity comparison of SM and STBC for $N_t \in \{2, 4, 8\}$ and $N_r = 1$. Here, the capacity curves are reproduced from [3].

Gaussian distributed, the second term in the right-hand side of (2.42) becomes [3]

$$C_2^* = \frac{1}{N_t} \sum_{i=1}^{N_t} \int_y f_{Y|X_{ch}=h_i}(y|x_{ch} = h_i) \log_2 \frac{f_{Y|X_{ch}=h_i}(y|x_{ch} = h_i)}{f_Y(y)} dy \tag{2.43}$$

where $f_{Y|X_{ch}=h_i}(y|x_{ch} = h_i) = \frac{1}{\pi\sigma_i^2} \exp\left(-\frac{|y|^2}{\sigma_i^2}\right)$, $\sigma_i^2 = |h_i|^2\sigma_X^2 + N_0$, and $f_Y(y) = \frac{1}{N_t}\sum_{i=1}^{N_t} \frac{1}{\pi\sigma_i^2} \exp\left(-\frac{|y|^2}{\sigma_i^2}\right)$. Thus, the capacity can be approximated by $C = C_1 + C_2^*$.[10]

Fig. 2.6 compares the ergodic capacity of the SM and space-time block coding (STBC) techniques for $N_t \in \{2, 4, 8\}$ and $N_r = 1$ [3]. To obtain the ergodic capacity of SM, we take the average of C over instantaneous channel fading coefficients. Here, the code rates are $1, 3/4$, and $1/2$ for the STBC schemes having two, four, and eight transmit antennas, respectively. For an STBC concept utilizing complex signals, these rates are the maximum code rates. As seen from the figure, the SM curves are on top of each other, and the STBC technique achieves a higher capacity than the SM

[10]Note that it is not clear if C is the actual SM capacity as it is not proved that the second term in the right-hand side of (2.42) is maximized when the input is Gaussian distributed. We use the superscript "*" to emphasize this important point.

technique when the number of transmit antennas is two. However, the capacity of the SM method outperforms that of the STBC scheme when the number of transmit antennas is more than two. Moreover, the superiority of SM against STBC becomes more evident when the number of transmit antennas increases. This is mainly because the decrease in the code rate results in a dramatic decrease in the capacity of the STBC method.

In [88], the authors investigate the potential gains of SM in a massive multiple-input multiple-output (MIMO) network in which the transmitter has no CSI. In this study, a closed-form upper-bound on the capacity of SM is derived, and it is shown that this upper-bound is almost completely achievable by the massive MIMO technique. Moreover, the achievable rate of SM for various input distributions is reported in [89], and *it is shown that the capacity of SM is higher than the Shannon capacity of SISO channels in the presence of a complex Gaussian input distribution.*

The authors of [90] applied Huffman coding to an adaptive SM scheme by assigning binary codes to antenna indices that take into account the frequency of occurrence of each symbol. The study explores the optimal antenna activation probabilities to obtain the capacity of the adaptive SM scheme and derives the closed-form upper and lower bound expressions on the SM capacity. The analytical and numerical results of [90] reveal that limiting the number of receive antennas to one is undesirable when sending the antenna information. Also, the adaptive SM technique in [90] outperforms the conventional SM and TAS techniques in terms of capacity and symbol error rate.

On the other hand, the authors of [91] treated channel fading coefficients as input symbols to a communication system. Thus, the resulting capacity expression does not include channel fading. To achieve the SM capacity, the multiplication of channel coefficient vector $\mathbf{h}_i$ and the input signal x_m should follow a Gaussian distribution [91].

Based on this observation, the resulting capacity expression is [91]

$$C = N_r \log_2(1+\rho). \tag{2.44}$$

Notice that this is independent of channel conditions, i.e., type of fading, and the number of transmit antennas; however, it is dependent on the number of receive antennas and SNR. Hence, it is equivalent to the AWGN capacity. Although SM has the potential to convert a fading channel to an AWGN channel and achieve even higher rate than the well-known MIMO concept, treating channel fading as input symbols may not be a valid approach since our ability to manipulate the channel distribution is limited.

It is of paramount importance to study mutual information of the IM techniques for finite alphabet inputs since assuming Gaussian input is impractical in real-time applications. In this regard, [4] investigates the mutual information of an SM concept and compares it with that of multiple-input single-output beamforming (MISO-BF) and transmit antenna selection (TAS) schemes.

In [4], a MISO-SM transmission scheme, in which the number of transmit and receive antennas is N_t and one, respectively, is considered. The mutual information for this system can be expressed based on the chain rule as in (2.42). Assuming the antennas at the transmitter are activated with equal probabilities, i.e., $1/N_t$, and the noise at the receiver is AWGN, the received signal follows a Gaussian mixture distribution. Hence, the conditional mutual information between X_{ch} and Y, when X_{sy} is given, as well as between X_{sy} and Y can be expressed, respectively, as [4]

$$\mathrm{I}(X_{ch};Y|X_{sy}) = \log_2 N_t - \frac{1}{N_t M}\sum_{i=1}^{N_t}\sum_{m=1}^{M}\mathrm{E}_w\left[\log_2\sum_{l=1}^{N_t}\exp\left(-\frac{|d_i^l+w|^2-|w|^2}{N_0}\right)\right], \tag{2.45}$$

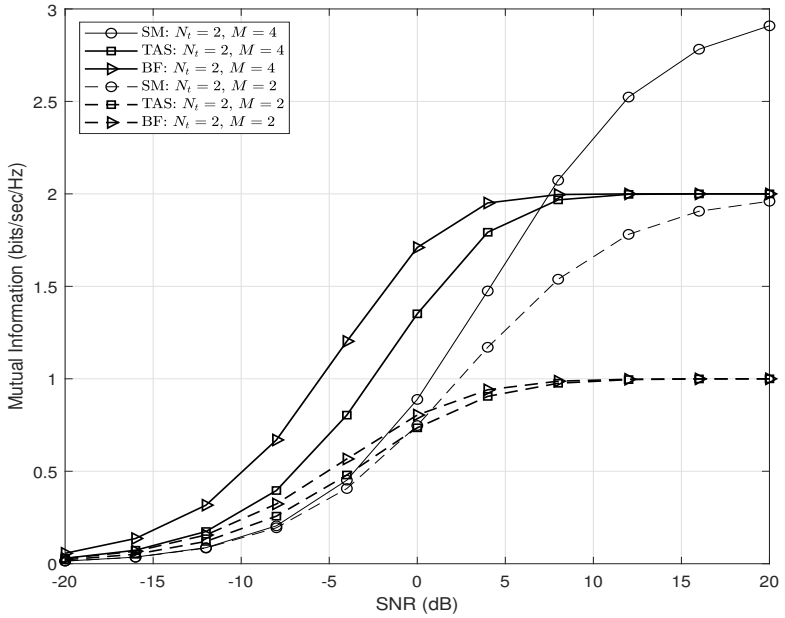


Figure 2.7. Mutual information comparison of MISO-BF, TAS, and SM for $N_t \in \{2, 4\}$ and $M \in \{2, 4\}$. Here, the mutual information curves are reproduced from [4].

$$\mathrm{I}(X_{sy}; Y) = \log_2 M - \frac{1}{N_t M} \sum_{i=1}^{N_t} \sum_{m=1}^{M} \mathrm{E}_w \left[\log_2 \frac{\sum_{l=1}^{N_t} \sum_{k=1}^{M} \exp\left(- \frac{|d_{i,m}^{l,k} + w|^2}{N_0} \right)}{\sum_{l=1}^{N_t} \exp\left(- \frac{|d_i^l + w|^2}{N_0} \right)} \right] \quad (2.46)$$

$d_{i,m}^{l,k} = h_i x_m - h_l x_k$ and $d_i^l = (h_i - h_l) x_m$. Substituting (2.45) and (2.46) into (2.42), one can obtain a mutual information expression for the SM mechanism. Such an expression can also be lower-bounded by using well-known Jensen inequality [4]. Moreover, [92] generalize these results to a MIMO-SM structure.

Fig. 2.7 compares the mutual information of MISO-BF, TAS, and SM for $N_t \in \{2, 4\}$ and $M \in \{2, 4\}$ [4]. At low SNR, MISO-BF and TAS outperform SM, since the MISO-BF and TAS techniques are capable of providing additional diversity gain. However, at high SNR, SM outperforms both of these benchmarks as well as provides additional multiplexing gain up to $\log_2 N_t$ bits/s/Hz due to the additional information carried by the antenna indices.

2.5.1.2 OFDM with Index Modulation

By denoting the random variables of signaling spaces related to the conventional symbols and subcarrier combinations as X_{sy} and X_{im}, respectively, the (ergodic) achievable rate of OFDM-IM can be given by

$$R_{OFDM-IM} = \frac{1}{N}\mathrm{E}_{\mathbf{h}}[\mathrm{I}(X;Y|\mathbf{h})] = \frac{1}{N}[\mathbb{H}(X_{sy}, X_{im}) - \mathrm{E}_{\mathbf{h}}[\mathbb{H}(X_{sy}, X_{im}|Y, \mathbf{h})]] \tag{2.47}$$

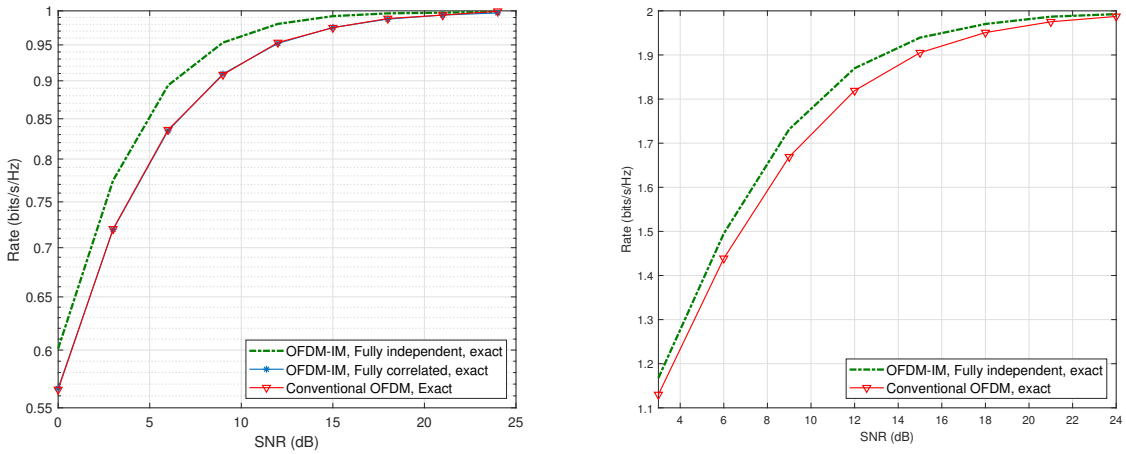
where $\mathbb{H}(\cdot)$ is the entropy operator.[11] Assuming equally likely codewords for OFDM-IM and considering the size of the OFDM-IM constellation, one can rewrite (2.47) for a finite input (M-ary symbol) as

$$R_{OFDM-IM} = \frac{K\log_2 M + \log_2 \binom{N}{K}}{N} - \frac{\mathrm{E}_{\mathbf{h}}[\mathbb{H}(X_{sy}, X_{im}|Y, \mathbf{h})]}{N}. \tag{2.48}$$

The expectation operation over the conditional entropy, $\mathrm{E}_{\mathbf{h}}[\mathbb{H}(X_{sy}, X_{im}|Y, \mathbf{h})]$, does not yield a closed-form expression and can be numerically evaluated using the Monte-Carlo method [5]. However, it is possible to obtain closed-form asymptotics and lower-bound expressions [5].

The simulation results in [5] show that OFDM-IM is capable of outperforming conventional OFDM when the number of subcarriers in a block is equal to the constellation size, i.e., $N = M$ (see Fig. 2.8 for some of these results). Also, as seen from Fig. 2.8a, OFDM-IM can yield almost the same achievable rate as conventional OFDM when the channel coefficients related to the subcarriers are fully correlated, i.e., $h_1 = h_2 = \ldots = h_N$. It is important to note that the SNR range in the figures of this thesis starts from the SNR levels where the corresponding methods cannot conduct reliable transmission or achieve high data rates. Such a range also incorporates the required SNR levels to achieve low error and high information rates. In the BER

[11] Entropy is a measure of randomness or uncertainty for a RV. For a continuous RV, this measure is called *differential entropy*, and we denote the differential entropy operator by $\mathbb{h}(\cdot)$.



(a) $M = N = 2$, $K = 1$, and $\eta = 1$.

(b) $M = N = 4$, $K = 3$, and $\eta = 2$.

Figure 2.8. Achievable rate comparison of OFDM-IM and conventional OFDM when $\eta = 1$ bps and $\eta = 2$ bps. Here, the rate curves are reproduced from [5].

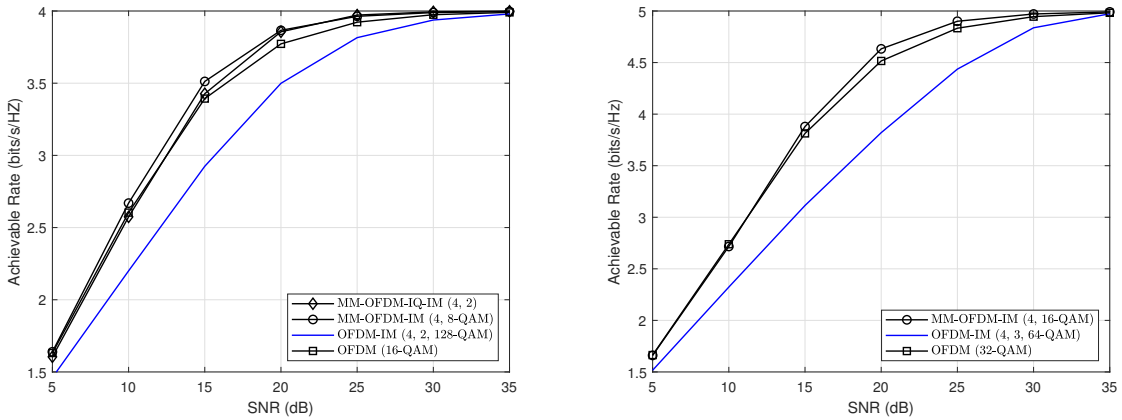
figures, this corresponds to the SNR levels where BER is less than 10^{-3}, whereas, in the rate figures, this implies the SNR levels required to achieve the information rates close to the limits of the considered methods.

On the other hand, to gain a practical insight into the achievable rate as well as the coded performance of the IM techniques, one can use the same approach as in [76] and compute the achievable rate numerically by assuming equally likely codewords and evaluating

$$R = \frac{1}{N}\left(f - \frac{1}{2^f}\sum_{i=1}^{2^f} \mathrm{E}_{\mathbf{h},\mathbf{w}}\left[\log_2 \sum_{j=1}^{2^f} e^{\delta(i,j)}\right]\right) \tag{2.49}$$

where $\delta(i,j) = \frac{-||\mathrm{diag}(\mathbf{h})(\mathbf{x}_i - \mathbf{x}_j)+\mathbf{w}||^2 + ||\mathbf{w}||^2}{N_0}$. Here, $\mathbf{x}_i$ and $\mathbf{x}_j$ are the ith and jth codewords of the IM codebook where $i, j \in \{1, \ldots, 2^f\}$.

By substituting the codewords of the multi-mode OFDM with IQ-IM (MM-OFDM-IQ-IM), MM-OFDM-IM, OFDM-IM, and OFDM techniques into (2.49), we obtain the related achievable rates and compare them in Fig. 2.9 for $\eta = 4$ bits per subcarrier (bps) and $\eta = 5$ bps. In the figure, "MM-OFDM-IM $(N, M-\mathrm{QAM})$" stands



(a) $\eta = 4$ bps (b) $\eta = 5$ bps

Figure 2.9. Achievable rate comparison of the OFDM-IM and conventional OFDM techniques when $\eta = 4$ bps and $\eta = 5$ bps.

for the MM-OFDM-IM technique with N subcarriers in each subcarrier group and N disjoint M-QAM constellations to specify the permutations, whereas "MM-OFDM-IQ-IM (N, M)" is the MM-OFDM-IQ-IM scheme of [1] with in-phase and quadrature components of N subcarriers employing N disjoint M-PAM constellations. Moreover, "OFDM $(M-$ QAM)" and "OFDM-IM $(N, K, M-$ QAM)" represent respectively the classical OFDM technique employing M-QAM and the OFDM-IM system that activates K subcarriers out of N subcarriers and carries M-ary QAM symbols on the activated subcarriers. As seen from the figure, although the OFDM-IM system fails to achieve competitive performance, the MM-OFDM-IM concepts exhibit the best performance and outperform the OFDM and OFDM-IM concepts at medium SNR. These results prove the effectiveness of the multi-mode concepts against the OFDM-IM and OFDM methods.

Now, let us focus on an ideal case where the activated subcarriers convey Gaussian distributed input symbols. Thus, the random variables related to the conventional symbols have the Gaussian distribution. Similar to (2.42), the mutual information

between the input and output of the OFDM-IM mechanism is given by

$$\mathrm{I}(X;Y) = \mathrm{I}(X_{sy};Y|X_{im}) + \mathrm{I}(X_{im};Y). \tag{2.50}$$

One can simply write the first term in the right-hand side of (2.50) as

$$\mathrm{I}(X_{sy};Y|X_{im}) = \sum_{c=1}^{C_{N,K}} \Pr(X_{im}=\alpha_c)\mathrm{I}(X_{sy};Y|X_{im}=\alpha_c) = \sum_{k=1}^{K} \log_2\left(1+\frac{E_S|h_c(k)|^2}{N_0}\right) \tag{2.51}$$

where $C_{N,K} = 2^{\lfloor \log_2 \binom{N}{K} \rfloor}$, $\Pr(X_{im}=\alpha_c)$ is the probability that the indices of activated subcarriers are α_c, and $h_c(k)$ is the channel fading coefficient associated with the subcarrier whose index is the kth element of α_c. Here, the second equation follows from the mutual information of the Gaussian input and the fact that each index set is equally likely, i.e., $\Pr(X_{im}=\alpha_c) = \frac{1}{C_{N,K}}$. Note also that the resulting expression is equivalent to the capacity of K parallel Gaussian channels.

On the other hand, the mutual information between X_{im} and Y can be expressed by [76]

$$\mathrm{I}(X_{im};Y) = \mathbb{H}(X_{im}) - \mathbb{H}(X_{im}|Y) = \log_2 C_{N,K} - \mathbb{H}(X_{im}|Y) \tag{2.52}$$

where $\mathbb{H}(X_{im}|Y)$ is not a closed-form expression and can be numerically calculated by using the Monte Carlo integration [5, 76].

It is important to stress that the mutual information of OFDM-IM in the presence of a Gaussian input cannot exceed the OFDM capacity [5]. This is clear from the fact that selecting the combinations of activated subcarriers uniformly and loading Gaussian distributed signals on the activated subcarriers will form a Gaussian mixture model. However, it is well-known that Gaussian distribution is the capacity-achieving distribution for parallel channels. As observed from Fig. 2.10, the mutual information

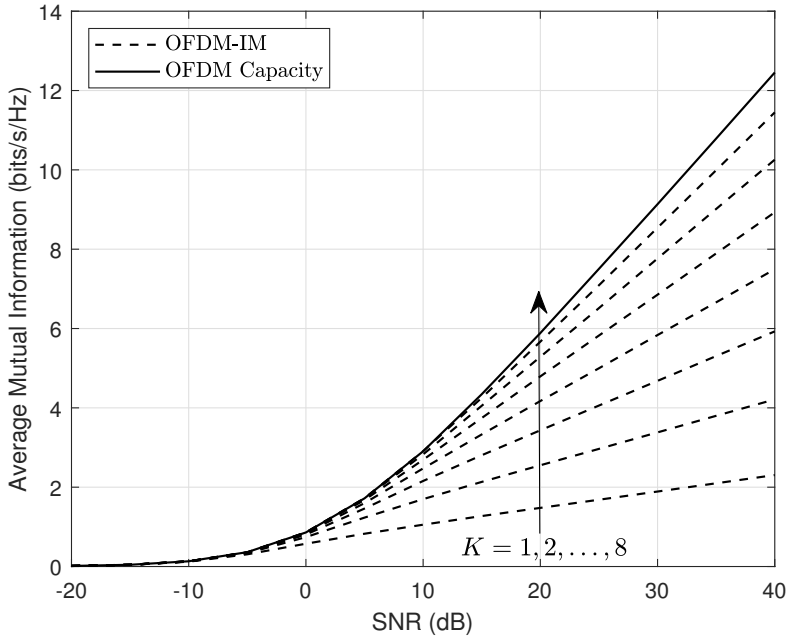
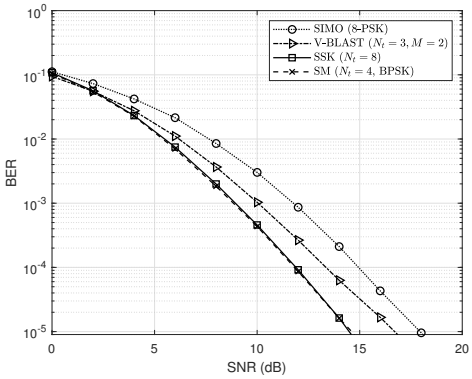


Figure 2.10. Comparison of the average mutual information of OFDM-IM and the OFDM capacity when $K = 1, \ldots, 8$. Here, the case with $K = 8$ corresponds to the OFDM capacity.

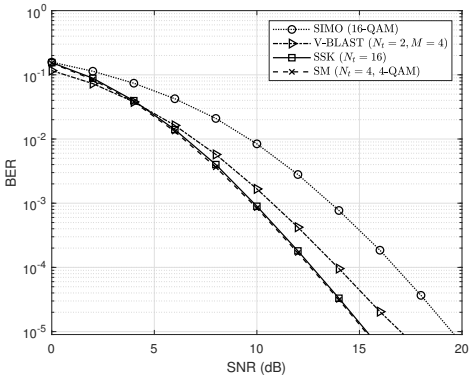
of OFDM-IM is upper-bounded by the OFDM capacity. Moreover, it increases when the number of activated subcarriers increases. This is because the information carried by the subcarrier combinations cannot compensate for the decrease in the information arising from the deactivation of subcarriers.

2.5.2 Bit Error Rate

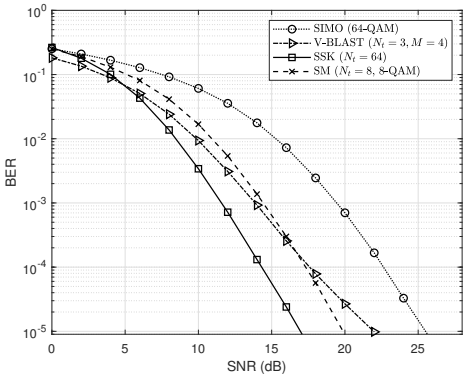
By using the union bounding technique in (2.21), one can obtain an upper-bound on the bit error probability of an IM codebook, which sums the contributions of the PEPs over all error events. However, the error probability depends also on the intersection probabilities of such error events. At low SNR, the intersection probabilities of the error events become relatively high due to relatively high noise power. Hence, the union bound remains incapable of explaining error performance at low SNR. Although such intersection probabilities become very small and the bound becomes tight at high SNR, we provide computer simulation results in this section to illustrate the BER of



(a) $\eta = 3$ bits/s/Hz



(b) $\eta = 4$ bits/s/Hz



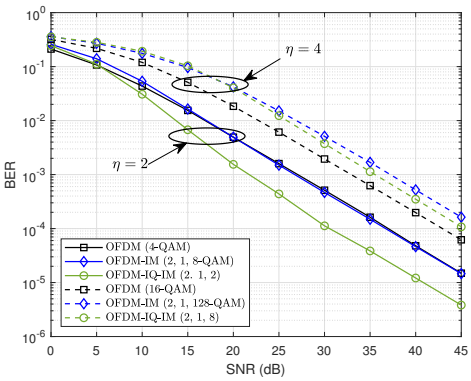
(c) $\eta = 6$ bits/s/Hz

Figure 2.11. BER comparison of the SM and SSK techniques with the SIMO and V-BLAST techniques for $N_r = 4$ and $\eta \in \{3, 4, 6\}$. Here, "V-BLAST (N_t, M)" stands for the V-BLAST system that employs M-PSK modulation over N_t transmit antennas, whereas "SIMO (M-PSK/QAM)" is the SIMO concept transmitting M-PSK or M-QAM signals over single transmit antenna. "SM (N_t, M-PSK/QAM)" and SSK (N_t) are the SM concept employing M-PSK/QAM over N_t transmit antennas and the SSK scheme having N_t transmit antennas, respectively.

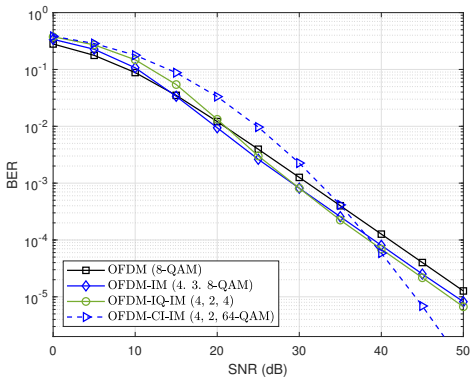
the IM concepts at both low and high SNR.

In our simulations, we assume all schemes operate over a Rayleigh fading channel, whose elements are i.i.d. RVs. Detection at the receiver is applied under the assumption that channel estimation is perfect. We also assume that the codeword elements are interleaved sufficiently so that each of them gets affected by independent fading.

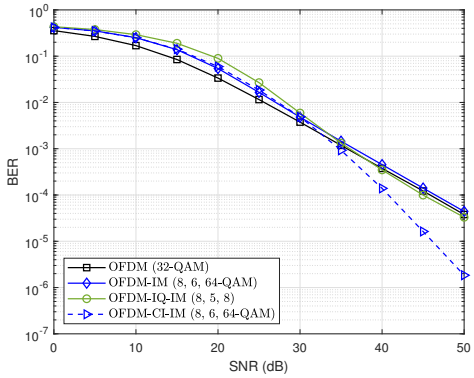
In Fig. 2.11, we compare the BER performance of the SM and SSK systems with that of the vertical Bell Laboratories layered space-time (V-BLAST) and single-



(a) $N = 2$, $\eta \in \{2, 4\}$ bps



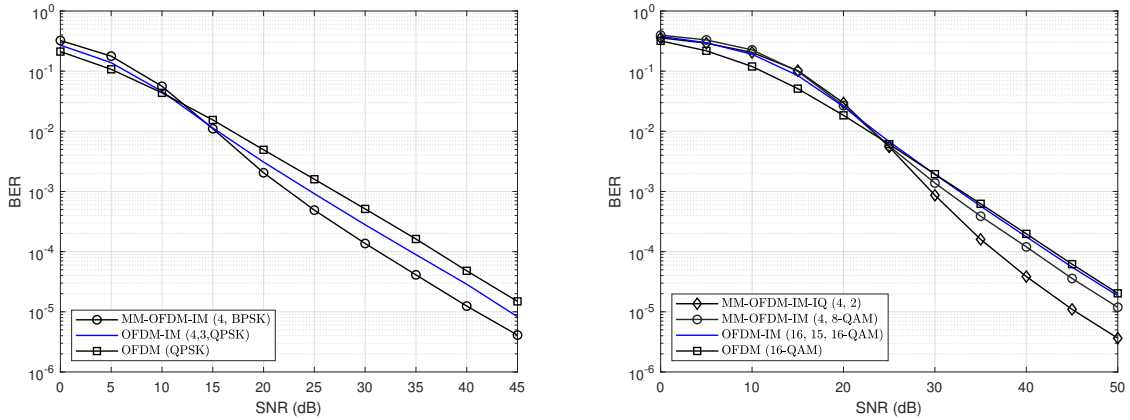
(b) $N = 4$, $\eta = 3$ bps



(c) $N = 8$, $\eta = 5$ bps

Figure 2.12. BER comparison of the OFDM-IM techniques with the conventional OFDM technique.

input multiple-output (SIMO) architectures for $\eta \in \{3, 4, 6\}$ and $N_r = 4$. Here, the V-BLAST architecture exploits the minimum mean square error (MMSE) detection based on successive interference cancellation as defined in [93], whereas the remaining concepts employ the optimal ML detection. The figure depicts that the SM and SSK systems achieve the best error performance for almost all cases and outperform the SIMO and V-BLAST systems. Although the SM scheme cannot defeat the V-BLAST mechanism for $\eta = 6$ bits/s/Hz at relatively low SNR, it starts providing a lower BER than the V-BLAST concept when SNR is around 16.5 dB. These results prove the effectiveness of the SSK and SM concepts against the conventional SIMO and V-BLAST techniques.



(a) $\eta = 2$ bps (b) $\eta = 4$ bps

Figure 2.13. BER comparison of the MM-OFDM-IM and OFDM-IM techniques with the conventional OFDM technique.

In Fig. 2.12, we compare the BER performance of the OFDM-IM, OFDM-IQ-IM, and OFDM-CI-IM concepts with that of the conventional OFDM technique for different values of SE, η, and codeword length, N. Here, we employ the optimal ML detectors for the systems in Figs. 2.12a and 2.12b, whereas we perform low-complexity LLR detectors in [2,36,37,75] for the OFDM-IM techniques in Fig. 2.12c. The OFDM-IM and OFDM-IQ-IM methods outperform the conventional OFDM at relatively low SE, i.e., $\eta \leq 3$ bps. Once we increase the SE beyond $\eta = 3$ bps, they lose the effectiveness against the classical OFDM. On the other hand, although the OFDM-CI-IM schemes in Figs. 2.12b and 2.12c are capable of achieving an additional diversity gain and outperforming all techniques at high SNR, they fail to provide competitive performance compared to the OFDM and OFDM-IM benchmarks at the practical SNR values.

We provide additional numerical results in Fig. 2.13 and compare the BER performance of the MM-OFDM-IM and OFDM-IM techniques with that of the conventional OFDM technique when $\eta = 2$ bps and $\eta = 4$ bps. Here, we employ the optimal ML detectors for the systems in Fig. 2.13a, whereas we exploit LLR detectors described in [2] and [1] for the IM concepts of Fig. 2.13b. As seen from Figs. 2.13a and 2.13b,

the multi-mode structure is beneficial since the MM-OFDM-IM techniques perform with better error performance than the OFDM-IM and OFDM techniques at medium and high SNR. Moreover, although the OFDM-IM technique cannot preserve its effectiveness at relatively high SE, this is not the case for the multi-mode counterparts.

2.5.3 Minimum Codeword Distances

Let us start with defining the *minimum L-product distance* as distance

$$d_1 = \min_{\substack{\mathbf{x},\hat{\mathbf{x}}\in\mathcal{X} \\ \mathbf{x}\neq\hat{\mathbf{x}}}} \prod_{x_n\neq\hat{x}_n}^{L} |x_n - \hat{x}_n| \tag{2.53}$$

between any two constellation points $\mathbf{x}$ and $\hat{\mathbf{x}}$ in the vector space $\mathcal{X}$ where L is the modulation diversity,[12] and x_n and $\hat{x}_n$ are the nth elements of the vectors $\mathbf{x}$ and $\hat{\mathbf{x}}$, respectively. Note that d_1 is of paramount importance for coding gain of a codebook in Rayleigh fading channels. On the other hand, the second distance that we study is the MED between the N-element codewords, i.e.,

$$d_2 = \min_{\substack{\mathbf{x},\hat{\mathbf{x}}\in\mathcal{X} \\ \mathbf{x}\neq\hat{\mathbf{x}}}} \sqrt{\sum_{n=1}^{N} |x_n - \hat{x}_n|^2}. \tag{2.54}$$

Note also that this distance is one of the main factors affecting the performance of a codebook when the channel is the AWGN channel.

In IM, constellation symbols related to the combinations of the activated (deactivated) codeword elements have a higher minimum Hamming distance than those of the conventional modulation symbols, e.g., QAM or PSK symbols [2]. More specifically, the minimum Hamming distance between these symbols is two since each combination will differ from another one in at least two positions. However, such a distance is only one for the conventional modulation symbols, limiting the minimum Hamming

[12]This is also defined as the minimum Hamming distance between the codewords of a codebook.

distance of an IM codebook to one. Thus, the minimum L-product distance of the IM techniques can be rewritten as

$$d_1 = \min_{\substack{\mathbf{x},\hat{\mathbf{x}}\in\mathcal{X} \\ \mathbf{x}\neq\hat{\mathbf{x}},x_n\neq\hat{x}_n}} |x_n - \hat{x}_n|. \tag{2.55}$$

As a result, IM maps incoming data bits to two different symbols with different distance properties. The main advantage of such a mapping is that the MED between the conventional modulation symbols becomes higher in IM than in the classic modulation concepts since IM distributes total symbol energy among only the activated codeword elements. It is important to note that one can achieve a higher minimum Hamming distance between index symbols, thus a better error performance for IM by carefully choosing the indices of active codeword elements as in [94]. However, in this case, the number of index symbols reduces and the SE is sacrificed. It is also not clear if the IM scheme can preserve its beneficial properties in terms of low-complexity implementation when the minimum Hamming distance increases.

On the other hand, following the useful design guidelines in Section III.B of [76] and assuming an ($M = 2^\eta$)-PSK/QAM constellation of unit power, d_1 of the OFDM-IM concept can be upper-bounded by

$$d_1 \leq \begin{cases} \frac{2}{\sqrt{\zeta}} \sin(\pi/2^\eta) & \text{(PSK)}, \\ \sqrt{\frac{6}{\zeta(2^\eta-1)}} & \text{(QAM)} \end{cases} \tag{2.56}$$

where η is the target SE and $\zeta = \frac{K}{N} \approx \frac{2^{\eta+1}-e}{2^{\eta+1}+e}$ is the subcarrier activation ratio. Moreover, it can be shown that $d_2 = d_1$ for almost all practical setups of OFDM-IM.

Returning our attention to the multi-mode counterparts of OFDM-IM, d_1 and d_2 of the MM-OFDM-IQ-IM technique based on the unit power assumption are given

by [1][13]

$$d_1 = \frac{2\sqrt{3}}{\sqrt{M^2 - N^{-2}}} \approx \frac{2\sqrt{3}}{M}, \tag{2.57}$$

$$d_2 = \frac{2\sqrt{6}}{\sqrt{(MN)^2 - 1}} \tag{2.58}$$

where d_2 follows from the MED between the points in an (MN)-PAM, i.e., $\frac{2\sqrt{3}}{\sqrt{(MN)^2-1}}$, and the fact that such a distance is observed across the codeword pairs that differ in two elements.

In Fig. 2.14, we conduct numerical comparisons among the minimum codeword distances of the MM-OFDM-IQ-IM, MM-OFDM-IM, OFDM-IM, and OFDM systems. For the MM-OFDM-IQ-IM and MM-OFDM-IM schemes of the figure, we assume $N = 4$ and adjust the constellation size, M, suitably to achieve a target SE, η. Then, we calculate the associated d_1 and d_2 according to [1]. On the other hand, we provide numerical distance results for the OFDM-IM technique based on the upper-bound in (2.56). Moreover, except for the MM-OFDM-IQ-IM concept, all of the techniques employ the QAM signaling. As seen from the figure, the OFDM-IM technique is capable of providing higher codeword distances than the classical OFDM technique at relatively low SE, i.e, $1 \leq \eta \leq 4$; however, the superiority of the OFDM-IM scheme against the OFDM method vanishes when $\eta > 4$. In the meantime, the MM-OFDM-IM and MM-OFDM-IQ-IM techniques can deliver considerably higher L-product distance, d_1, than the OFDM-IM and OFDM concepts, while their MEDs, d_2, are comparable to that of these concepts.

[13]Here, to keep the discussion brief, we choose to show the distance expressions related to the MM-OFDM-IQ-IM scheme of [1]. However, one can find similar expressions in [1] related to the MM-OFDM-IM concept applying disjoint M-PSK/QAM constellations.

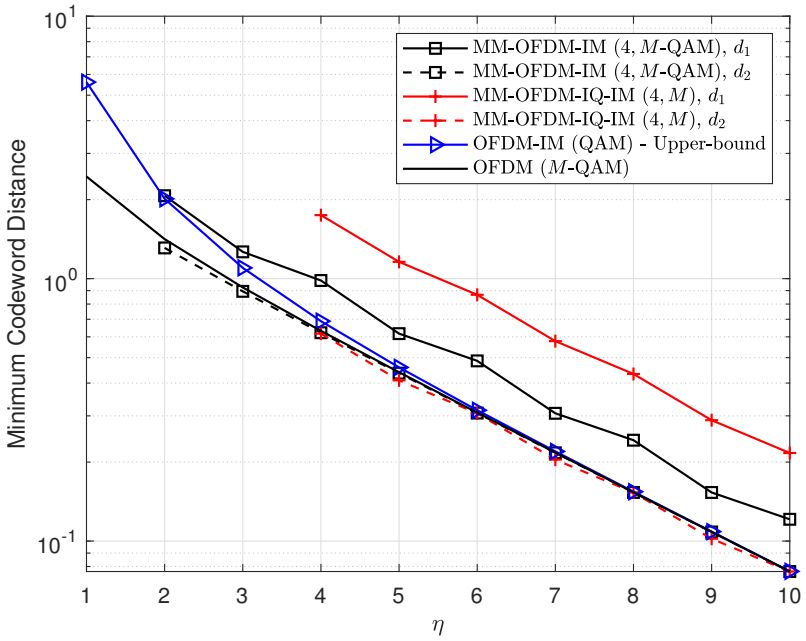


Figure 2.14. Comparison of the minimum codeword distances for MM-OFDM-IQ-IM, MM-OFDM-IM, OFDM-IM, and OFDM.

2.5.4 Detection Complexity

The optimum ML detection of the SM techniques becomes impractical when the constellation size is large, i.e., the number of transmit antennas and points in the M-ary constellation is high. Thus, to achieve high SEs at an affordable complexity, these techniques need to perform low-complexity detection algorithms at the receiver. In this regard, there exist several detection algorithms in the SM literature based on well-known methods such as compressed sensing [95, 96] and sphere decoding [97, 98]. Among these algorithms, the sphere decoding-based detection algorithms seem to be the most beneficial ones since they provide near ML performance with considerably reduced complexity [98, 99]. Also, the sphere decoding-based SM structure is capable of attaining an error performance comparable to that of the V-BLAST method with sphere decoding while having a much lower detection complexity [97].

On the other hand, the OFDM-IM concepts require low-complexity subcarrier-wise detectors to overcome the prohibitive complexity of the optimum ML detection.

Table 2.2. Comparison of detection complexities for OFDM-IM, MM-OFDM-IM, and OFDM.

System Parameters	OFDM-IM (N, K, M_1 – PSK/QAM)		MM-OFDM-IM (N, M_2 – PSK/QAM)	OFDM (M_3 – PSK/QAM)
	ML	LLR	LLR	ML
$N = 4, K = 3, M_1 = 4, M_2 = 2, M_3 = 4$	64	4	5	4
$N = 8, K = 4, M_1 = 91, M_2 = 5, M_3 = 16$	5.37×10^8	91	22.5	16
$N = 16, K = 12, M_1 = 115, M_2 = 8, M_3 = 54$	3.09×10^{26}	115	68	54

In this regard, the LLR detectors of the OFDM-IM schemes are expedient as they achieve near-optimum error performance with considerably reduced complexity. The OFDM-IM technique uses the LLR values to decide the activated subcarriers [2, 36], whereas the multi-mode schemes use them to determine disjoint constellations carried by a subcarrier [1, 38]. The orders of complexity in terms of the number of metric calculations per subcarrier for the LLR detectors of the OFDM-IM and MM-OFDM-IM techniques are $O(M_1)$ and $O(M_2N/2 + M_2/2)$, while the complexity of the classical OFDM concept is of order $O(M_3)$. Here, M_1, M_2, and M_3 are the size of the conventional modulation for the OFDM-IM, MM-OFDM-IM, and OFDM schemes.[14] To understand how these complexities scale compared to one another, we provide numerical complexity results regarding the ML detector of OFDM-IM, the LLR detectors of OFDM-IM and MM-OFDM-IM, and the ML detector of OFDM in Table 2.2.[15] As seen from the table, the LLR detectors are much less complex than the ML detector of the OFDM-IM system, and they have comparable complexity to that of the ML detection in OFDM.

2.6 Concluding Remarks

In this chapter, we provided an overview of digital communications by illustrating the basic structure and defining baseband modulation and demodulation as well as

[14]Here, we use different letters to specify the modulation orders of the concepts since they usually need to employ unequal modulation orders to achieve the same SE.

[15]To send the same number of the information bits for each scheme, thus to make fair comparisons, we adjust M_1, M_2, and M_3. The associated values in the table may not be practical due to not being a power of two; however, the table aims to provide insights into the complexity orders.

channel models. Then, we briefly described the channel capacity and focused on the design criteria required to make a constellation desirable in the fading channels. In this regard, we showed examples of well-known constellations in one and more than one dimension. Finally, we turned our attention to combinatorial modulation concepts that have desirable properties in terms of the design criteria.

As we demonstrated in this chapter, the combinatorial structure enables the IM techniques to achieve desirable properties in terms of achievable rate, BER, codeword distance, and low-complexity detection. Motivated by these properties, we propose more reliable and spectrally efficient modulation concepts by uncovering different combinatorial tools such as set partitions and integer combinations. To see the actual potential of the combinatorial mechanisms, we further investigate these concepts and provide even better modulation concepts based on a well-known coding mechanism.

Chapter 3

Encoding Using Set Partitions

As mentioned in the previous chapter, one can design promising modulation methods by exploiting combinations or permutations in encoding. More specifically, IM uses the combinations of activated/deactivated codeword elements and achieves a better performance than the classical modulation techniques in terms of BER and achievable rate. Moreover, MM-IM utilizes the permutations of codeword elements drawn from disjoint constellations and provides superior error and rate performance against the classical modulation and IM methods. In addition to IM and MM-IM, several modulation concepts benefit from combinatorial encoding and the application of disjoint constellations. Although most of these concepts have competitive error and rate performance, they exploit only permutations or combinations of codeword elements in encoding. They do not fully explore the potential of combinatorial encoding as they lack the application of other tools, e.g., set and integer partitioning.

Motivated by this fact and the beneficial properties of having a combinatorial structure, this chapter uncovers the application of set partitioning in encoding. In this regard, we develop a new codebook design method, which we name *set partition modulation* (SPM), and show the performance of SPM against the IM and classical benchmarks. Our efforts prove that the SPM concept could be even more beneficial

than the MM-IM and IM concepts in terms of error and rate performance.

The contributions of this chapter can be summarized comprehensively as follows:

- The proposed method uses a novel combinatorial tool, which is partitions of codeword elements rather than permutations or combinations of such elements. By using the set partitioning, we aim to design a more general codebook that can achieve a better error performance as well as a higher data rate than the conventional modulation and IM techniques.

- We define different variants of SPM and give a practical model for applying SPM in parallel channels.

- The proposed SPM schemes employ distinguishable constellations, similar to MM-IM; however, in SPM, the different constellations are used to distinguish between subsets in a set partition, unlike MM-IM where distinguishable constellations are used to construct permutations.

- The proposed SPM variants are capable of encompassing DM-IM and MM-IM schemes as special cases. Moreover, they are also capable of achieving a better data rate than such schemes.

- We define a codebook selection problem for the SPM variants to design efficient codebooks, which are at least as good as the IM methods. We formulate such a problem as a clique problem in graph theory and develop an efficient solution for such a problem.

- The achievable data rate and BER of the SPM variants are investigated in this chapter, and an upper-bound on the BER is obtained. We also provide asymptotic expressions for data rate and BER.

- Our analytical findings show that the SPM variants can outperform conventional modulation, IM, DM-IM, and MM-IM in terms of data rate and BER.

Apart from the beneficial properties listed above, the proposed SPM concepts have a few points that need more investigation to make these concepts more practical. To the best of the author's knowledge, there does not exist any number system that maps set partitions to decimal numbers or binary digits. Therefore, one has to use look-up tables to map the information bits to set partitions at the transmitter and the decoded partitions to the information bits at the receiver. This situation may limit the application of the proposed methods when the codebook size is large. That is because the complexity of the look-up table implementation becomes high when mapping the information bits to a large codebook.

On the other hand, to calculate the optimum LLR values related to the bits of the SPM techniques, the receiver must go through all SPM codewords and calculate the squared Euclidean distances. Therefore, such a calculation can be prohibitive when the SPM codebooks have many codewords. Moreover, calculating the LLR values in a way that exhibits low complexity and high reliability is a challenging problem that requires further research. The lack of practical algorithms may limit the application of the SPM techniques when there is channel encoding in the system.

3.1 Basic Definitions

Before going into details of SPM, let us provide some useful definitions which will make it easier for us to describe the SPM methods.

Definition 2. Set Partition: A set partition is the grouping the elements of a set in a way that the groups are disjoint and the union of the groups gives the set. The set $\{u_1, u_2, u_3\}$, for example, has the following five partitions $\{\{u_1, u_2, u_3\}\}$, $\{\{u_1\}, \{u_2, u_3\}\}$, $\{\{u_2\}, \{u_1, u_3\}\}$, $\{\{u_3\}, \{u_1, u_2\}\}$, and $\{\{u_1\}, \{u_2\}, \{u_3\}\}$.

Definition 3. Stirling Number of the Second Kind: The Stirling number of the second kind, denoted by $\left\{{N \atop K}\right\}$, can be defined as the number of ways to partition an N-

element set into K non-empty subsets. For example, the set $\{u_1, u_2, u_3\}$ has the following three set partitions with two non-empty subsets $\{\{u_1\}, \{u_2, u_3\}\}, \{\{u_2\}, \{u_1, u_3\}\}$ and $\{\{u_3\}, \{u_1, u_2\}\}$, thus $\left\{{3 \atop 2}\right\} = 3$.

Definition 4. Bell Number: The Bell number B_N enumerates the total number of partitions of an N-element set. The Bell number is related to Stirling numbers of the second kind through the equation $B_N = \sum_{K=1}^{N} \left\{{N \atop K}\right\}$. Since a three-element set has five partitions, $B_3 = 5$.

Definition 5. Ordered Bell Number: The ordered Bell number $\breve{B}_N$ enumerates the total number of partitions of an N-element set considering all permutations of subsets for each partition. The ordered Bell number satisfies the equation $\breve{B}_N = \sum_{K=1}^{N} K! \left\{{N \atop K}\right\}$. A three-element set has one set partition with one subset, three set partitions with two subsets, and one set partitions with three subsets. Hence, the ordered Bell number, $\breve{B}_3$, can be calculated as $\breve{B}_3 = 1!(1) + 2!(3) + 3!(1) = 13$.

3.2 Set Partition Modulation

3.2.1 SPM

In an SPM system, a codebook of L_{SPM} codewords $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{L_{SPM}}$ is constructed such that each codeword is a sequence of N elements, which are drawn from a constellation diagram in the complex plane, i.e., $\mathbf{x}_l = \{x_{l1}, x_{l2}, \ldots, x_{lN}\}$ where $x_{ln} \in \mathbb{C}$, $l = 1, \ldots, L_{SPM}$, and $n = 1, \ldots, N$. Each codeword is mapped to a partition of an N-element set $\mathcal{U}$ into K subsets. Since the number of ways one can form the partitions is $\left\{{N \atop K}\right\}$, the SPM codebook size is given by $L_{SPM} = \left\{{N \atop K}\right\}$. We call $\mathcal{U}$ the generator set of an SPM codeword. To produce unique codewords for SPM, an N-element codeword should have at least K distinguishable elements, and each element, $\mu_k, k = 1, \ldots, K$, is used to specify which element in the codewords belongs to which subset. Therefore,

distinguishable elements differentiate each subset from the other(s).

3.2.2 Ordered SPM

Ordered SPM (OSPM) is an extended version of SPM in which the codebook is enlarged by considering permutations of the subsets, thus the codebook size is given by $L_{OSPM} = K!\left\{{N \atop K}\right\}$.

3.2.3 Full SPM

In Full SPM (FSPM), the codewords $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{L_{FSPM}}$ are generated by partitioning an N-element set $\mathcal{U}$ into non-empty disjoint subsets in such a way that the number of these subsets takes any possible value, $K \in \{1, \ldots, N\}$. In other words, all partitions of an N-element set $\mathcal{U}$ into non-empty disjoint subsets are used to form the FSPM codebook. For FSPM, the codebook size is equal to the Bell number B_N, i.e., $L_{FSPM} = B_N = \sum_{K=1}^{N} \left\{{N \atop K}\right\}$.

3.2.4 Ordered Full SPM

We can further increase the number of codewords in an FSPM codebook by considering the permutations of the subsets in a partition. In this regard, we define ordered full SPM (OFSPM) as a modulation scheme that forms its codebook by using all partitions of an N-element set $\mathcal{U}$ along with all permutations of the subsets in each partition. Hence, the OFSPM codebook size is given by the ordered Bell number $\breve{B}_N$, i.e., $L_{OFSPM} = \breve{B}_N = \sum_{K=1}^{N} K!\left\{{N \atop K}\right\}$.

Example 1. In Table 3.1,[1] we list unordered and ordered set partitions of a three-element set $\mathcal{U} \coloneqq \{u_1, u_2, u_3\}$ to give an example of a mapping between the variants of SPM and the corresponding set partitions for $N = 3$. For SPM and OSPM, we set

[1] In the table, "-" signifies that the specific SPM method (SPM, OSPM, FSPM, or OFSPM) does not have a symbol constructed from the corresponding set partition.

Table 3.1. An example of codeword generation for the SPM variants when $N = 3$ and $K = 2$.

	SPM	OSPM	FSPM	OFSPM
$\mathcal{S}_1 = \{\{u_1, u_2, u_3\}\}$	-	-	$\mathbf{x}_1 = \{\mu_1, \mu_1, \mu_1\}$	$\mathbf{x}_1 = \{\mu_1, \mu_1, \mu_1\}$
$\mathcal{S}_2 = \{\{u_1\}, \{u_2, u_3\}\}$	$\mathbf{x}_1 = \{\mu_1, \mu_2, \mu_2\}$	$\mathbf{x}_1 = \{\mu_1, \mu_2, \mu_2\}$	$\mathbf{x}_2 = \{\mu_1, \mu_2, \mu_2\}$	$\mathbf{x}_2 = \{\mu_1, \mu_2, \mu_2\}$
$\mathcal{S}_3 = \{\{u_2\}, \{u_1, u_3\}\}$	$\mathbf{x}_2 = \{\mu_2, \mu_1, \mu_2\}$	$\mathbf{x}_2 = \{\mu_2, \mu_1, \mu_2\}$	$\mathbf{x}_3 = \{\mu_2, \mu_1, \mu_2\}$	$\mathbf{x}_3 = \{\mu_2, \mu_1, \mu_2\}$
$\mathcal{S}_4 = \{\{u_3\}, \{u_1, u_2\}\}$	$\mathbf{x}_3 = \{\mu_2, \mu_2, \mu_1\}$	$\mathbf{x}_3 = \{\mu_2, \mu_2, \mu_1\}$	$\mathbf{x}_4 = \{\mu_2, \mu_2, \mu_1\}$	$\mathbf{x}_4 = \{\mu_2, \mu_2, \mu_1\}$
$\mathcal{S}_5 = \{\{u_2, u_3\}, \{u_1\}\}$	-	$\mathbf{x}_4 = \{\mu_2, \mu_1, \mu_1\}$	-	$\mathbf{x}_5 = \{\mu_2, \mu_1, \mu_1\}$
$\mathcal{S}_6 = \{\{u_1, u_3\}, \{u_2\}\}$	-	$\mathbf{x}_5 = \{\mu_1, \mu_2, \mu_1\}$	-	$\mathbf{x}_6 = \{\mu_1, \mu_2, \mu_1\}$
$\mathcal{S}_7 = \{\{u_1, u_2\}, \{u_3\}\}$	-	$\mathbf{x}_6 = \{\mu_1, \mu_1, \mu_2\}$	-	$\mathbf{x}_7 = \{\mu_1, \mu_1, \mu_2\}$
$\mathcal{S}_8 = \{\{u_1\}, \{u_2\}, \{u_3\}\}$	-	-	$\mathbf{x}_5 = \{\mu_1, \mu_2, \mu_3\}$	$\mathbf{x}_8 = \{\mu_1, \mu_2, \mu_3\}$
$\mathcal{S}_9 = \{\{u_1\}, \{u_3\}, \{u_2\}\}$	-	-	-	$\mathbf{x}_9 = \{\mu_1, \mu_3, \mu_2\}$
$\mathcal{S}_{10} = \{\{u_2\}, \{u_1\}, \{u_3\}\}$	-	-	-	$\mathbf{x}_{10} = \{\mu_2, \mu_1, \mu_3\}$
$\mathcal{S}_{11} = \{\{u_2\}, \{u_3\}, \{u_1\}\}$	-	-	-	$\mathbf{x}_{11} = \{\mu_3, \mu_1, \mu_2\}$
$\mathcal{S}_{12} = \{\{u_3\}, \{u_1\}, \{u_2\}\}$	-	-	-	$\mathbf{x}_{12} = \{\mu_2, \mu_3, \mu_1\}$
$\mathcal{S}_{13} = \{\{u_3\}, \{u_2\}, \{u_1\}\}$	-	-	-	$\mathbf{x}_{13} = \{\mu_3, \mu_2, \mu_1\}$

$K = 2$. As seen from the table, to obtain the codewords in an SPM codebook, we first partition the elements of the generator set $\mathcal{U}$ into subsets, and then permute the orders of the subsets if we are constructing the codewords for one of the ordered counterparts of SPM, i.e., OSPM or OFSPM. The unordered and ordered set partitions, $\mathcal{S}_l$, are used to form SPM codewords, $\mathbf{x}_l$ where $l = 1, \ldots, L_S$ and $L_S = 3, 6, 5$, and 13 for SPM, OSPM, FSPM and OFSPM, respectively, since $\left\{{3 \atop 2}\right\} = 3, 2!\left\{{3 \atop 2}\right\} = 6, B_3 = 5$, and $\breve{B}_3 = 13$. Then, we use the subset identifier μ_n, $n \in \{1, \ldots, N\}$, to represent the elements that belong to the nth subset. For example, for the first codeword of SPM in Table 3.1, we have the partition $\mathcal{S}_2 \coloneqq \{\{u_1\}, \{u_2, u_3\}\}$. Since the element u_1 and the elements u_2 and u_3 are in the first and second subsets, respectively, we assign μ_1 to first element and μ_2 to remaining elements, i.e., $\mathbf{x}_1 = \{\mu_1, \mu_2, \mu_2\}$. For OSPM, the first three codewords are SPM codewords and we further obtain an extended codebook by taking the permutations of the subsets in a set partition into account. In FSPM, we consider all set partitions without permuting the orders of the subsets. However, in OFSPM, we considerably expand the codebook by considering all set partitions as well as the permutations of the subsets in those of the partitions.

3.3 Practical Model for Parallel Channels

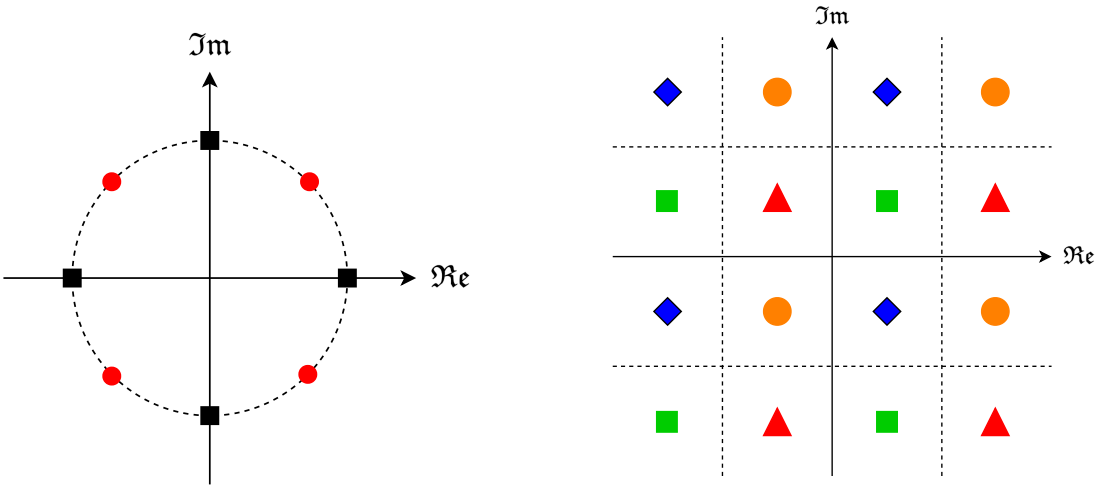
We present a practical system model in which we apply the SPM concepts to parallel channels.

3.3.1 Transmitter

In this model, similar to the OFDM-IM techniques, b input bits enter the SPM transmitter, and these bits are divided into $P = b/f$ blocks, each having f input bits. Similarly, the total number of parallel channels N_T is also divided into $P = N_T/N$ blocks, each having N channels. For each block of input bits, f information bits are modulated by an SPM encoder, and the resulting modulated symbols are conveyed over N channels.

Since each bit and each channel block have the same mapping operation, we focus on a single block, the pth block (where $p \in \{1, 2, \ldots, P\}$), in what follows. In the pth block, the SPM encoder further divides f information bits into two parts, one of them having f_1 bits and the other one having f_2 bits with $f_1 + f_2 = f$. The first f_1 bits are used to determine the specific set partition $\mathcal{S}_l$, $l = 1, \ldots, L_S$ ($L_S \in \{L_{SPM}, L_{OSPM}, L_{FSPM}, L_{OFSPM}\}$), of the N-element generator set $\mathcal{U} \coloneqq \{u_1, u_2, \ldots, u_N\}$ belonging to one of the SPM variants defined above. The chosen partition in the pth block is mapped to the corresponding SPM codeword $\mathbf{x}_l$. The remaining f_2 bits are used to modulate symbols on the N channels, considering the corresponding mapping of the set partition determined by the first f_1 bits. As discussed in the previous subsection, we use different subset identifiers in order to produce unique SPM codewords. To preserve the uniqueness property of the SPM codewords and modulate the symbols on each channel, we further assume that each subset identifier μ_k is an element of a disjoint M-ary signal constellation $\mathcal{M}_k$, i.e., $\mu_k \in \mathcal{M}_k$ and $\mathcal{M}_k \cap \mathcal{M}_{\hat{k}} = \emptyset$, where $k, \hat{k} \in \{1, 2, \ldots, N\}$ and $k \neq \hat{k}$.[2] For convenience, we choose the size of each constel-

[2]Note that the subset identifiers are not necessarily elements of disjoint M-ary signal constella-



(a) Two disjoint QPSK constellations. (b) Four disjoint 4-QAM constellations.

Figure 3.1. An example of disjoint PSK and QAM constellations.

lation as M, thus $f_2 = N \log_2 M$. By following the useful design guidelines in [1], we obtain the distinguishable PSK constellations $\{\mathcal{M}_k\}$ by rotating an ordinary M-PSK constellation with the angle of $2(k-1)\pi/(MN)$, $k = 1, \ldots, N$, to maximize the distance between constellation points. To obtain distinguishable QAM constellations, likewise [1], we employ the well-known set partitioning technique in [100]. To better illustrate how we form the disjoint constellations, we give an example of two disjoint QPSK constellations and four disjoint 4-QAM constellations in Fig. 3.1. Here, the constellation points with the same color and marker are related to the same disjoint constellation. One can observe from Fig. 3.1a that the second QPSK constellation is obtained by rotating the other one with an angle of 45°. Fig. 3.1b shows how we attain four disjoint 4-QAM constellations by set partitioning the classical 16-QAM constellation.

The mapping of f_1 bits to the set partitions can be performed by using a look-

tions and each of them may be chosen as a single constellation point in the same M-ary constellation. For example, four constellation points in a single QPSK modulation can be used to specify the subsets of set partitions regarding a four-element set. Hence, assigning a unique constellation point to each subset identifier would be enough to constitute an SPM scheme. However, in this special case, the number of information bits transmitted by a block is decreased by f_2 bits since f_2 bits are not used to modulate the subset identifiers.

up table in a similar manner as was proposed to map information bits to subcarrier activation patterns in [2], or to permutation indices as detailed in [1]. Once the mapping between f_1 bits and set partitions has been completed, $f_2 = N \log_2 M$ bits are used to determine the modulated symbols, or in other words subset identifiers, on each channel. Hence, one of the SPM codewords along with the corresponding modulation symbols $\{\mu_k\}$ constitutes the symbol vector $\mathbf{x}^p \in \{\mathbf{x}_1, \ldots, \mathbf{x}_{L_S}\}$ in the pth block. After obtaining symbol vectors for all blocks, a block creator forms the overall symbol vector $\mathbf{x} \coloneqq [x(1), x(2), \ldots, x(N_T)]^T = [\mathbf{x}^1, \ldots, \mathbf{x}^p, \ldots, \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$. Here, we assume that each element of $\mathbf{x}$ is distributed among equally spaced channels to ensure statistical independence, and each modulated symbol carried by a channel has unit average energy, i.e., $\mathrm{E}[|x(t)|^2] = 1$, $t = 1, \ldots, N_T$.

One can employ the SPM encoding over any slotted transmission mechanism since the codeword elements can be independently sent over different slots. For example, one can use time or frequency slots to send the codeword elements related to the SPM symbols. Moreover, we can also find antenna-domain applications where separate antennas convey different codeword elements. Therefore, the SPM technique can find applications in well-known transmission techniques such as OFDM, time division multiplexing (TDM), and V-BLAST. More specifically, one can use N subcarriers in an OFDM system, N time slots in a TDM technique, and N transmit antennas in a V-BLAST method to convey an SPM codeword of length N.

Example 2. In Fig. 3.2, we show a practical application of the SPM mechanism in parallel channels where the SPM encoding is performed over the subcarriers of the well-known OFDM technique. The resulting technique is called *OFDM with SPM (OFDM-SPM)*. Here, we divide the incoming information bits and subcarriers of the OFDM method into P groups. Each group has N subcarriers and $f = f_1 + f_2$ information bits. We map the first f_1 bits to the N-element SPM codewords and corresponding disjoint constellations. Then, in each group, we use the remaining f_2 bits to modu-

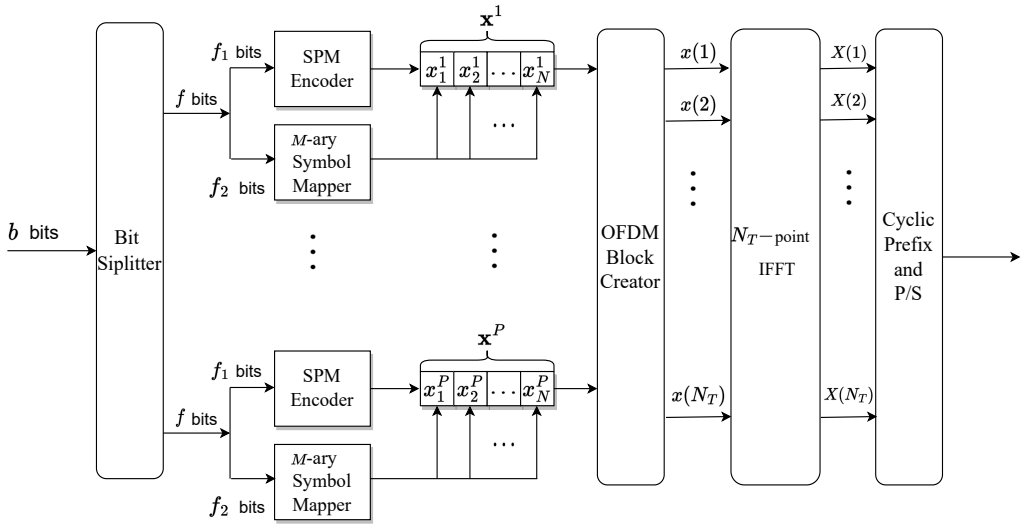


Figure 3.2. Transmitter structure of the OFDM-SPM scheme.

late the symbols on the subcarriers according to the disjoint constellations determined by the first f_1 bits. Each codeword element of the OFDM-SPM scheme, x_n^p, is conveyed by a different subcarrier where $n \in \{1, 2, \ldots, N\}$ and $p \in \{1, 2, \ldots, P\}$. Therefore, the resulting OFDM symbol $\mathbf{x} = [x_1^1, x_2^1, \ldots, x_N^1, x_1^1, x_2^2, \ldots, x_N^2, \ldots, x_1^P, x_2^P, \ldots, x_N^P]$ is conveyed by the subcarriers whose center frequencies are $f_1^1, f_2^1, \ldots, f_N^1, f_1^2, f_2^2, \ldots, f_N^2, \ldots, f_1^P, f_2^P, \ldots, f_N^P$ where f_n^p is the center frequency of the nth subcarrier in the pth group.[3] After the SPM encoding, exactly the same operations as conventional OFDM are applied in OFDM-SPM. The symbol vector is processed with an N_T-point IFFT, and a cyclic prefix of sufficient length, which is not lower than the memory of the discrete channel impulse response, is attached to the beginning of each time-domain symbol vector. After parallel-to-serial and up-conversion, the transmission is operated over a frequency-selective Rayleigh fading channel.

[3]The OFDM technique is widely used in the IEEE 802.11 family of standards. In this family of standards, typical operation bands are 2.4 and 5 GHz bands. As a member of this family, IEEE 802.11ac, also known as Wi-Fi 5, employs the OFDM technique with a subcarrier spacing of 312.5 kHz and supports channel bandwidths of 20, 40, 80, and 160 MHz. Here, the subcarrier spacing and channel bandwidth signify the gap between the center frequencies of adjacent subcarriers and the bandwidth of an OFDM symbol, respectively.

Algorithm 1: The LC-ML detection of the SPM techniques

Step 1: Initialize $1 \times N$ decoding vector, $\mathbf{d}$, of all zeros and $L_S \times N$ index matrix $\mathbf{Z}$;
Step 2: Sort the channel gains related to the codeword elements in descending order, i.e., $|h(\lambda_1)|^2 \geq \ldots \geq |h(\lambda_N)|^2$ where $\lambda_n \in \{1, \ldots, N\}$;
Step 3: LC-ML detection;
for $n = 1$ **to** N **do**
  $\mathcal{I}_n \coloneqq \text{unique}(\mathbf{Z}(:, \lambda_n))$;
  $\mathcal{M} \coloneqq \cup_{i \in \mathcal{I}_n} \mathcal{M}_i$;
  $\hat{x}_{\lambda_n} = \arg\min_{x \in \mathcal{M}} |y(\lambda_n) - \sqrt{E_S} x h(\lambda_n)|^2$, $\hat{I}_{\lambda_n} = \arg_{i \in \mathcal{I}_n} \hat{x}_{\lambda_n} \in \mathcal{M}_i$;
  $\mathbf{r} = \text{find}(\mathbf{Z}(:, \lambda_n) == \hat{I}_{\lambda_n})$;
  $\mathbf{Z} = \mathbf{Z}(\mathbf{r}, :)$;
  $\mathbf{d}(\lambda_n) = \hat{x}_{\lambda_n}$;
  $n \leftarrow n + 1$;
end
return $\mathbf{d}$;

3.3.2 Receiver

At the receiver, the received signal vector can be written as

$$\mathbf{y} \coloneqq [y(1), y(2), \ldots, y(N_T)]^T = \sqrt{E_S}\mathbf{X}\mathbf{h} + \mathbf{w} \tag{3.1}$$

where E_S is the energy of the transmitted symbol vector and $\mathbf{X} = \text{diag}(\mathbf{x})$. Moreover, $\mathbf{h}$ and $\mathbf{w}$ are $N_T \times 1$ channel and noise vectors, respectively.

Since the encoding procedure for each block is independent of others, decoding can be performed independently at the receiver. Hence, using the ML detection, the detected symbol vector for the pth block can be written as

$$\hat{\mathbf{x}}^p = \arg\min_{\mathcal{S}_l, \mu_k} ||\mathbf{y}^p - \sqrt{E_S}\mathbf{X}^p\mathbf{h}^p||^2 \tag{3.2}$$

where $\mathbf{y}^p = [y((p-1)N+1), \ldots, y(pN)]^T$, $\mathbf{X}^p = \text{diag}(\mathbf{x}^p)$, and $\mathbf{h}^p = [h((p-1)N+1), \ldots, h(pN)]^T$.

The optimum ML detector in (3.2) performs $8L_S M^N$ real multiplications since we

have L_S SPM as well as M^N modulation symbols on N codeword elements, and we perform eight real multiplications for each squared Euclidean distance calculation. To overcome the excessive complexity of the optimum ML detector when M is high, we design a look-up table-based low-complexity ML (LC-ML) detector. The proposed LC-ML detector for the SPM schemes is summarized in Algorithm 1.[4] The algorithm starts by initializing a $1 \times N$ decoding vector, $\mathbf{d}$, of all zeros and an $L_S \times N$ index matrix, $\mathbf{Z}$, which contains the indices of disjoint constellations that define all admissible partitions.[5] Since there are N distinguishable constellations, the lth row and nth column element, z_{ln}, of $\mathbf{Z}$ is drawn from the set $\{1, 2, \ldots, N\}$, i.e., $z_{ln} \in \{1, 2, \ldots, N\}$. Then, the channel gains regarding the codeword elements are sorted in descending order, i.e., $|h(\lambda_1)|^2 \geq \ldots \geq |h(\lambda_N)|^2$ where $\lambda_n \in \{1, \ldots, N\}$. Starting from the codeword element with the largest channel gain, the algorithm determines the possible constellation indices that a codeword element can adopt. To do that, the algorithm looks at the unique elements in the λ_nth column of the index matrix, $\mathbf{Z}(:, \lambda_n)$, by using the function "unique(.)". Next, the signals, x, drawn from the union of the possible distinguishable constellations, $\mathcal{M}$, are used in an ML detector to determine the constellation symbol, $\hat{x}_{\lambda_n}$, and index, $\hat{I}_{\lambda_n}$ on the λ_nth codeword element. Then, the row indices, $\mathbf{r}$, corresponding to elements of $\mathbf{Z}(:, \lambda_n)$ that equal to $\hat{I}_{\lambda_n}$ are found by using the function "find(.)". To prevent illegal constellation indices, the index matrix is limited with only these rows, i.e, $\mathbf{Z} = \mathbf{Z}(\mathbf{r}, :)$. $\hat{x}_{\lambda_n}$ is assigned to the λ_nth element of decoding vector $\mathbf{d}$, i.e., $\mathbf{d}(\lambda_n) = \hat{x}_{\lambda_n}$. The algorithm continues until all symbols on N codeword elements are detected.

Note that ordering the channel gains and starting the decoding algorithm from the codeword elements related to the strongest channels improve the error performance

[4] Here, we omit the block superscripts for convenience since the LC-ML decoding can be performed independently for each block.

[5] For example, for the OFSPM scheme having $N = 3$ codeword elements, $\mathbf{Z} \in \mathbb{Z}^{13\times3}$ matrix consists the subscripts of subset identifiers, or in other words, the indices of disjoint constellations, shown in Table 3.1. Here, $\mathbb{Z}^{L_S \times N}$ stands for the set of integer numbers which is of the size $L_S \times N$.

considerably. Although we assume that the codeword elements fade independently, each codeword has a set partitioning structure. After deciding the codeword elements with the highest channel gains, we choose the remaining ones according to the previously decoded codeword elements and the SPM structure. This approach improves the reliability of the SPM symbol decoding as the decoding of the codeword elements related to the weakest channels are reinforced by the ones that have higher channel gains. It also reduces the error propagation resulting from the erroneous detection of the SPM codewords.

The number of operations performed by the LC-ML detectors of SPM and OSPM can be upper-bounded by $(NK - K + N + 2L_S - 1)\log_2 L_S + 8NKM$ operations, whereas it can be upper-bounded by $(N^2 + 2L_S - 1)\log_2 L_S + 8N^2M$ operations for the LC-ML detectors of FSPM and OFSPM. These complexities become relatively small compared to the complexity of the optimum ML detection when especially M is high. For example, when $N = 4$, $K = 2$ and $M = 8$, the proposed LC-ML detector performs at least 99.75 %, 99.87 %, 99.76 %, and 99.92 % less operations, respectively, for the SPM, OSPM, FSPM, and OFSPM schemes than the corresponding optimum ML detectors.

In this algorithm, the function "unique(.)" can be performed in $(l_n + N)\log_2 l_n$ operations where l_n is the number of rows of the index matrix $\mathbf{Z}$ in the nth step,[6] $n = 1, \ldots, N$, since an array of size l_n can be sorted in $l_n \log l_n$ operations and N distinct elements in an sorted array of size l_n can be found by performing $N\log_2 l_n$ binary search operations. In the worst-case, the constellation indices can take all possible values of the set $\{1, 2, \ldots, K\}$ for the SPM and OSPM schemes and the set $\{1, 2, \ldots, N\}$ for the FSPM and OFSPM schemes. Hence, $8KM$ real multiplications for the SPM and OSPM schemes and $8NM$ real multiplications for the FSPM and OFSPM schemes are performed during the calculations of squared Euclidean

[6]For example, in the first iteration, $l_1 = L_S$.

distances. Then, using the binary search, the function "find(.)" requires $\log_2 l_n$ operations in the nth iteration. Hence, in the worst-case, the proposed LC-ML detector performs $(\sum_{n=1}^{N}(K + l_n + 1)\log_2 l_n) + 8NKM$ operations for SPM and OSPM, and $(\sum_{n=1}^{N}(N + l_n + 1)\log_2 l_n) + 8N^2M$ operations for FSPM and OFSPM. Considering the facts that $l_1 = L_S$, $\sum_{n=2}^{N} l_n < L_S$, and $\sum_{n=2}^{N}\log_2 l_n \leq (N-1)\log_2 \frac{L_S}{N-1}$, the number of operations for the LC-ML detector of the SPM and OSPM schemes can be upper-bounded by $(NK - K + N + 2L_S - 1)\log_2 L_S + 8NKM$ and such a number can be upper-bounded by $(N^2 + 2L_S - 1)\log_2 L_S + 8N^2M$ for the FSPM and OFSPM methods.

3.4 Codebook Selection

As explained in Section 2.3, for a fading channel, the BER performance of a codebook at a high SNR is limited by the minimum product distance between codeword pairs that have the minimum Hamming distance. For the IM schemes, such codeword pairs are formed by the modulation symbols instead of the symbols carried by the indices [1]. More explicitly, the minimum Hamming distance is observed when the IM codewords $\mathbf{x}_m$ and $\mathbf{x}_{m'}$ have the same indices and different modulation symbols on one of their codeword elements. This is why the minimum Hamming distance for the pairs of the IM codewords is unity. However, such a distance is two for the index symbols [1]. Although the statement on the modulation symbols is valid for the SPM techniques, the statement on the index symbols is not necessarily valid for the SPM methods. This can be shown by comparing the OSPM codewords $\mathbf{x}_1$ and $\mathbf{x}_5$ in Table 3.1. Here, the only difference between the two codewords is observed in the third elements. Such a condition limits the diversity order of the SPM symbols to one if we include both codewords in the same codebook. However, one can also construct the codebook in a way that the unit distance codeword pairs are not included. This ensures the same minimum Hamming distance characteristics for the SPM codewords

as for the IM and MM-IM concepts. Hence, our aim in codebook selection for the SPM variants can be summarized as obtaining codewords that are at least as good as those of the IM methods. We are well aware of the fact that the optimum codebook selection is based on maximizing diversity order and product-distances; however, our selection algorithms do not consider product-distances due to the complexity of the codebook selection problem.

The codebook selection problem can be formulated as a *maximum clique problem*[7] or *k-clique problem* in a graph. Before formulating such problem, we begin with specifying some definitions and notations related to graph theory for clarity. Consider an undirected graph $G = (V, E)$, where $V = \left\{v_1, v_2, \ldots, v_{L_S}\right\}$ is the vertex set and $E \subseteq V \times V$ is the edge set of G. A graph is complete if all of its vertices are connected by an edge, i.e., we have $(v_l, v_{\hat{l}}) \in E$ for $\forall v_l, v_{\hat{l}} \in V$ and $v_l \neq v_{\hat{l}}$. A clique can be defined as a complete subgraph of a graph [101, 102]. The clique number or the size of the maximum clique is denoted by $w(G)$. Moreover, the symmetric $L_S \times L_S$ matrix $A_G = (a_{v_l v_{\hat{l}}})_{v_l, v_{\hat{l}} \in V}$, where $a_{v_l v_{\hat{l}}} = 1$ if $(v_l, v_{\hat{l}}) \in E$ and $a_{v_l v_{\hat{l}}} = 0$ if $(v_l, v_{\hat{l}}) \notin E$, stands for the adjacency matrix of the graph G.

Now, we formulate our codebook selection problem as maximum clique and k-clique problems as follows. Assume that a graph's vertices represent the SPM codewords, and the graph's edges represent the Hamming distance between such codeword pairs. To ensure similar Hamming distance properties to the IM benchmarks, we set a condition on the vertex connections such that two vertices are connected to each other if the Hamming distance between vertices (codewords) is greater or equal to two, i.e. $E = \left\{(v_l, v_{\hat{l}}) |\ v_l, v_{\hat{l}} \in V, \text{HamDist}(v_l, v_{\hat{l}}) \geq 2\right\}$ where the "HamDist" function is used to measure the Hamming distance between vertices. Here, our aim is to finding the maximum clique that includes the largest possible number of vertices.

On the other hand, if the problem is formulated as a *k-clique problem*, one should

[7]This problem is also equivalent to some other important graph problems such as the *maximum independent set problem* and *minimum vertex cover problem.*

consider the cliques that includes a specific number, k, of vertices. Considering the number of the set partitions, the clique size, k, can be determined. For example, we have $\breve{B}_N = 75$ ordered set partitions for $N = 4$ and we are interested in choosing 64 partitions out of 75 to send six data bits. Hence, we can formulate such a problem in the same way as the maximum clique problem. However, in this case, we are interested in the cliques with 64 vertices. One drawback of the k-clique problem is that we do not know whether the current graph has a clique of size k. However, we can proceed to check the cliques by decreasing the clique size in the absence of a k-clique in the given graph.

3.4.1 Brute-force Search for k-clique Problem

One straightforward way of selecting good codewords, in other words, finding a k-clique, is exhaustively searching among all possible subgraphs. In this regard, the brute-force search algorithm for the k-clique problem is given in Algorithm 2.

The algorithm starts by initializing the graph $G(V, E)$ where $V = \{v_1, v_2, \ldots, v_{L_S}\}$ is the vertex set representing the SPM codewords, and $E = \{(v_l, v_{\hat{l}}) |\ v_l, v_{\hat{l}} \in V,$ $\text{HamDist}(v_l, v_{\hat{l}}) \geq 2\}$ is the edge set of G. In the second step, we set $\kappa = 0$ where κ is used to adjust the clique size. Then, in the third step, we divide the graph constituted by the partitions into subgraphs, G_c, with the corresponding vertex set V_c and edge set E_c, in a way that each subgraph G_c includes $k = 2^{\lfloor \log_2 w_U(G) \rfloor - \kappa}$ vertices where $w_U(G) = N_{-1} + 1$ is an upper-bound on the size of the maximum clique $w(G)$, i.e., $w(G) \leq w_U(G) = N_{-1} + 1$ [103].[8] Here, we choose the number of vertices, k, in a subgraph as a power of two since the complexity of the algorithm would be high if we take all possible values of k into consideration. The choice of k can also be justified by the use of a fixed-length bit mapping scheme, since we wish to send the

[8] One may find many bounds on the clique number [102]. The safest strategy would be calculating all bounds and using the tightest. We found the upper-bound in [103] sufficiently tight for our problem.

Algorithm 2: Brute-force search for the k-clique problem

Step 1: Initialize a graph $G(V,E)$ with $V=\{v_1,v_2,\ldots,v_{L_S}\}$, the vertex set of the SPM codewords, and, $E=\{(v_l,v_{\hat{l}})|\ v_l,v_{\hat{l}}\in V,\mathrm{HamDist}(v_l,v_{\hat{l}})\geq 2\}$;
Step 2: $\kappa=0$;
Step 3: Initialize subgraphs, $G_c(V_c,E_c)$, $c=1,\ldots,C$, with $V_c=\{v_1^c,v_2^c,\ldots,v_k^c\}$ and $E_c=\{(v_l^c,v_{\hat{l}}^c)|\ v_l^c,v_{\hat{l}}^c\in V_c,\mathrm{HamDist}(v_l^c,v_{\hat{l}}^c)\geq 2\}$ where $C=\binom{L_S}{k}$, $k=2^{\lfloor\log_2 w_U(G)\rfloor-\kappa}$, $v_l^c,v_{\hat{l}}^c\in V_c$, $l,\hat{l}=\{1,2,\ldots,k\}$, and $v_l^c\neq v_{\hat{l}}^c$;
Step 4: Initialize adjacency matrix, A_{G_c}, of G_c and $\hat{V}\in\emptyset$ where $A_{G_c}=(a_{l,\hat{l}}^c)_{(v_l^c,v_{\hat{l}}^c)\in V_c}$;
Step 5: Finding the k-clique;
for $c=1$ **to** C **do**
  if $a_{l,\hat{l}}^c=1,\ \forall l,\hat{l}=\{1,2,\ldots,k\}$ *and* $l\neq\hat{l}$ **then**
    $\hat{V}=V_c$;
    break;
  end
  $c\leftarrow c+1$;
end
Step 6: Checking the presence of k-clique in G;
if $\hat{V}\in\emptyset$ **then**
  $\kappa\leftarrow\kappa+1$;
  Go to Step 3;
end
return $\hat{V}$;

same amount of data bits even if we consider all possible values of k. Moreover, N_{-1} denotes the number of eigenvalues of the adjacency matrix A_G that do not exceed -1. Then, for each subgraph, we calculate adjacency matrix, A_{G_c}, and initialize a selected graph as $\hat{V}\in\emptyset$. In the fifth step, we check each adjacency matrix to see whether the subgraph is a k-clique or not. If the subgraph is a k-clique, the algorithm returns the current subgraph as the selected graph. If no k-clique is encountered after checking all subgraphs, we increase κ by one and go to Step 3 to check whether there is a clique of smaller size. In that way, the brute-force search algorithm guarantees reaching the maximum number of bits carried by set partitions for the fixed-length binary encoding.

Note that subgraph initialization in Step 3 may require excessive memory when

Algorithm 3: Vertex exclusion for maximum clique problem

Step 1: Initialize a graph $G(V, E)$ with $V = \{v_1, v_2, \ldots, v_{L_S}\}$ and $E = \{(v_l, v_{\hat{l}}) |\ v_l, v_{\hat{l}} \in V, \text{HamDist}(v_l, v_{\hat{l}}) \geq 2\}$;
Step 2: Calculate the cardinality of the neighborhood, $|N(v_l)|$, for each node where $v_l \in V$;
Step 3: Finding a clique by excluding vertices;
while $|N(v_l)| \neq |V| - 1, \forall v_l \in V$ **do**
    $v_{min} = \arg\min_{v_l \in V} |N(v_l)|$;
    $G \leftarrow G - \{v_{min}\}$;
    $V \leftarrow V - \{v_{min}\}$;
    Calculate the neighborhood cardinality, $|N_c(v_l)|$, for each node of V where $\forall v_l \in V$;
    $|N(v_l)| \leftarrow |N_c(v_l)|$;
end
return V;

the number of vertices is high. To avoid such memory requirements, one can use a combinatorial number system to obtain vertex combinations from natural numbers. The combinatorial number system provides a bijective mapping between the natural numbers and k-combinations [104]. Hence, Steps 3 and 4 can be integrated into the for-loop in Step 5 and each subgraph can be constructed in each loop by using a combinatorial number system. Note also that the for-loop in Step 5 can be parallelized to speed up the algorithm.

3.4.2 Vertex Exclusion for the Maximum Clique Problem

Since the maximum clique problem is NP-complete, the time of execution of exact algorithms will increase exponentially with the number of vertices in the graph [102]. In this regard, the brute-force search algorithm exhibits high complexity when the number of vertices is high. To overcome this problem, we propose a more practical algorithm based on excluding the vertices that have the minimum number of connections. Although such an algorithm provides a sub-optimal solution for the codeword selection problem, it is much more efficient than a brute-force search in terms of complexity.

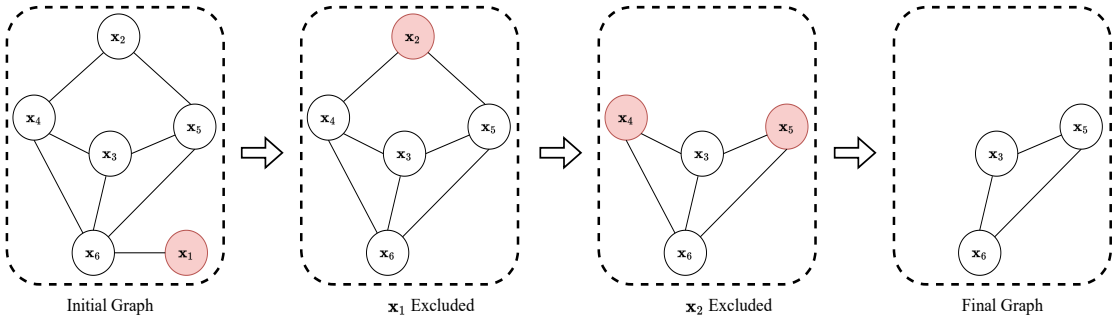


Figure 3.3. Vertex exclusion example for SPM when $N = 4$ and $K = 3$.

Our proposed algorithm is given in Algorithm 3. The algorithm starts by initializing the Hamming graph. Then, in the second step, we calculate the number of neighbors, $|N(v_l)|$, of each node $v_l \in V$ where $N(v_l) = \left\{v_{\hat{l}} \in V;\ a_{v_l v_{\hat{l}}} = 1\right\}$ is the neighborhood of v_l in G. In the third step, we remove the nodes that have the minimum number of neighbors until the remaining graph is a clique, i.e. $|N(v_l)| = |V| - 1, \forall v_l \in V$. In this step, $|N_c(v_l)|$ is used to update the cardinality of the neighborhood of v_l after removing the node that has the minimum number of neighbors. As a result, the algorithm removes the set partitions having the maximum number of unit Hamming distance pairs one-by-one until we have no set partition pairs with unit Hamming distances.

We give an illustrative example of Algorithm 3 in Fig. 3.3. We start forming an initial graph by a set of vertices and edges. We label the vertices by codewords, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_6$, of an SPM codebook with $N = 4$ and $K = 3$. Here, there exists an edge between two particular nodes if the Hamming distance between the corresponding SPM codewords is greater than or equal to two. After forming the initial graph, we proceed by removing the nodes which have the minimum number of edges until the remaining graph is a clique. We highlight such nodes in red color in the figure. In the first step, we cut the node labeled $\mathbf{x}_1$ out of the initial graph since such a node has the minimum number of edges. Then, for the same reason, we exclude the one labeled $\mathbf{x}_2$ from the resulting graph. In the next step, we can remove either $\mathbf{x}_4$ or

$\mathbf{x}_5$ as they have the minimum number of edges. In the final graph, we encounter a clique, meaning that the Hamming distance between the codewords $\mathbf{x}_3, \mathbf{x}_5$, and $\mathbf{x}_6$ is at least two.

3.4.3 Complexity Comparison

In this subsection, we compare the complexities of Algorithms 2 and 3 in terms of the algorithm run-times and the worst-case time-complexities.

The algorithm run-time results in units of seconds are shown in Tables 3.2 and 3.3. In the tables, we also depict the numbers of vertices achieved by each algorithm in the brackets along with the upper-bound results on the achievable number of vertices according to [103]. Moreover, results obtained by using a built-in function called "FindClique[G]" in Wolfram Language are provided as a benchmark. Such a function searches for the maximal set of vertices where the corresponding subgraph is a clique [105]. Finally, we include the number of index symbols achieved by the DM-IM and MM-IM methods in the rightmost columns of Tables 3.2 and 3.3, respectively, to show the effectiveness of the proposed schemes against these schemes.[9] Here, to reach the maximum number of index symbols, DM-IM modulates half of the codeword elements by one of the distinguishable constellations and the other half by the other constellation.

In Table 3.2, the complexity results are provided for the OSPM schemes having $N = 4, 6$ and, 8 codeword elements and $K = 2$ distinguishable constellations. In this case, the OSPM systems can produce 14, 62, and 254 vertices (SPM codewords) for $N = 4, 6$, and, 8, respectively. On the other hand, in Table 3.3, we provide complexity results for the OFSPM schemes having $N = 3, 4, 5$, and 6. In that case, these schemes are capable of producing 13, 75, 541, and 4683 vertices (SPM codewords) for $N =$

[9]Here, we include the number of index symbols produced by the DM-IM and MM-IM techniques to compare the number of symbols related to combinations/permutations for these schemes and set partitions for the proposed schemes. Hence, we do not employ the codebook selection algorithms to get the same number of index symbols as those of these schemes.

Table 3.2. The run-time complexities (in seconds) of the vertex selection algorithms along with the numbers of achieved vertices (in brackets) for the OSPM concept when $N \in \{4, 6, 8\}$ and $K = 2$.

	Algorithm 2	Algorithm 3	FindClique[G]	Upper-bound in [103]	DM-IM
$N = 4$	0.4343 (8)	0.0033 (8)	0.05 (8)	(9)	(6)
$N = 6$	-	0.0176 (32)	0.06 (32)	(32)	(20)
$N = 8$	-	0.0761 (128)	0.20 (128)	(129)	(70)

$3, 4, 5$, and 6, respectively. As can be seen from Tables 3.2 and 3.3, Algorithm 2 is not able to provide a practical solution for most of the cases.[10] However, Algorithm 3 provides efficient solutions for all of the cases by outperforming both Algorithm 2 and the "FindClique[G]" function in terms of algorithm run-time. The effectiveness of Algorithm 3 becomes more evident when the number of codeword elements, and therefore the number of vertices, increases. Furthermore, the clique number achieved by Algorithm 3 is consistent with the clique numbers achieved by the "FindClique[G]" function and obtained by the upper-bound in [103]. Also, the OSPM and OFSPM schemes employing Algorithm 3 attain more symbols associated with set partitions than index symbols of the DM-IM and MM-IM schemes, respectively.

The worst-case time-complexities of Algorithms 2 and 3, on the other hand, can be calculated as follows. To perform Steps 1, 2, 4 and 6 of Algorithm 2, the required time is almost constant and small compared to the time required to perform Steps 3 and 5. The upper-bound $\omega_U(G)$ in Step 3 can be executed in $O(L_S^3)$ operations [103]. In the worst case, it can be shown that the dominant term in the time-complexity becomes order $O((2e)^{L_S/2})$. Moreover, Algorithm 3 makes a total of $\frac{L_S(L_S+1)}{2} + \binom{L_S+1}{3}$ comparisons in the worst-case. Hence, Algorithm 3 can be executed in $O(L_S^3)$ operations which is much lower than the complexity of Algorithm 2.

[10] Here, "-" denotes that the algorithm is not able to return a solution within a reasonable amount of time.

Table 3.3. The run-time complexities (in seconds) of the vertex selection algorithms along with the numbers of achieved vertices (in brackets) for the OFSPM concept when $N \in \{3, 4, 5, 6\}$.

	Algorithm 2	Algorithm 3	FindClique[G]	Upper-bound in [103]	MM-IM
$N = 3$	0.024 (4)	0.0022 (7)	0.03 (7)	(7)	(6)
$N = 4$	-	0.0088 (32)	0.06 (32)	(33)	(24)
$N = 5$	-	0.4929 (181)	-	(225)	(120)
$N = 6$	-	477.0625 (1321)	-	(1876)	(720)

3.5 Performance Analysis

In this section, we analyze the data rate and the BER of the proposed SPM schemes.

3.5.1 Data Rate

We analyze the data rate of the proposed schemes in terms of the number of bits corresponding to the SPM codewords normalized by the number of channels used to convey each codeword. We also provide some useful expressions for the number of partitions obtained by the SPM schemes along with comparisons regarding the IM benchmarks. We assume $f_2 = N \log_2 M$ for each SPM variant.

3.5.1.1 SPM

As discussed in Section 3.2, the number of set partitions produced by SPM is given by the Stirling number of the second kind $\left\{ {N \atop K} \right\}$. The achievable data rate per channel for an SPM scheme with N channels and K subset identifiers in each block is

$$\eta_{SPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 \left\{ {N \atop K} \right\} \rfloor + N \log_2 M}{N} \tag{3.3}$$

where $\lfloor . \rfloor$ is the floor operation.

The Stirling number of the second kind, $\left\{ {N \atop K} \right\}$, can be written as [106]

$$\left\{ {N \atop K} \right\} = \frac{1}{K!} \sum_{j=0}^{K} (-1)^j \binom{K}{j} (K - j)^N. \tag{3.4}$$

It is also straightforward to show that the following recurrence holds:

$$\left\{ {N \atop K} \right\} = K \left\{ {N-1 \atop K} \right\} + \left\{ {N-1 \atop K-1} \right\}. \tag{3.5}$$

Remark. Consider an SPM block having two distinguishable constellations, i.e., $K = 2$, on N channels and a DM-IM block having two distinguishable constellations on N channels in which $N - d$ of the channels are modulated by one of two different constellations and the remaining d of them are modulated by the other constellation.[11] Except for the case where $N = 2$ and $d = 1$, the number of set partitions obtained by the SPM encoder is equal to or greater than the number of the element combinations obtained by the DM-IM encoder, i.e., $\left\{ {N \atop 2} \right\} \geq \binom{N}{d}$ for $N \geq 2$. The equality holds for $N = 2$ and $d = 2$. This means that the achievable data rate for SPM is equal to or greater than that of DM-IM when the channels of both schemes convey symbols that have the same modulation order. One may check that $\left\{ {N \atop 2} \right\} = 2^{N-1} - 1$. Moreover, from the recurrence relation of the binomial coefficient, we have $\binom{N}{d} = \binom{N-1}{d} + \binom{N-1}{d-1}$. If we compare $\binom{N}{d}$ with $\left\{ {N \atop 2} \right\}$, we see that $\left\{ {N \atop 2} \right\} \geq \binom{N}{d}$ for $N \geq 2$ except for the case where $N = 2$ and $d = 1$.

For a fixed and relatively small value of K, the asymptotic value of the Stirling number of the second kind as $N \to \infty$ can be written as $\left\{ {N \atop K} \right\} \sim \frac{K^N}{K!}$. Hence, the asymptotic value of the achievable data rate per channel for an SPM system, when $N \to \infty$, can be written as

$$\eta_{SPM} \sim \frac{N \log_2 K - \log_2 K! + N \log_2 M}{N} \sim \log_2(KM). \tag{3.6}$$

Remark. Consider a special MM-IM concept as in [83] having N channels and K distinguishable constellations along with M-PSK/QAM/PAM symbols carried on each

[11] It is fair to compare these two schemes since both of them have two distinguishable constellations.

channel. To compare the achievable rate of such an MM-IM scheme, we assume that each N/K codeword elements employ the same disjoint constellation. Hence, for this scenario, we have $\frac{N!}{(N/K)!^K}$ combinations related to the disjoint constellations. Applying the Stirling approximation,[12] the achievable data rate per channel for this MM-IM scheme as $N \to \infty$ can be written as $\eta_{MM-IM} \sim \log_2(KM)$. Hence, the proposed SPM scheme is capable of providing the same asymptotic achievable rate as this special MM-IM scheme.

When N is large, the value of K that maximizes $\left\{{N \atop K}\right\}$ satisfies $K_N \sim \frac{N}{\ln N}$ [107]. More precisely, the following relation holds for sufficiently large N [108]

$$K_N \in \left\{\lfloor e^{W(N)} - 1 \rfloor, \lceil e^{W(N)} - 1 \rceil\right\} \tag{3.7}$$

where $\lfloor . \rfloor$ and $\lceil . \rceil$ are floor and ceiling operations, respectively. $W(N)$ is the Lambert W function satisfying $W(N)\exp(W(N)) = N$ [109]. There is no exception to the relation in (3.7) for $1 \leq N \leq 1200$ [108]. More importantly, the maximum value of the Stirling number of the second kind adheres to the relation $\ln \left\{{N \atop K_N}\right\} \sim N \ln N - N \ln \ln N - N$ [107]. Hence, the asymptotic maximum achievable rate of the SPM scheme satisfies

$$\eta_{SPM}^{max} \sim \frac{N \log_2(N/e \ln N) + N \log_2 M}{N} = \log_2(N/\ln N) + \log_2(M) - \log_2 e. \tag{3.8}$$

This asymptotic result shows that a substantially improved data rate is attainable when we use K_N distinguishable constellations in the SPM technique.

[12] $\ln N! \sim N \ln N - N$ as $N \to \infty$.

3.5.1.2 OSPM

The achievable data rate per channel for an OSPM scheme having N codeword elements and K subset identifiers can be written as

$$\eta_{OSPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 K! \left\{ {N \atop K} \right\} \rfloor + N \log_2 M}{N}. \tag{3.9}$$

Remark. It is straightforward to show that $2! \left\{ {N \atop 2} \right\} \geq \binom{N}{d}$. However, it is important to note that an OSPM codebook, which incorporates partitions with $K = 2$ subsets and their ordered counterparts, subsumes a DM-IM codebook. This can easily be proved by considering set partitions along with permutations when $K = 2$. Hence, it can be concluded that OSPM is a more general scheme, which encompasses symbols related to the element combinations generated by an DM-IM encoder. It is also important to note that an OSPM encoder produces the same index patterns as an MM-IM encoder for $K = N$. Hence, MM-IM is a special case of OSPM when $K = N$. Moreover, the partitions of an N-element set into two subsets would result in the same index symbols as GDM-IM when the set, $\mathcal{K}$, containing possible numbers of codeword elements having one of a number of distinguishable constellations is given by $\mathcal{K} = \{1, 2, \ldots, N-1\}$ for GDM-IM. *Despite these similarities, it is important to recognize that OSPM is inherently different due to the use of set partitions to encode information instead of index patterns or permutations.*

Assuming a fixed K and using the asymptotic representation for the Stirling number of the second kind, we can write the following achievable data rate expression as $N \to \infty$:

$$\eta_{OSPM} \sim \frac{N \log_2 K + N \log_2 M}{N} \sim \log_2(KM). \tag{3.10}$$

Hence, OSPM achieves the same asymptotic data rate as that of SPM when $K << N$.

On the other hand, it is known that, the value of K that maximizes $K!\left\{{N \atop K}\right\}$ satisfies $\breve{K}_N \sim \frac{N}{2\ln 2}$ as $N \to \infty$ [110]. Moreover, we have the asymptotic relation $\breve{K}_N!\left\{{N \atop \breve{K}_N}\right\} \sim N!/2(\ln 2)^{N+1}$ [110]. Hence, the asymptotic maximum achievable rate of the OSPM method satisfies

$$\eta_{OSPM}^{max} \sim \frac{\log_2(N!/2(\ln 2)^{N+1}) + N\log_2 M}{N} \sim \log_2(N) + \log_2(M) - \log_2(e\ln 2). \quad (3.11)$$

where the second asymptotic relation follows from the Stirling's approximation.

We can conclude that the OSPM concept is capable of achieving asymptotically higher achievable rate than the MM-IM scheme while the number of exploited distinguishable constellations is less than N.

3.5.1.3 FSPM

The achievable data rate for an FSPM scheme having N channels can be written as

$$\eta_{FSPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 B_N \rfloor + N\log_2 M}{N}. \quad (3.12)$$

Using the asymptotic expression for the Bell number given in [111], the associated data rate can be written as

$$\eta_{FSPM} \sim \log_2(N/\ln N) + \log_2(M) - \log_2 e. \quad (3.13)$$

3.5.1.4 OFSPM

The achievable data rate per channel for an OFSPM system having N channels can be written as

$$\eta_{OFSPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 \breve{B}_N \rfloor + N\log_2 M}{N}. \quad (3.14)$$

Remark. It is straightforward to show that $\breve{B}_N > N!$ for $N \geq 2$, since the ordered set partitions include all permutations of N-element set partitions and the number of the permutations of N-element partitions is equal to $N!$. In other words, for an N-element set $\mathcal{U} \coloneqq \{u_1, u_2, \ldots, u_N\}$ with $K = N$, $\mathcal{S} = \{\{u_1\}, \{u_2\}, \ldots, \{u_N\}\}$ is a valid set partition, and ordering the subsets of this set would result in $N!$ different partitions. Hence, the resulting OFSPM codebook contains the codeword $\mathbf{x}_l = \{\mu_1, \mu_2, \ldots, \mu_N\}$ along with the codewords representing all permutations of the elements of $\mathbf{x}_l$. Note that these codewords form the MM-IM codebook, and OFSPM is a more general scheme compared to MM-IM since it covers all codewords formed by an MM-IM encoder.

Using the asymptotic relation for the ordered Bell numbers given in [112], the asymptotic achievable data rate of the OFSPM system as $N \to \infty$ can be written as

$$\eta_{OFSPM} \sim \frac{\log_2(N!/2(\ln 2)^{N+1}) + N \log_2 M}{N} \sim \log_2(N) + \log_2(M) - \log_2(e \ln 2). \quad (3.15)$$

As can be observed from (3.8), (3.11), (3.13), and (3.15), the maximum achievable data rate of each SPM variant exhibits the same asymptotic behavior as its full SPM counterpart while exploiting fewer distinguishable constellations. Hence, SPM and OSPM are capable of achieving the same asymtotic date rate as FSPM and OFSPM, respectively, by utilizing fewer disjoint constellations.

3.5.2 Bit Error Rate

Let $P(\mathbf{X}_i \to \mathbf{X}_j)$ denote the PEP associated with the erroneous detection of $\mathbf{X}_i$ as $\mathbf{X}_j$ where $\mathbf{X}_i = \text{diag}(\mathbf{x}_i)$ and $\mathbf{X}_j = \text{diag}(\mathbf{x}_j)$. From (3.2), the PEP conditioned on the

channel coefficients is given by[13]

$$P(\mathbf{X}_i \rightarrow \mathbf{X}_j|\mathbf{h}) = Q\left(\sqrt{\frac{E_S||(\mathbf{X}_i - \mathbf{X}_j)\mathbf{h}||^2}{2N_0}}\right). \tag{3.16}$$

By using the identity $Q(r) \approx \frac{1}{12}e^{-r^2/2} + \frac{1}{4}e^{-2r^2/3}$ and averaging over $\mathbf{h}$, an approximate unconditional PEP expression can be obtained [2]:

$$P(\mathbf{X}_i \rightarrow \mathbf{X}_j) = \mathrm{E}_{\mathbf{h}}\left[P(\mathbf{X}_i \rightarrow \mathbf{X}_j|\mathbf{h})\right] \approx \frac{1/12}{\det(\mathbf{I}_N + \frac{E_S}{4N_0}\mathbf{C}\mathbf{Z}_{ij})} + \frac{1/4}{\det(\mathbf{I}_N + \frac{E_S}{3N_0}\mathbf{C}\mathbf{Z}_{ij})} \tag{3.17}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix, $\mathbf{C} = \mathrm{E}_{\mathbf{h}}[\mathbf{h}\mathbf{h}^H]$, and $\mathbf{Z}_{ij} = (\mathbf{X}_i - \mathbf{X}_j)^H(\mathbf{X}_i - \mathbf{X}_j)$.

An upper-bound on the average BER is given by the well-known union bound

$$P_b \leq \frac{1}{f2^f}\sum_{i=1}^{2^f}\sum_{j=1}^{2^f} P(\mathbf{X}_i \rightarrow \mathbf{X}_j)D(\mathbf{X}_i \rightarrow \mathbf{X}_j) \tag{3.18}$$

where $D(\mathbf{X}_i \rightarrow \mathbf{X}_j)$ is the number of bits in error for the corresponding pairwise error event. Note that the upper-bound expression given in (3.18) is valid for all SPM schemes.

Assuming $\mathbf{C} = \mathrm{E}_{\mathbf{h}}[\mathbf{h}\mathbf{h}^H] \approx \mathbf{I}_N$ and considering the fact that $\mathbf{Z}_{ij}$ is a diagonal matrix, one can rewrite (3.17) as

$$P(\mathbf{X}_i \rightarrow \mathbf{X}_j) \approx \frac{1/12}{\prod_{n=1}^{N}(1 + \frac{E_S}{4N_0}\delta_n)} + \frac{1/4}{\prod_{n=1}^{N}(1 + \frac{E_S}{3N_0}\delta_n)} \tag{3.19}$$

where δ_n is the nth diagonal element of $\mathbf{Z}_{ij}$. At high SNR, one can neglect the one in

[13]Here, we drop superscripts related to the block indices as the BER analysis is valid for all blocks.

the denominator of (3.19) and write the following approximation

$$P(\mathbf{X}_i \to \mathbf{X}_j) \approx \frac{1/12}{\prod_{n\in\Gamma} \frac{E_S}{4N_0}\delta_n} + \frac{1/4}{\prod_{n\in\Gamma} \frac{E_S}{3N_0}\delta_n} \tag{3.20}$$

where Γ is the set of nonzero diagonal elements of $\mathbf{Z}_{ij}$. Finally, an asymptotic expression for the BER of the SPM variants can be obtained by substituting (3.20) into (3.18).

As explained in the previous sections, the SPM codebooks can be designed in a way that the minimum Hamming distance between the symbols related to the set partitions is equal to two. Hence, assuming the SPM symbols satisfy this property, the average BER expression will be dominated by the modulation symbols at high SNRs. Therefore, the diversity order of the BER curves is limited to one.

3.6 Numerical Results

In this section, we provide numerical data rate and BER results for the proposed schemes. In figures, "SPM (N, K, M)" and "OSPM (N, K, M)" stand for the SPM and OSPM methods having N channels and K distinguishable constellations along with M-PSK symbols conveyed by each channel, whereas "FSPM (N, M)" and "OFSPM (N, M)" stand for the variants of SPM employing all set partitions and having N codeword elements with M-PSK symbols. "OSPM (N, K, M), Alg. 3" and "OFSPM (N, M), Alg. 3" form their codewords in the same way as OSPM, (N, K, M) and OFSPM, (N, M), respectively, and then, they employ Algorithm 3 to cut some of their codewords.[14] Moreover, "IM (N, K_a, M)" stands for the conventional IM scheme having K_a activated codeword element out of N and employing M-PSK symbols on the

[14]Here, we employ a slightly different version of Algorithm 3 to the overall codebook of the OSPM and OFSPM schemes. By defining the Hamming distance as the number of positions at which the corresponding symbols of the SPM schemes are different, we use Algorithm 3 to cut the OSPM and OFSPM codewords having the maximum number of unit Hamming distance pairs one-by-one until we have the same data rate as those of the benchmarks.

Table 3.4. Descriptions of the modulation concepts considered in this section.

Term	Description
SPM	Set partition modulation
OSPM	Ordered set partition modulation
FSPM	Full set partition modulation
OFSPM	Ordered full set partition modulation
DM-IM	Dual-mode index modulation
GDM-IM	Generalized dual-mode index modulation
MM-IM	Multi-mode index modulation
BPSK	Binary phase shift keying
8-PSK	8-ary phase shift keying

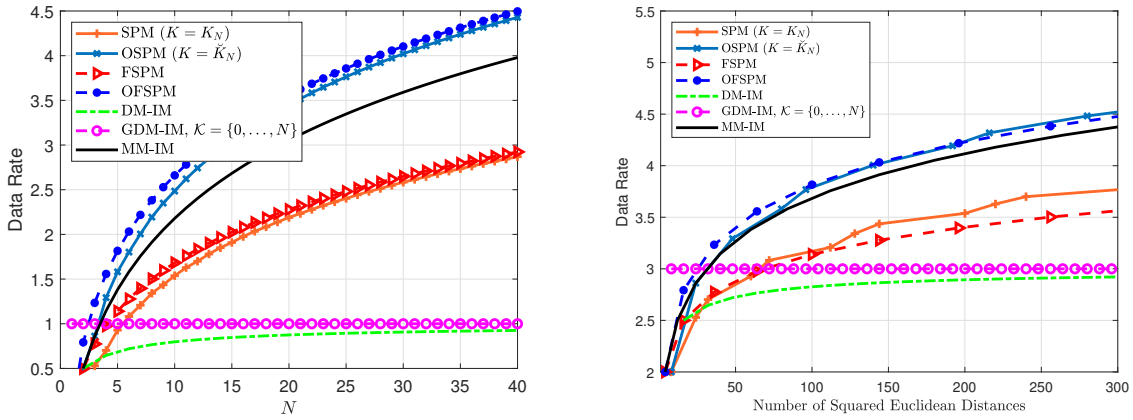
activated elements. Finally, "DM-IM (N, M)" signifies a dual-mode scheme having N codeword elements along with two distinguishable M-PSK constellations, and "MM-IM (N, M)" represents a multi-mode scheme having N codeword elements along with N distinguishable M-PSK constellations. In Table 3.4, we provide a list of definitions corresponding to the acronyms of the modulation methods to make it easier for the reader to follow the discussion in this section.

In our simulations, we assume all schemes operate over a Rayleigh fading channel, whose elements are i.i.d. RVs. The ML detection is applied under the assumption that channel estimation is perfect. The use of the Rayleigh model in the simulations is realistic when considering environments with a large number of scatterers [28].

3.6.1 Data Rate

In Fig. 3.4, we compare the data rates of the proposed SPM techniques with the DM-IM, GDM-IM, and MM-IM techniques. The data rate results in terms of the number of index bits per channel are given as a function of N in Fig. 3.4a and a function of detection complexity in Fig. 3.4b.[15] Since all the schemes considered in this figure activate all codeword elements, and the modulation order of the carried

[15]Here, by index bits, we imply the bits that are mapped to the symbols related to the set partitions and combinations/permuations in the SPM and IM systems, respectively.



(a) (b)

Figure 3.4. Data rate comparison of the SPM variants with DM-IM, GDM-IM, and MM-IM.

symbols on each element can be chosen to be the same, we ignore the modulation bits transmitted per channel in Fig. 3.4a. Also, we do not restrict the codebook sizes to a power of two, since all the codewords in a codebook can be utilized by a binary coding technique such as Huffman coding regardless of the number of codewords [80, 90]. To reach the maximum number of index bits, DM-IM modulates half of the codeword elements by one of the distinguishable constellations and the other half by the other constellation. Moreover, we assume $\mathcal{K} = \{0, 1, \ldots, N\}$ for GDM-IM. To reach the maximum numbers of set partitions for SPM and OSPM, we substitute $K_N \in \{\lfloor e^{W(N)} - 1 \rfloor, \lceil e^{W(N)} - 1 \rceil\}$ and $\breve{K}_N = \frac{N}{2 \ln 2}$, respectively, into K. The data rate results in Fig. 3.4a verify the remarks in the previous section and indicate that all SPM variants outperform DM-IM and GDM-IM for almost all values of N. Moreover, although MM-IM considerably outperforms SPM and FSPM, it is outperformed by OSPM and OFSPM for almost all values of N. On the other hand, in Fig. 3.4b, we calculate the detection complexity as the number of squared Euclidean distance calculations performed by the low-complexity LLR or LC-ML detectors of the considered schemes. In this regard, the LLR detector of MM-IM performs $(N^2 + N)M/2$ calculations, whereas the LLR detectors of DM-IM and GDM-IM perform $2NM$ calculations for

a codeword length N. Moreover, in the worst-case, the LC-ML detectors[16] of SPM and OSPM, respectively, perform NK_NM and $N\breve{K}_NM$ calculations, whereas the LC-ML detectors of FSPM and OFSPM perform N^2M calculations for a codebook with N-dimensional codewords. Here, we assume that $M = 4$ for all schemes. As seen from Fig. 3.4b, OSPM and OFSPM exhibit superior performance compared to other schemes. Moreover, the SPM mechanism outperforms the FSPM, DM-IM, and GDM-IM methods when the complexity increases.

As discussed in the previous chapter, to gain more insight into the achievable rate of the proposed schemes, one can substitute a finite input symbol set for such schemes into the standard mutual information expression. Thus, assuming equally likely codewords and applying the same approach as discussed in [76, 113], the achievable rate of the proposed schemes can be computed by evaluating (2.49).

We compare the achievable rates of the SPM variants with those of the MM-IM, DM-IM, IM, and conventional modulation (M-PSK) schemes in Figs. 3.5 and 3.6. Here, the rate results are in terms of bits per channel use (bpcu). The achievable rate curves provided in these figures are obtained by using (2.49). The results in Fig. 3.5 are given for SPM $(4, 2, 2)$, OSPM $(4, 2, 2)$, MM-IM $(2, 2)$, DM-IM $(4, 2)$, IM $(4, 2, 4)$, and BPSK systems. Here, SPM $(4, 2, 2)$ and OSPM $(4, 2, 2)$ encoders produce $\left\{{4 \atop 2}\right\} = 7$ and $2!\left\{{4 \atop 2}\right\} = 14$ codewords, and then $2^{\lfloor \log_2 7 \rfloor} = 4$ and $2^{\lfloor \log_2 14 \rfloor} = 8$ set partitions are selected by Algorithm 2. As can be observed from the figure, SPM $(4, 2, 2)$, MM-IM $(2, 2)$, DM-IM $(4, 2)$, and IM $(4, 2, 4)$ exhibit almost the same achievable rate performance for all SNR values. On the other hand, OSPM $(4, 2, 2)$ outperforms all other benchmarks at high SNR. Moreover, to achieve 1.5 bpcu, the SNR requirement for OSPM $(4, 2, 2)$ is approximately 12 dB lower than for the IM schemes.

[16]It is important to note that the proposed LC-ML detector does not require a look-up table when the codebook selection algorithms are not employed. In that case, it is enough to check that the detected distinguishable constellations form an admissible index symbol that can be mapped to a set partition. For this case, the detection complexity of the LC-ML detector is dominated by the number of squared Euclidean distance calculations.

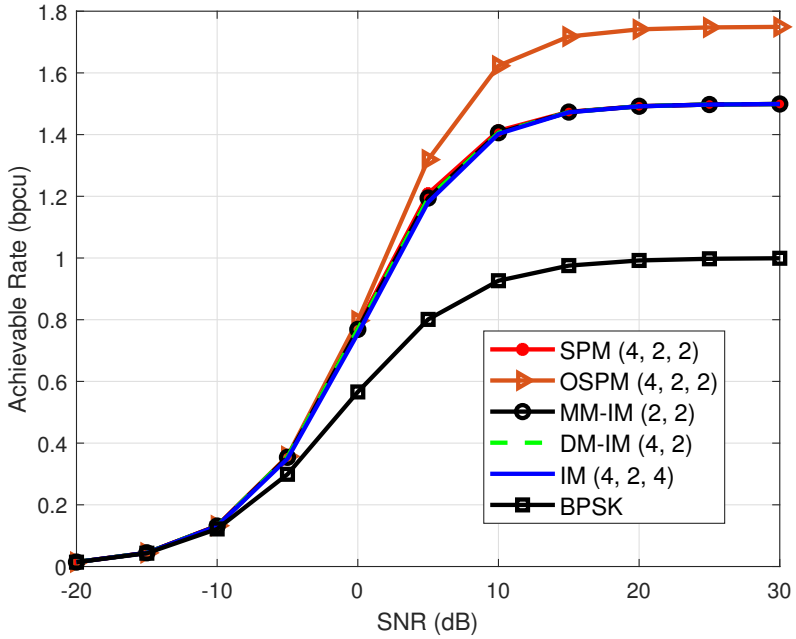


Figure 3.5. Achievable rate comparison of SPM $(4, 2, 2)$ and OSPM $(4, 2, 2)$ with MM-IM $(2, 2)$, DM-IM $(4, 2)$, IM $(4, 2, 4)$, and BSPK.

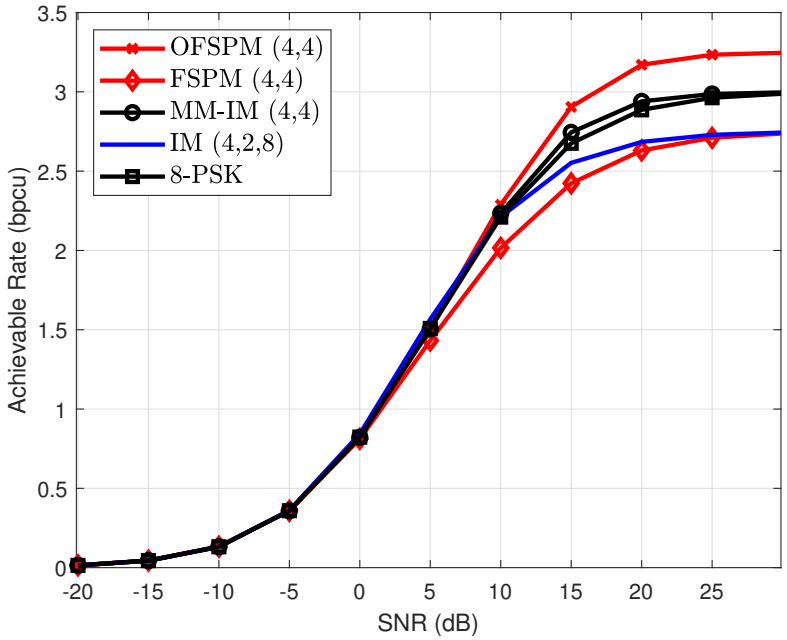


Figure 3.6. Achievable rate comparison of OFSPM $(4, 4)$ and FSPM $(4, 4)$ with MM-IM $(4, 4)$, IM $(4, 2, 8)$, and 8-PSK.

In Fig. 3.6, we compare the achievable rates of OFSPM $(4, 4)$ and FSPM $(4, 4)$ with MM-IM $(4, 4)$, IM $(4, 2, 8)$, and 8-PSK. Here, FSPM $(4, 4)$ and OFSPM $(4, 4)$

encoders produce $B_N = 15$ and $\breve{B}_N = 75$ codewords, respectively. 32 codewords are then selected according to Algorithm 3 for OFSPM $(4, 4)$ in a way that the rank of the difference matrices for the codeword pairs in the selected codebook is at least two. For FSPM $(4, 4)$, Algorithms 2 and 3 cannot return $2^{\lfloor \log_2 15 \rfloor} = 8$ codewords; they return only four and five codewords, respectively. To fully exploit the codewords generated by this scheme and to achieve a higher data rate at high SNR, we use three other codewords in the FSPM (4, 4) scheme in addition to the codewords generated by Algorithm 3. This limits the minimum Hamming distance of the codeword pairs to one. This explains why the FSPM (4, 4) scheme is outperformed by all other schemes. However, Fig. 3.6 shows that OFSPM $(4, 4)$ is capable of providing a better achievable rate than all other schemes at high SNR. Moreover, OFSPM $(4, 4)$ is, theoretically, able to achieve 3 bpcu with an SNR requirement that is almost 13 dB lower than that of MM-IM $(4, 4)$.

3.6.2 Bit Error Rate

In this subsection, we compare the proposed SPM methods with the conventional modulation and IM methods in terms of BER performance. We also investigate the effectiveness of the proposed LC-ML detector compared to the optimum ML detector.

In Fig. 3.7, we compare the BER performance of SPM $(4, 2, 2)$, OSPM $(4, 2, 2)$, and OSPM $(4, 2, 2)$, Alg. 3 with MM-IM $(2, 2)$, DM-IM $(4, 2)$, IM $(4, 2, 4)$, and classic BSPK. Except for OSPM $(4, 2, 2)$ and BSPK, all schemes exhibit the same SE which is 1.5 bpcu. SEs for OSPM $(4, 2, 2)$ and BSPK are 1.75 bpcu and 1 bpcu, respectively. Here, SPM $(4, 2, 2)$ and OSPM $(4, 2, 2)$ encoders produce $\left\{ {4 \atop 2} \right\} = 7$ and $2!\left\{ {4 \atop 2} \right\} = 14$ SPM codewords, and then $2^{\lfloor \log_2 7 \rfloor} = 4$ and $2^{\lfloor \log_2 14 \rfloor} = 8$ set partitions are selected by Algorithm 2. To make a fair comparison in terms of data rate, OSPM $(4, 2, 2)$, Alg. 3 employs Algorithm 3 to cut the codewords of OSPM $(4, 2, 2)$ until it has the same SE as those of SPM $(4, 2, 2)$, MM-IM $(2, 2)$, DM-IM $(4, 2)$, and

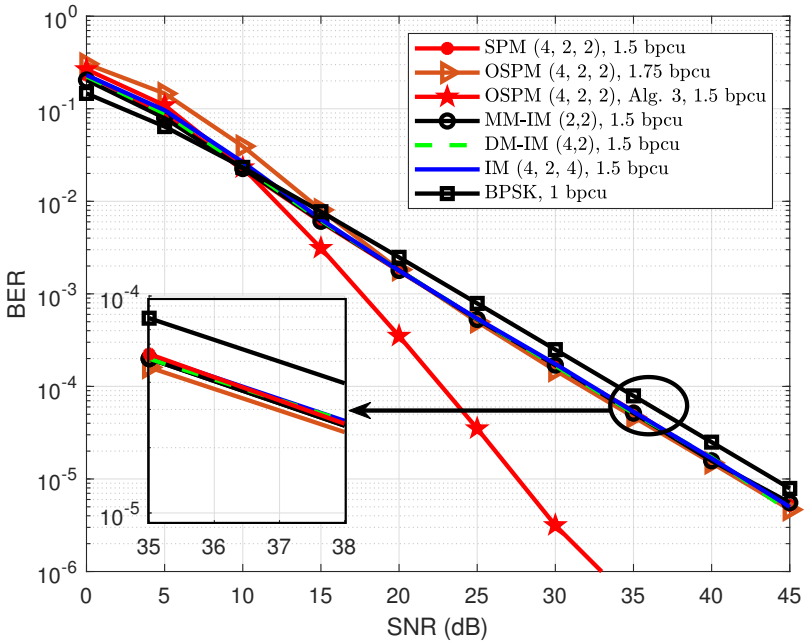


Figure 3.7. BER comparison of SPM $(4, 2, 2)$, OSPM $(4, 2, 2)$, and OSPM $(4, 2, 2)$, Alg. 3 with MM-IM $(2, 2)$, DM-IM $(4, 2)$, IM $(4, 2, 4)$, and BPSK.

IM $(4, 2, 4)$. As seen from the figure, SPM $(4, 2, 2)$ exhibits almost the same BER performance as IM $(4, 2, 4)$, DM-IM $(4, 2)$, and MM-IM $(2, 2)$, and it outperforms BSPK at medium-to-high SNR. More importantly, although OSPM $(4, 2, 2)$ has higher SE than other schemes, it provides marginally better BER performance compared to other techniques at high SNR. The reason behind this is the codewords related to set partitions in the OSPM scheme have the capability of preserving the same minimum Hamming distance properties as the codewords related to the permutations in the MM-IM scheme or the combinations in the DM-IM scheme, and the number of such codewords in the OSPM scheme is larger than those in the DM-IM and MM-IM schemes. This is desirable at high SNR since these codewords are capable of introducing diversity order of two unlike the codewords related to conventional modulation, which limit the diversity order to one. Therefore, they mitigate the effect of the codewords related to conventional modulation on the BER performance. OSPM $(4, 2, 2)$, Alg. 3, on the other hand, considerably outperforms the other schemes by introducing an additional diversity gain, as the codebook of OSPM $(4, 2, 2)$ is

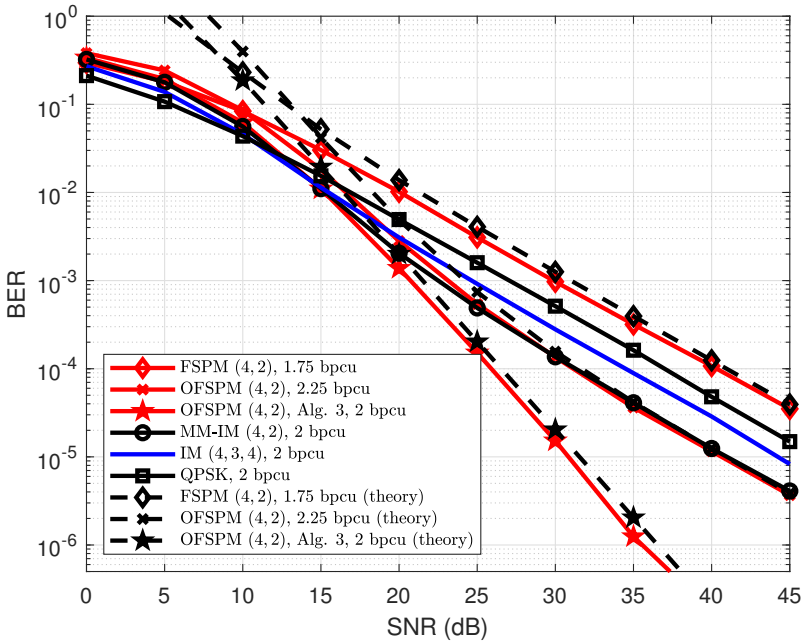


Figure 3.8. BER comparison of FSPM $(4, 2)$, OFSPM $(4, 2)$, and OFSPM $(4, 2)$, Alg. 3 with MM-IM $(4, 2)$, IM $(4, 3, 4)$, and QPSK.

culled enough by Algorithm 3 that the minimum Hamming distance for the remaining codebook becomes two. This is only possible because the combinatorial structure that partitioning offers produces more index symbols than the benchmark schemes.

In Fig. 3.8, FSPM $(4, 2)$, OFSPM $(4, 2)$, and OFSPM $(4, 2)$, Alg. 3 are compared with MM-IM $(4, 2)$, IM $(4, 3, 4)$, and QPSK. Except for FSPM $(4, 2)$ and OFSPM $(4, 2)$, all schemes have the same SE of 2 bpcu. SEs for FSPM $(4, 2)$ and OFSPM $(4, 2)$ are 1.75 bpcu and 2.25 bpcu, respectively. Here, FSPM $(4, 2)$ and OFSPM $(4, 2)$ encoders produce $B_N = 15$ and $\breve{B}_N = 75$ codewords. Then, 32 codewords are selected according to Algorithm 3 for OFSPM $(4, 2)$. As discussed in the previous section, Algorithms 2 and 3 are not capable of providing $2^{\lfloor \log_2 15 \rfloor} = 8$ codewords for FSPM $(4, 2)$. However, we use eight codewords of FSPM $(4, 2)$, including the five FSPM codewords provided by Algorithm 3, to achieve a higher data rate. OFSPM $(4, 2)$, Alg. 3 employs Algorithm 3 to remove the codewords of OFSPM $(4, 2)$ until it has the same SE as those of MM-IM $(4, 2)$, IM $(4, 3, 4)$, and QPSK. We also provide results on the theoretical upper-bound for the SPM schemes. As observed from

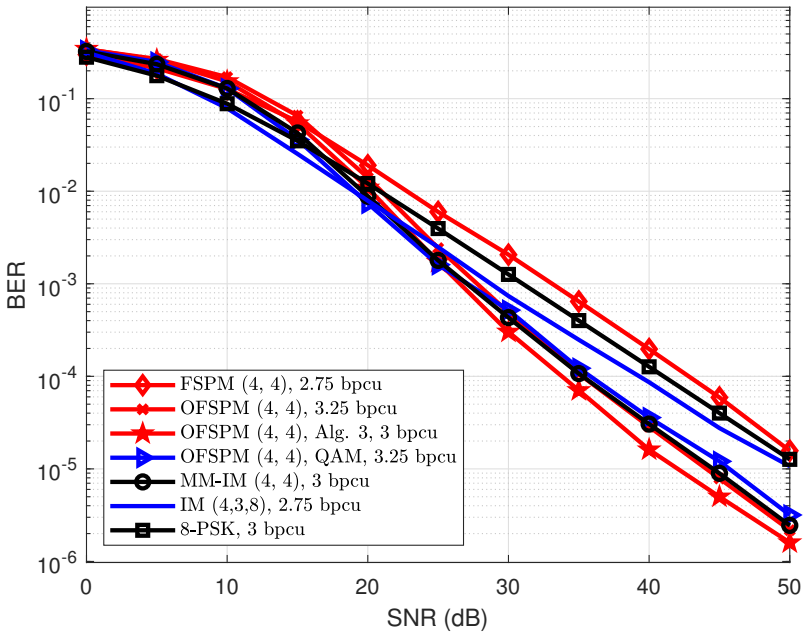


Figure 3.9. BER comparison of FSPM $(4, 4)$, OFSPM $(4, 4)$, and OFSPM $(4, 4)$, Alg. 3 with MM-IM $(4, 4)$, IM $(4, 3, 8)$, and 8-PSK.

the figure, upper-bound curves are consistent with computer simulations, especially at high SNR. Although FSPM $(4, 2)$ cannot provide a BER advantage relative to IM $(4, 3, 4)$ and MM-IM $(4, 2)$, OFSPM $(4, 2)$ exhibits superior BER performance relative to all benchmarks at high SNR while achieving an enhanced SE. These results arise from the fact that the set partitions in the selected FSPM codebook exhibit lower minimum Hamming distance than the set partitions in the selected OFSPM codebook. Moreover, the codewords related to set partitions in OFSPM are capable of preserving the same minimum Hamming distance property as the codewords related to the permutations in the MM-IM scheme, although the number of index bits for the former increases. As also seen from the figure, when we remove the codewords of OFSPM $(4, 2)$ by using Algorithm 3, the remaining codebook, in other words, OFSPM $(4, 2)$, Alg. 3, achieves a higher diversity order as well as considerably outperforms other methods.

In Fig. 3.9, we compare the BER performance of FSPM $(4, 4)$, OFSPM $(4, 4)$, and OFSPM $(4, 4)$, Alg. 3 with MM-IM $(4, 4)$, IM $(4, 3, 8)$, and 8-PSK schemes.

We also provide a BER curve labeled "OFSPM $(4, 4)$, QAM" for an OFSPM concept employing four different 4-QAM constellations to form the OFSPM codewords. These constellations are obtained by employing the set partitioning technique described in [100] to a 16-QAM constellation. We use the same SPM codewords as in Fig. 3.8 for the SPM variants. In this figure, FSPM $(4, 4)$ and IM $(4, 3, 8)$; OFSPM $(4, 4)$, Alg. 3, MM-IM $(4, 4)$, and 8-PSK; OFSPM $(4, 4)$ and OFSPM $(4, 4)$, QAM schemes have the same SEs which are 2.75 bpcu, 3 bpcu, and 3.25 bpcu, respectively. Similar results to those of Fig. 3.8 can be observed in Fig. 3.9 for a higher modulation order. On the other hand, OFSPM $(4, 4)$, QAM outperforms both OFSPM $(4, 4)$ and MM-IM $(4, 4)$ at low SNR; however, OFSPM $(4, 4)$ exhibits slightly better BER performance than OFSPM $(4, 4)$, QAM and MM-IM $(4, 4)$ at high SNR. The behavior of the BER curves at low SNR can be explained by the MEDs between the codewords. The MED between "OFSPM $(4, 4)$, QAM" codewords is 0.8944; however, the MED between OFSPM $(4, 4)$ and MM-IM $(4, 4)$ codewords is just 0.5518. On the other hand, the BER performance at high SNR will be dominated by the Euclidean distances between the minimum Hamming distance codewords. In that case, the Euclidean distance between OFSPM $(4, 4)$ and MM-IM $(4, 4)$ codewords is given by 1.4142; however, the Euclidean distance between "OFSPM $(4, 4)$, QAM" codewords is just 1.2649. The comparison between OFSPM $(4, 4)$ and MM-IM $(4, 4)$ results from the number of index bits. OFSPM $(4, 4)$ produces more index bits than MM-IM $(4, 4)$ by using set partitions rather than permutations. Such index bits become undesirable at low SNR due to the MED between the associated codewords. However, at high SNR, they become desirable since these codewords provide better Hamming distance properties than the codewords associated with conventional modulation. It is also important to note that OFSPM $(4, 4)$, Alg. 3 achieves the best performance at high SNR by providing almost 1.4 dB SNR gain compared to OFSPM $(4, 4)$ and MM-IM $(4, 4)$ at a BER value of 10^{-4}.

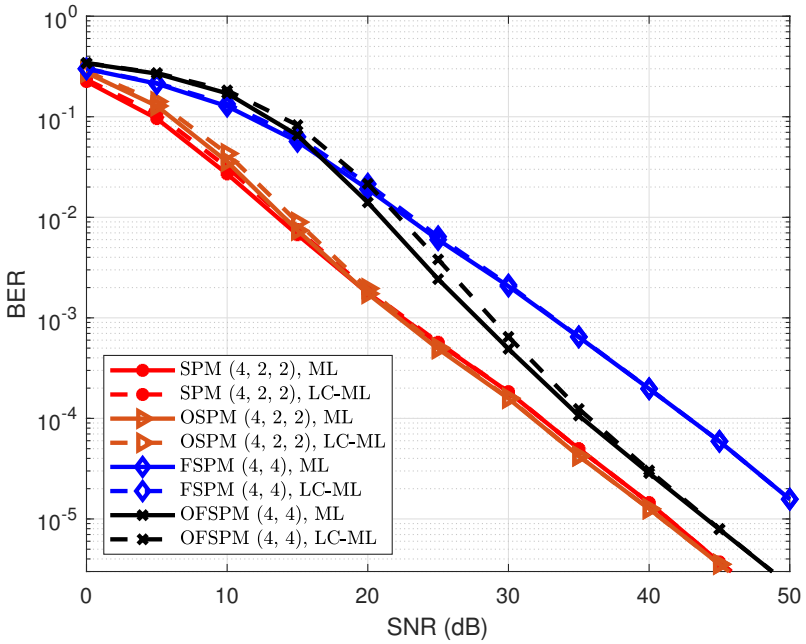


Figure 3.10. BER comparison of the proposed LC-ML and the optimum ML detectors for the SPM techniques.

Fig. 3.10 demonstrates the BER performance of the proposed LC-ML detector compared to the optimum ML detector for SPM $(4, 2, 2)$, OSPM $(4, 2, 2)$, FSPM $(4, 4)$, and OFSPM $(4, 4)$. In this figure, curves with extensions "LC-ML" and "ML" in the legend relate to the SPM schemes applying the proposed LC-ML and optimum ML detectors, respectively. As seen from the figure, the performance of the proposed LC-ML detector is very close to that of the optimum ML detector, especially at low and high SNR.

3.7 Summary

In this chapter, we showed a novel modulation concept, which we call SPM. We represented several variants of the concept and showed a practical implementation in the context of parallel channels. Moreover, we defined a codebook selection problem for the proposed techniques and expressed such a problem as a clique problem in graph theory. We further provided an efficient solution for the codebook selection.

Then, we investigated the performance of the new techniques in terms of their data rates and BER characteristics, and we presented asymptotic results regarding the performance. We compared the proposed SPM variants with the appropriate benchmarks. Through computer simulations and theoretical calculations, it is shown that the proposed SPM techniques can provide noteworthy improvements compared to conventional modulation, IM, DM-IM, and MM-IM in terms of data rate and BER.

Chapter 4

Encoding Using Compositions

Due to the remarkable growth in mobile data services, intelligent devices, and connectivity, the need for spectrally/energy efficient and reliable communication systems increases day by day [114]. In this context, IM has attracted remarkable attention due to its attractive advantages like better error performance and improved energy/spectral efficiency over conventional modulation and multiplexing schemes. Considering the conspicuous success of OFDM in cellular networks, local area networks, audio broadcasting, and power line networks, the adaptation of IM to OFDM has also attracted several researchers' attention [1, 2, 15, 19, 38]. It has been shown by these studies that OFDM-IM is capable of achieving better error performance, higher data rate, and energy efficiency than OFDM.

On the other hand, 5G's key enabling technologies are expected to satisfy the user demands for a considerably high data rate, ultra-reliable transmission, and very low latency. However, it is anticipated that 6G and the following next-generation networks will require even more data rate, better error performance, and lower latency than what 5G offers. In this regard, the existing OFDM and OFDM-IM schemes require further improvement in error performance and SE to be deployed efficiently in the next-generation networks.

In the previous chapter, motivated by the advantages of the combinatorial techniques as well as the requirements of the next-generation networks, we designed a novel modulation concept based on a combinatorial tool called set partitioning. The chapter unveiled a novel application of set partitioning in encoding. It also showed that the resulting SPM techniques attain a superior performance compared to the existing methods.

As in the previous chapter, we continue to work on the combinatorial encoding methods in this chapter and uncover a novel application of another combinatorial tool called composition (or integer composition) in the modulation design. In this regard, we propose two different modulation concepts, *weak composition modulation (WCM)* and *composition modulation (CM)*, that exploit weak and strict compositions of an integer, respectively, to form their codewords. Moreover, we propose another concept called *index and composition modulation (ICM)* that combines the beneficial properties of IM and CM in one technique. In ICM, we encode information using the combinations of active/deactive codeword elements as in IM and the combinatorial framework of integer compositions as in CM. Then, we depict practical parallel channel implementations of the proposed schemes and use the compositions to embed information into the energies of the codeword elements. We investigate the performance of the proposed methods in this chapter and obtain closed-form BER and mutual information expressions. For the ICM scheme, we analyze SE and design an efficient low-complexity detector to overcome the high complexity arising from the optimum ML detector. Our analytical, as well as numerical, findings indicate that the proposed concepts can achieve a substantially better performance than the IM and conventional modulation techniques.

4.1 Basic Definitions

In this section, we describe the basic idea of composition encoding. We begin with some useful definitions and relations. Then, we propose different modulation concepts based on CM.

Definition 6. Multiset: A multiset can be defined as a modified concept of a set, in which each element is allowed to be repeated a number of times. An example of a multiset can be given by

$$\{\overbrace{a_1,\ldots,a_1}^{\mu_1},\overbrace{a_2,\ldots,a_2}^{\mu_2},\ldots,\overbrace{a_N,\ldots,a_N}^{\mu_N}\}.$$

Here, we have N different elements, $a_1, a_2, \ldots, a_N$, in the multiset. These elements are repeated by $\mu_1, \mu_2, \ldots, \mu_N$ times, respectively, where $\mu_n \in \mathbb{Z}_{\geq 0}$, $n \in \{1, 2, \ldots, N\}$, and $\mathbb{Z}_{\geq 0}$ is the set of non-negative integers. We denote the cardinality of this set by I and define the cardinality as the total number of elements, i.e., $I = \mu_1 + \mu_2 + \ldots + \mu_N$.

Definition 7. Multiset number: It is the number of multisets of cardinality I that an N-set has. To select a multiset of cardinality I, we have to select non-negative integers $\mu_1, \mu_2, \ldots, \mu_N$ such that $I = \mu_1 + \mu_2 + \ldots + \mu_N$. The number of solutions to that equation is $\binom{I+N-1}{I}$. The proof follows from the fact that the problem is equivalent to finding the number of ways of placing I indistinguishable balls into N distinguishable boxes [115].

Definition 8. Partition: A partition of a positive integer is a multiset of positive integers that sum to that integer. Hence, a partition of an integer I can be written as $\{\nu_1, \nu_2, \ldots, \nu_N\}$ where $I = \nu_1 + \nu_2 + \ldots + \nu_N$ and $\nu_1 \geq \nu_2 \geq \ldots \geq \nu_N \geq 1$. We denote such a partition by $(\nu_1, \nu_2, \ldots, \nu_N)$.

Definition 9. Composition: A composition of an integer is an ordered list of positive integers that sum to that integer. Integer composition is sometimes called

ordered integer partition. Hence, ordering the elements of an integer partition yields the same partition; however, it yields a different integer composition. It is sometimes called a strict composition since the integers that form a composition are strictly positive.

Let $C(I)$ be the total number of compositions of an integer I and let $C_N(I)$ be the number of compositions of I which have exactly N summands where $I \geq N$.[1] One can check that $C(I) = 2^{I-1}$ and $C_N(I) = \binom{I-1}{N-1}$ [115].

Definition 10. Weak composition: A weak composition of an integer is an ordered list of non-negative integers that sum to that integer. Unlike an integer composition, we allow the list elements to be zero for a weak composition. If we do not limit the number of the parts, there are infinitely many weak compositions of an integer.

Definition 11. A weak N-composition of I: A weak N-composition of an integer I is an ordered list of N non-negative integers that sum to that integer. The number of weak compositions of I into N parts is $\binom{I+N-1}{N-1}$.

4.2 Composition Modulation

4.2.1 Weak Composition Modulation

In WCM, a codebook of L_{WCM} codewords $\mathbf{x}_1, \ldots, \mathbf{x}_{L_{WCM}}$ is formed in a way that each codeword has N elements that are determined by the weak N-compositions of an integer I. Hence, when we form N-tuples related to the codewords of a WCM codebook, we take the elements in an N-tuple equal to the square roots of the elements in a weak N-composition of an integer I. More explicitly, the ith element x_i, $i = 1, \ldots, N$, of an N-tuple is chosen as $x_i = \sqrt{\mu_i}$ where $\mu_i \in \{0, 1, \ldots, I\}$ is the ith

[1] For this case, we say that such a composition has N parts.

summand of the composition.[2] For example, two weak 4-compositions of $I = 4$ can be given as $4 = 1 + 1 + 1 + 1$ and $4 = 4 + 0 + 0 + 0$. In this case, the corresponding N-tuples are given as $(1, 1, 1, 1)$ for the former and $(2, 0, 0, 0)$ for the latter. Notice that we ensure the same total energy for each N-tuple since we incorporate compositions of the same integer throughout the codebook. Moreover, the number of codewords in a WCM codebook can be given as the number of weak N-compositions of I, i.e., $L_{WCM} = \binom{I+N-1}{N-1}$.

4.2.2 Composition Modulation

Unlike WCM which uses the weak compositions of an integer to form its codewords, we use only strict compositions of an integer to form the CM codewords. Hence, we form the N-tuples for CM in a way that the elements are equal to the square roots of the elements related to the compositions of an integer I that have N summands. More explicitly, the ith element x_i, $i = 1, \ldots, N$, of an N-tuple is chosen as $x_i = \sqrt{v_i}$ where $v_i \in \{1, \ldots, I - N + 1\}$ is the ith summand of the composition. Note that one needs to pick $I \geq N$ to be able to construct the CM codewords. The number of codewords, L_{CM}, in this scheme, can be given as the number of compositions of an integer I with N parts, i.e., $L_{CM} = \binom{I-1}{N-1}$. Like WCM, we ensure the same total energy across the CM codebook.

4.2.3 Index and Composition Modulation

An ICM codebook consists of L_{ICM} codewords, and each codeword is an N-tuple of nonnegative real numbers which can be regarded as a vector $\mathbf{x}_l = [x_{l1}, x_{l2}, \ldots, x_{lN}]$, $l = 1, 2, \ldots, L_{ICM}$, in an Euclidean space $\mathcal{S}$ of N dimensions where $x_{ln} \in \mathbb{R}_{\geq 0}$, $n = 1, 2, \ldots, N$, and $\mathbb{R}_{\geq 0}$ is the set of nonnegative real numbers. In an ICM codeword,

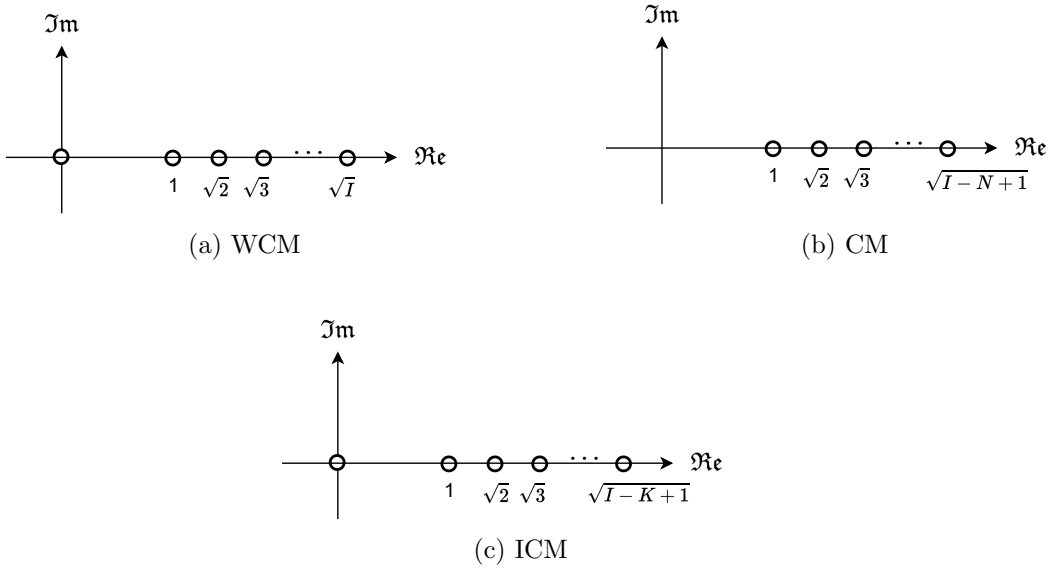
[2] Here, we take the square root of the summands of the compositions to ensure that the energies of the related codeword elements are equal to the summands. That is also the case for the other modulation concepts described in this section.

K elements are positive real numbers where $K \leq N$, whereas the remaining $N-K$ elements are zeros. Here, we call the K positive real and $N-K$ zero elements as activated and deactivated elements, respectively, since no energy is used to form the zero elements. All other codewords of the codebook can be obtained by permuting the order of the N elements in the codeword. The values of K activated elements in each codeword are chosen according to the integers that form the compositions of an integer I with K parts. Hence, one can obtain different codewords by using different compositions of an integer. More explicitly, when we form $\mathbf{x}_l$, we first determine the indices of K activated elements in an N-tuple and denote the set of these indices as $\mathcal{I} \coloneqq \{\alpha_1, \alpha_2, \ldots, \alpha_K\}$ where $\alpha_k \in \{1, 2, \ldots, N\}$ and $k = 1, 2, \ldots, K$. Then, the kth activated element $x_{l\alpha_k}$ of the N-tuple is chosen as $x_{l\alpha_k} = \sqrt{\nu_k}$ where $\nu_k \in \{1, \ldots, I-K+1\}$ is the kth summand of the composition and $I = \nu_1 + \nu_2 + \ldots + \nu_K$. Note that one needs to pick $I \geq K$ to be able to construct the ICM codewords. Since the number of combinations regarding the order of K activated and $N-K$ deactivated elements is $\binom{N}{K}$ and the number of integer compositions of I with K parts is $\binom{I-1}{K-1}$, the number of codewords in an ICM codebook is $L_{ICM} = \binom{N}{K}\binom{I-1}{K-1}$. In Table 4.1, we give a codebook generation example for an ICM scheme when $N = I = 3$ and $K = 2$. As seen from the leftmost column of the table, we first determine the indices of $K = 2$ activated elements. Then, we determine the compositions of an integer $I = 3$ with $K = 2$ parts and map the summands in those compositions to the elements of the ICM codewords according to the indices of the activated elements. The resulting ICM codewords are shown in the rightmost column.

In Fig. 4.1, we illustrate the constellation points used to form the WCM, CM, and ICM codewords. We use black circles to represent the points. As seen from the figure, all the techniques exploit non-negative real numbers to form their codewords. However, as we will show in the next section, the applications of the proposed encoding techniques are not limited to the real constellation points. On the other hand,

Table 4.1. Codebook generation example for ICM when $N = I = 3$ and $K = 2$.

Indices of Activated Elements, $\mathcal{I}$	Compositions	ICM Codeword
$\{1, 2\}$	3=1+2	$\mathbf{x}_1 = [1, \sqrt{2}, 0]$
$\{1, 3\}$		$\mathbf{x}_2 = [1, 0, \sqrt{2}]$
$\{2, 3\}$		$\mathbf{x}_3 = [0, 1, \sqrt{2}]$
$\{1, 2\}$	3=2+1	$\mathbf{x}_4 = [\sqrt{2}, 1, 0]$
$\{1, 3\}$		$\mathbf{x}_5 = [\sqrt{2}, 0, 1]$
$\{2, 3\}$		$\mathbf{x}_6 = [0, \sqrt{2}, 1]$



(a) WCM

(b) CM

(c) ICM

Figure 4.1. The constellation points that are used to form the WCM, CM, and ICM symbols.

the WCM and ICM methods utilize the origin to construct their N-dimensional codewords, whereas this is not the case for CM as shown in Fig. 4.1b.

4.3 Practical Model For Parallel Channels

This section presents different system models in which we apply the proposed WCM, CM, and ICM methods to parallel channels. As mentioned in the previous chapter, one can form parallel channels across different frequencies, time slots, and antennas. Here, we consider a practical model similar to the IM and SPM techniques defined

in the previous chapters. Therefore, once again, we assume b input bits enter the transmitter, and these bits are separated into $P = b/f$ blocks, each having f input bits. Similarly, the total number of channels N_T is divided into $P = N_T/N$ blocks, each having N channels.[3] For each block of input bits, f information bits are modulated by the WCM, CM, and ICM encoders. The resulting modulated symbols are conveyed through N channels.

Since each bit and each channel block have the same mapping operation for the proposed schemes, we focus on a single block, the pth block (where $p \in \{1, 2, \ldots, P\}$), in the rest of this section. In the blocks of WCM and CM, the f information bits are further divided into two parts, one of them having f_1 bits and the other one having f_2 bits with $f_1 + f_2 = f$. Besides, in each block of the ICM method, we divide f information bits into three parts, having f_1, f_2, and f_3 bits with $f_1 + f_2 + f_3 = f$. As we will explain in the next subsections, the remaining transmitter and receiver operations will be different for the proposed methods.

4.3.1 WCM

4.3.1.1 Transmitter

At the transmitter, the first f_1 bits are used to determine a specific weak N-composition of a positive integer I in the pth block which is denoted by $\omega^p := \{\mu_1, \mu_2, \ldots, \mu_N\}$ where $\mu_i \in \{0, 1, \ldots, I\}$ and $I = \mu_1 + \mu_2 + \ldots + \mu_N$. Therefore, $f_1 = \lfloor \log_2 \binom{I+N-1}{N-1} \rfloor$. Then, we map the ith summand in the selected weak composition to the energy of the symbol carried by the ith channel in a way that the total energy is preserved. Hence, the energy related to the ith codeword element can be given as $E_i = \mu_i E_T / I$ where $E_T = \sum_{i=1}^{N} |x_i^p|^2$ is the total energy of a block and x_i^p is the symbol carried by the ith channel in the pth block. Let us denote the set that contains the energies of the

[3] As exemplified in the previous section, such channels can be regarded as the subcarriers of the well-known OFDM mechanism.

Table 4.2. Codeword generation and look-up table implementation example for WCM when $N = I = 3$ and $\lambda = 1$.

Weak Compositions	$\mathcal{E}^p$	WCM Codeword	f_1 bits
3=0+0+3	$\{0, \quad 0, \quad E_T\}$	(0, 0, 8-PSK)	[0 0 0]
3=0+1+2	$\{0, \quad \frac{E_T}{3}, \frac{2E_T}{3}\}$	(0, BPSK, QPSK)	[0 0 1]
3=0+2+1	$\{0, \quad \frac{2E_T}{3}, \quad \frac{E_T}{3}\}$	(0, QPSK, BPSK)	[0 1 0]
3=0+3+0	$\{0, \quad E_T, \quad 0\}$	(0, 8-PSK, 0)	[0 1 1]
3=1+0+2	$\{\frac{E_T}{3}, \ 0, \ \frac{2E_T}{3}\}$	(BPSK, 0, QPSK)	[1 0 0]
3=1+1+1	$\{\frac{E_T}{3}, \ \frac{E_T}{3}, \frac{E_T}{3}\}$	(BPSK, BPSK, BPSK)	[1 0 1]
3=1+2+0	$\{\frac{E_T}{3}, \frac{2E_T}{3}, \quad 0\}$	(BPSK, QPSK, 0)	[1 1 0]
3=2+0+1	$\{\frac{2E_T}{3}, \ 0, \ \frac{E_T}{3}\}$	(QPSK, 0, BPSK)	[1 1 1]
3=2+1+0	$\{\frac{2E_T}{3}, \frac{E_T}{3}, \quad 0\}$	(QPSK, BPSK, 0)	unused
3=3+0+0	$\{E_T, \quad 0, \quad 0\}$	(8-PSK, 0, 0)	unused

codeword elements as $\mathcal{E}^p \coloneqq \{\mu_1 E_T/I, \mu_2 E_T/I, \ldots, \mu_N E_T/I\}$. Moreover, the mapping of f_1 bits to the energies can be implemented by using a look-up table. Then, the remaining f_2 bits are used to determine M-PSK symbols on the activated channels. Note that since we utilize weak compositions in this structure, some codewords elements may have no energy, or in other words, some channels may be deactivated, depending on the selected weak composition. On the other hand, to send the same amount of information bits for all weak compositions, we adjust the modulation order on the activated channels according to the summands of a weak composition and choose the size of the constellation on the ith channel as $2^{\lambda\mu_i}$ if the ith codeword element has nonzero energy, i.e., $\mu_i \neq 0$. Here, λ is a positive integer that we use to alter the number of transmitted bits. In that way, we ensure that $f_2 = \lambda I$.

Let us consider an example of how we decide the modulation orders on the activated channels and implement a look-table based mapping between the energies of the codeword elements and the first f_1 information bits for the proposed scheme when $N = I = 3$ and $\lambda = 1$. We list the weak 3-compositions of $I = 3$ along with the energy sets, $\mathcal{E}^p$, and the corresponding WCM codewords in Table 4.2. As seen from the table, we map the summands of the weak compositions to the codeword elements and

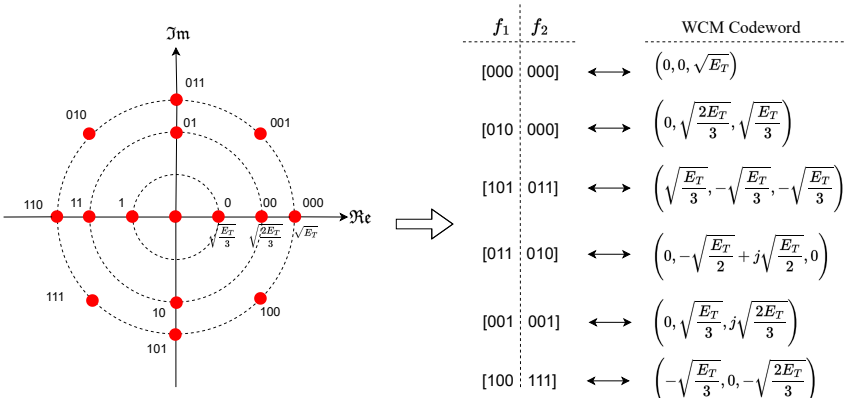


Figure 4.2. Constellation diagram and associated codewords of WCM when $N = I = 3$ and $\lambda = 1$.

use them to determine the modulation order on the activated channels. To better illustrate the codewords of the WCM method in Table 4.2, we demonstrate the corresponding constellation diagram (on the left) and associated WCM codewords (on the right) in Fig. 4.2.[4] The constellation diagram indicates the information bits mapped to the constellation points. We perform the mapping between $f_1 + f_2$ information bits and the WCM codewords by exploiting the look-up table in Table 4.2 and the mapping on the constellation diagram given in the figure. Here, the constellation points are located on the circles whose radii are $\sqrt{\frac{E_T}{3}}$, $\sqrt{\frac{2E_T}{3}}$, and $\sqrt{E_T}$. As seen from the figure, the outermost circle has the highest number of points, i.e., eight points. The number of points decreases when we proceed from the outermost circle to the innermost one. We also have a point at the origin since we deactivate some codeword elements in the WCM codebook.

The SE of the WCM method is given by

$$\eta_{WCM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 \binom{I+N-1}{N-1} \rfloor + \lambda I}{N}. \tag{4.1}$$

By assuming $\mathcal{M}_i$ is a $2^{\lambda\mu_i}$-ary constellation, the WCM symbol vector corresponding

[4]Note that we have 64 WCM codewords when $N = I = 3$ and $\lambda = 1$. However, the figure shows only a limited number of codewords due to the limited space.

to the pth block can be written as $\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p]$ where $x_i^p \in \mathcal{M}_i$ when $\mu_i \neq 0$, otherwise $x_i^p = 0$. After obtaining symbol vectors for all blocks, a block creator forms the overall symbol vector $\mathbf{x} \coloneqq [x(1), x(2), \ldots, x(N_T)]^T = [\mathbf{x}^1, \ldots, \mathbf{x}^p, \ldots, \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$. Note that the energy of the symbol carried by the ith channel can be written as $E_i^p = |x_i^p|^2 = \mu_i E_T / I$.

4.3.1.2 Receiver

At the receiver, the received signal vector can be written as

$$\mathbf{y} \coloneqq [y(1), y(2), \ldots, y(N_T)]^T = \mathbf{X}\mathbf{h} + \mathbf{w} \tag{4.2}$$

where $\mathbf{X} = \text{diag}(\mathbf{x})$. Moreover, $\mathbf{h}$ and $\mathbf{w}$ are $N_T \times 1$ channel and noise vectors, respectively.

Since the encoding procedure for each block is independent of others, decoding can be performed independently at the receiver. Hence, using the ML detection, the detected symbol vector for the pth block can be written as

$$(\hat{\mathcal{E}}^p, \hat{\mathbf{x}}^p) = \arg \min_{\mathcal{E}^p, \mathbf{x}^p} ||\mathbf{y}^p - \mathbf{X}^p \mathbf{h}^p||^2 \tag{4.3}$$

where $\mathbf{y}^p = [y((p-1)N+1), \ldots, y(pN)]^T$, $\mathbf{X}^p = \text{diag}(\mathbf{x}^p)$, and $\mathbf{h}^p = [h((p-1)N+1), \ldots, h(pN)]^T$.

4.3.2 CM

4.3.2.1 Transmitter

The CM encoder uses the first f_1 bits to determine a specific (strict) composition of a positive integer I with N parts. Therefore, $f_1 = \lfloor \log_2 \binom{I-1}{N-1} \rfloor$. We, then, map the ith element in the selected composition to the energy of the ith codeword element, i.e.,

$E_i = \nu_i E_T/I$ where $\nu_i \in \{1, 2, \ldots, I-N+1\}$ and $I = \nu_1 + \nu_2 + \ldots + \nu_N$. Unlike WCM, all parts of a composition are positive, thus all channels are activated in CM. We denote the set that contains the associated energies as $\xi^p := \{\nu_1 E_T/I, \nu_2 E_T/I, \ldots, \nu_N E_T/I\}$. Like WCM, the mapping of f_1 bits to the set, ξ^p, can be done by using a look-up table. Once we decide ξ^p according to the incoming f_1 bits, $f_2 = N \log_2 M$ bits are used to determine M-PSK symbols carried by the channels. Thus, the SE of the CM approach can be written as

$$\eta_{CM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 \binom{I-1}{N-1} \rfloor + N \log_2 M}{N}. \tag{4.4}$$

Here, different from WCM, we use the same modulation order, M, on each channel since we activate all of the channels during a transmission interval. For brevity, we use the same notation as the WCM concept to denote the CM symbol vector. Hence, such a vector for the pth block can be written as $\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p]$ where $x_i^p \in \mathcal{M}$ and $\mathcal{M}$ is an M-PSK constellation. After obtaining symbol vectors of all blocks, the overall CM vector is formed as $\mathbf{x} := [x(1), x(2), \ldots, x(N_T)]^T = [\mathbf{x}^1, \ldots, \mathbf{x}^p, \ldots, \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$. Note that the energy of the symbol carried by the ith channel can be given as $E_i^p = |x_i^p|^2 = \nu_i E_T/I$. Note also that the proposed scheme is equivalent to parallel transmission of the M-PSK symbols when $I = N$. For this case, we have only $\binom{N-1}{N-1} = 1$ composition, and $\nu_i = 1$, $\forall i \in \{1, 2, \ldots, N\}$, thus all the codeword elements have the same energy.

In Fig. 4.3, we show the constellation points that form the CM codewords when $N = 3$, $I = 5$, and $M = 4$. As in the WCM scheme, we place the constellation points on the circles with different radii. However, unlike the WCM method, the number of points on each circle remains unchanged, i.e., we have $M = 4$ points on each circle. Moreover, the CM scheme does not have a constellation point at the origin since this scheme activates all of its codeword elements.

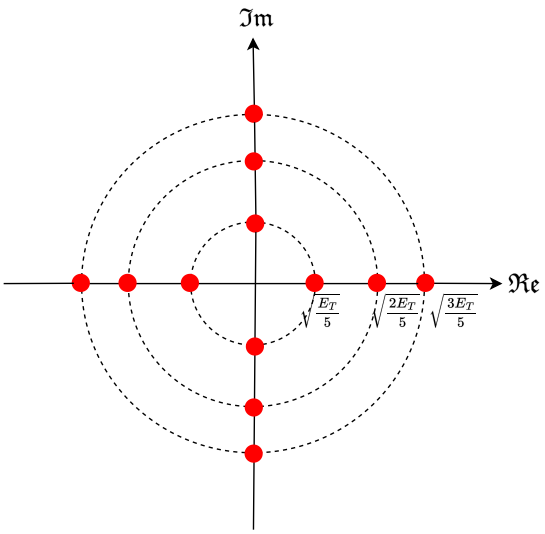


Figure 4.3. The constellation points that are used to construct the CM codewords when $N = 3, I = 5$, and $M = 4$.

4.3.2.2 Receiver

At the receiver, the received signal can be obtained by substituting the CM vector into (4.2) instead of the WCM vector. Moreover, the ML detection can be performed in the same way as (4.3), and the detected symbol can be written as

$$(\hat{\xi}^p, \hat{\mathbf{x}}^p) = \arg \min_{\xi^p, \mathbf{x}^p} ||\mathbf{y}^p - \mathbf{X}^p \mathbf{h}^p||^2. \tag{4.5}$$

4.3.3 ICM

In this subsection, we will show the transmitter and receiver operations related to the practical implementation of the ICM concept in parallel channels. Moreover, we will present a low-complexity ML detector for the ICM concept and a simple bit-to-composition mapping algorithm that does not require look-up table implementation for the mapping of the weak compositions of I with N summands (equivalently compositions of $I + N$ with N summands) to the decimal numbers and bits. It is worth mentioning that the mapping algorithm is applicable to all of the proposed

CM techniques.

4.3.3.1 Transmitter

The first $f_1 = \lfloor \log_2 \binom{N}{K} \rfloor$ bits are used to determine the activated K channels. Then, the $f_2 = \lfloor \log_2 \binom{I-1}{K-1} \rfloor$ bits are used to determine the specific composition of an integer I with K parts where $I \geq K$. The energies of the symbols on the activated channels are chosen according to the composition of an integer I with K parts. Let us denote the sets that comprise the indices of the activated channels and the energies related to the activated channels in the pth block, respectively, as $\mathcal{I}^p := \{\alpha_1, \alpha_2, \ldots, \alpha_K\}$ and $\beta^p := \{\nu_1 E_T/I, \nu_2 E_T/I, \ldots, \nu_K E_T/I\}$ where $\alpha_k \in \{1, 2, \ldots, N\}$, $k = 1, 2, \ldots, K$, $\alpha_1 < \alpha_2 < \ldots < \alpha_K$, $\nu_k \in \{1, 2, \ldots, I-K+1\}$, and $I = \nu_1 + \nu_2 + \ldots + \nu_K$. Here, the kth element in β^p is related to the α_kth channel. Once we decide $\mathcal{I}^p$ and β^p according to the f_1 and f_2 bits, respectively, $f_3 = K \log_2 M$ bits are used to determine M-PSK symbols carried by the activated channels. Hence, the SE per channel for the ICM system can be given as

$$\eta_{ICM} = \frac{f_1 + f_2 + f_3}{N} = \frac{\lfloor \log_2 \binom{N}{K} \rfloor + \lfloor \log_2 \binom{I-1}{K-1} \rfloor + K \log_2 M}{N}. \tag{4.6}$$

The mapping of f_1 bits to the indices of the activated channels and f_2 bits to the transmit energy levels can be implemented by using a look-up table. In Table 4.3, we present an example of how these mappings are performed when $N = I = 4$ and $K = 3$. Since $f_1 = \lfloor \log_2 \binom{4}{3} \rfloor = 2$, the first two bits, p_1, in the corresponding bitstream $[p_2\ p_1]$ are used to determine the indices of the activated elements. Then, the remaining $f_2 = \lfloor \log_2 \binom{I-1}{K-1} \rfloor = \lfloor \log_2 \binom{3}{2} \rfloor = 1$ bit, p_2, is used to determine a specific composition of $I = 4$ with $K = 3$ parts. For example, when $[p_2\ p_1] = [0\ 1\ 1]$ bits enter the ICM encoder, the first two bits "$p_1 = 11$" choose the indices of the activated elements as $\mathcal{I}^p := \{2, 3, 4\}$, then the remaining bit "$p_2 = 0$" chooses the energy set

Table 4.3. Look-up table implementation for ICM when $N = I = 4$ and $K = 3$.

Compositions	β^p	$\mathcal{I}^p$	ICM Energy Vector	$[p_2 \vdots p_1]$
4=1+1+2	$\{\frac{E_T}{4}, \frac{E_T}{4}, \frac{2E_T}{4}\}$	$\{1,2,3\}$	$\frac{E_T}{4}\ \frac{E_T}{4}\ \frac{2E_T}{4}\ 0$	$[0 \vdots 0\ 0]$
		$\{1,2,4\}$	$\frac{E_T}{4}\ \frac{E_T}{4}\ 0\ \frac{2E_T}{4}$	$[0 \vdots 0\ 1]$
		$\{1,3,4\}$	$\frac{E_T}{4}\ 0\ \frac{E_T}{4}\ \frac{2E_T}{4}$	$[0 \vdots 1\ 0]$
		$\{2,3,4\}$	$0\ \frac{E_T}{4}\ \frac{E_T}{4}\ \frac{2E_T}{4}$	$[0 \vdots 1\ 1]$
4=1+2+1	$\{\frac{E_T}{4}, \frac{2E_T}{4}, \frac{E_T}{4}\}$	$\{1,2,3\}$	$\frac{E_T}{4}\ \frac{2E_T}{4}\ \frac{E_T}{4} 0$	$[1 \vdots 0\ 0]$
		$\{1,2,4\}$	$\frac{E_T}{4}\ \frac{2E_T}{4}\ 0\ \frac{E_T}{4}$	$[1 \vdots 0\ 1]$
		$\{1,3,4\}$	$\frac{E_T}{4}\ 0\ \frac{2E_T}{4}\ \frac{E_T}{4}$	$[1 \vdots 1\ 0]$
		$\{2,3,4\}$	$0\ \frac{E_T}{4}\ \frac{2E_T}{4}\ \frac{E_T}{4}$	$[1 \vdots 1\ 1]$
4=2+1+1	$\{\frac{2E_T}{4}, \frac{E_T}{4}, \frac{E_T}{4}\}$	$\{1,2,3\}$	$\frac{2E_T}{4}\ \frac{E_T}{4}\ \frac{E_T}{4}\ 0$	unused
		$\{1,2,4\}$	$\frac{2E_T}{4}\ \frac{E_T}{4}\ 0\ \frac{E_T}{4}$	unused
		$\{1,3,4\}$	$\frac{2E_T}{4}\ 0\ \frac{E_T}{4}\ \frac{E_T}{4}$	unused
		$\{2,3,4\}$	$0\ \frac{2E_T}{4}\ \frac{E_T}{4}\ \frac{E_T}{4}$	unused

$\beta^p := \{E_T/4, E_T/4, 2E_T/4\}$. Then, we map the set, β^p, to the energies of the activated channels and obtain the corresponding ICM energy vector as shown in the table.[5]

Once we decided the activated channels and their energies according to the first $f_1 + f_2$ bits, then the remaining f_3 bits are used to modulate the signals on the activated channels by using an M-PSK constellation, $\mathcal{M}$. Hence, in the pth block, the ICM symbol vector can be written as $\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p]$ where $x_i^p \in \{\emptyset\} \cup \mathcal{M}$. The energy of the symbol carried by the kth activated channel is $E_k^p = |x_k^p|^2 = \nu_k E_T/I$ where $k = 1, 2, \ldots, K$. After obtaining symbol vectors of all blocks, the overall ICM vector is formed as $\mathbf{x} := [x(1), x(2), \ldots, x(N_T)]^T = [\mathbf{x}^1, \ldots, \mathbf{x}^p, \ldots, \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$.[6]

To better illustrate the codewords of an ICM codebook, we show its constellation

[5] As seen from the table, some of the symbol vectors are not included in the codebook due to the mapping between the fixed-length bit sequences and the symbol vectors. However, one can extend the codebook by exploiting different concepts like in [116] to avoid unused symbol vectors and increase the data rate.

[6] We assume that the elements of $\mathbf{x}$ are interleaved sufficiently and the maximum spacing is achieved for the channels.

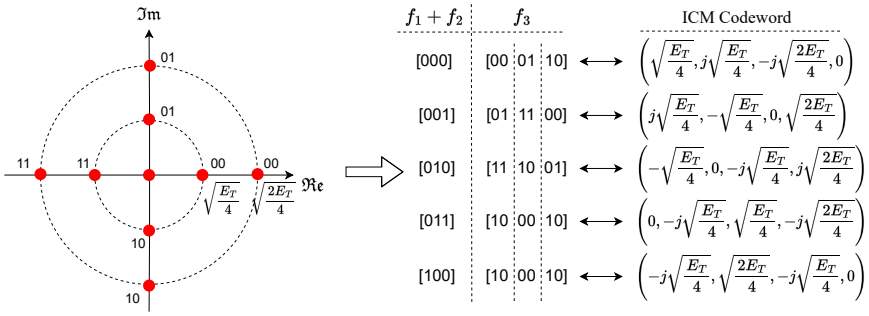


Figure 4.4. The constellation points and codewords that form the ICM codebook when $N = I = 4$, $K = 3$, and $M = 4$.

points (on the left) and codewords (on the right) in Fig. 4.4 for $N = I = 4$, $K = 3$, and $M = 4$.[7] The figure also indicates the information bits mapped to the constellation points. We perform the mappings between $f_1 + f_2$ information bits by exploiting the look-up table in Table 4.3, whereas we use the mapping illustrated on the constellation diagram for the remaining f_3 bits. Here, the first $f_1 + f_2 = 3$ bits are used to select the ICM energy vectors as shown in Table 4.3. The remaining $f_3 = 6$ bits are used to select the modulation symbols on the activated codeword elements. Notice from the figure that dashed lines divide $f_3 = 6$ bits into three groups where each group has two bits. We use the bits in each group to decide the modulation symbols on the activated codeword elements. The energy vectors and modulation symbols form the corresponding ICM codewords as we indicate on the right-hand side of the figure.

Example 3. In Fig. 4.5, we show a practical application of the ICM technique in parallel channels. Here, we execute the ICM encoding over the subcarriers of the OFDM technique. We call the resulting method as *OFDM with ICM (OFDM-ICM)*. The figure shows the transmitter structure of the OFDM-ICM method. As shown in the figure, we divide the incoming information bits and subcarriers of the OFDM method into P groups. Each group has N subcarriers and $f = f_1 + f_2 + f_3$ information bits. In each group, we map the first f_1 bits to the indices of the activated subcarriers, whereas

[7] Note that we have 512 ICM codewords when $N = I = 4$, $K = 3$, and $M = 4$. However, the figure shows only a limited number of codewords due to the limited space.

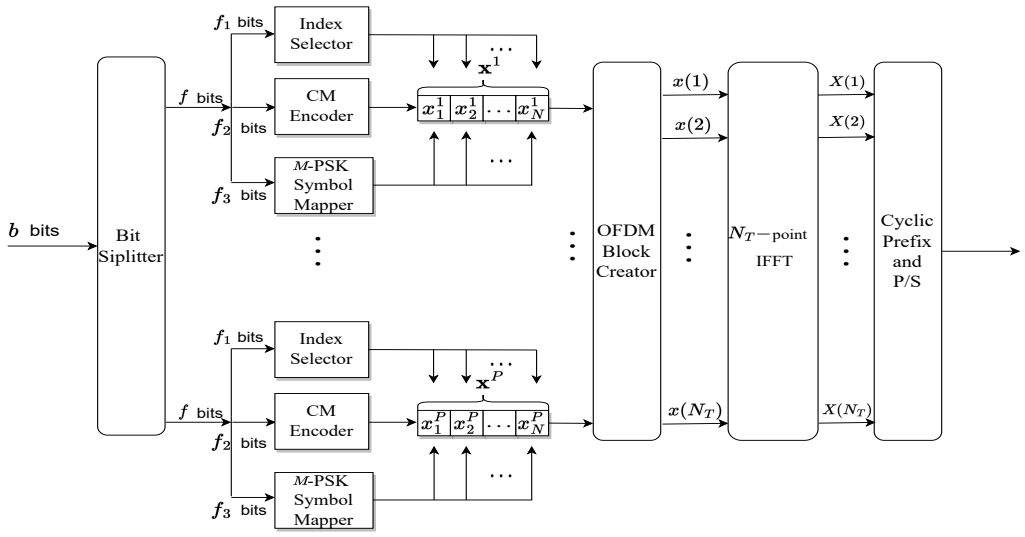


Figure 4.5. Transmitter structure of the OFDM-ICM scheme.

the preceding f_2 bits are used to decide the energies of the activated subcarriers according to a CM encoder. Then, we map the remaining f_3 bits to M-PSK symbols. Each codeword element of the OFDM-ICM scheme, x_n^p, is conveyed by a different subcarrier where $n \in \{1, 2, \ldots, N\}$ and $p \in \{1, 2, \ldots, P\}$. Therefore, the resulting OFDM symbol $\mathbf{x} = [x_1^1, x_2^1, \ldots, x_N^1, x_1^1, x_2^2, \ldots, x_N^2, \ldots, x_1^P, x_2^P, \ldots, x_N^P]$ is conveyed by the subcarriers whose center frequencies are $f_1^1, f_2^1, \ldots, f_N^1, f_1^2, f_2^2, \ldots, f_N^2, \ldots, f_1^P, f_2^P, \ldots, f_N^P$ where f_n^p is the center frequency of the nth subcarrier in the pth group. After these operations, we form the OFDM block and employ exactly the same operations as conventional OFDM. Note that one can employ the WCM and CM techniques to the OFDM mechanism in a similar way described in this example.

Remark. The ICM concept is equivalent to the classical IM and parallel transmission of the M-PSK symbols when $I = K$ and $I = K = N$, respectively. It is clear from (4.6) that the ICM method can provide a higher SE than the IM and CM techniques. Note also that ICM is fundamentally different from the IM mechanisms in [1,14,19,38] due to the use of energy levels to encode information into compositions rather than using different signals to encode data into index patterns or permutations.

4.3.3.2 Receiver

At the receiver, the received signal vector can be written as

$$\mathbf{y} \coloneqq [y(1), y(2), \ldots, y(N_T)]^T = \mathbf{X}\mathbf{h} + \mathbf{w}. \tag{4.7}$$

Since the encoding procedure for each block is independent of others, decoding can be performed independently at the receiver. Hence, using the ML detection, the detected symbol vector for the pth block can be written as

$$(\hat{\mathcal{I}}^p, \hat{\beta}^p, \hat{\mathbf{x}}^p) = \arg \min_{\mathcal{I}^p, \beta^p, \mathbf{x}^p} ||\mathbf{y}^p - \mathbf{X}^p \mathbf{h}^p||^2. \tag{4.8}$$

The optimum ML detector in (4.8) performs $2^{\lfloor \log_2 \binom{N}{K} \rfloor + \lfloor \log_2 \binom{I-1}{K-1} \rfloor} M^K$ squared Euclidean distance calculations, which result in an excessive receiver complexity when M, N, and I are high. To overcome the high complexity of the optimum ML detector, we design the following LC-ML detector for the ICM technique:

1. Calculate the following LLR value for the nth channel where $n \in \{1, 2, \ldots, N\}$:

$$\delta(n) = \ln \frac{\sum_{x \in \mathcal{M}} \sum_{E \in \Lambda} P(x(n) = \sqrt{E}x | y(n))}{P(x(n) = \emptyset | y(n))} \propto \ln \frac{\sum_{x \in \mathcal{M}} \sum_{E \in \Lambda} \exp\left(-|y(n) - h(n)\sqrt{E}x|^2 / N_0\right)}{\exp(-|y(n)|^2 / N_0)} \tag{4.9}$$

where x is a unit-energy M-PSK symbol.[8] $P(x(n) = \sqrt{E}x | y(n))$ stands for the probability that the symbol carried by the nth channel, $x(n)$, is equal to $\sqrt{E}x$ given the received signal regarding the nth channel, $y(n)$, whereas $P(x(n) = \emptyset | y(n))$ is the probability that the nth channel is deactivated given $y(n)$. Moreover, Λ is the set which consists of the possible energy levels, i.e., $\Lambda \coloneqq \{E_T/I, 2E_T/I, \ldots, (I -$

[8]Here, we drop the block superscript, p, for convenience since the proposed LC-ML detector can be operated for each block independently.

$K+1)E_T/I\}$. Note that the LLR value in (4.9) can be approximated by [1]

$$\delta(n) \approx \frac{\exp\left(-|y(n)-h(n)\sqrt{\hat{E}(n)}\hat{x}(n)|^2/N_0\right)}{\exp(-|y(n)|^2/N_0)} \tag{4.10}$$

where $(\hat{x}(n),\hat{E}(n)) = \arg\min_{x,E}|y(n)-h(n)\sqrt{E}x|^2$.

2. Sort the LLR values $(\delta(1),\delta(2),\ldots,\delta(N))$ in descending order, i.e., $\delta(\rho_1) > \delta(\rho_2) > \ldots > \delta(\rho_N)$ where $\rho_n \in \{1,2,\ldots,N\}$. Then, determine the indices of the activated channels as $\hat{\mathcal{I}} = \{\rho_1,\rho_2,\ldots,\rho_K\}$.

3. Sort the channel gains of the activated channels in descending order, i.e., $|h(\gamma_1)|^2 > |h(\gamma_2)|^2 > \ldots > |h(\gamma_K)|^2$ where $\gamma_k \in \hat{\mathcal{I}}$ and $\Gamma \coloneqq \{\gamma_1,\gamma_2,\ldots,\gamma_K\}$.

4. Determine the energy levels and M-PSK symbols on the activated K channels by following the order in Γ and using the ML detection. More explicitly, we start with the γ_1th channel and determine its energy level as well as M-PSK symbol by $(\hat{x}(\gamma_1),\hat{E}(\gamma_1)) = \arg\min_{x\in\mathcal{M},E\in\Lambda}|y(\gamma_1)-h(\gamma_1)\sqrt{E}x|^2$. The related composition can be obtained by $\hat{\nu}(\gamma_1) = \hat{E}(\gamma_1)I/E_T$. Then, we update Λ according to previously estimated energy levels and the fact that $\hat{\nu}(\gamma_k) \geq 1$ as $\Lambda \coloneqq \{E_T/I,\ldots,(I-(\hat{\nu}(\gamma_1)+K-2))E_T/I\}$, and we proceed with the γ_2th channel and determine its energy level and M-PSK symbol. For the γ_kth, $1 < k < K$, and γ_Kth channels, the sets in question can be written respectively as $\Lambda \coloneqq \{E_T/I,\ldots,(I-\Delta_k-(K-k)+1)E_T/I\}$ and $\Lambda \coloneqq \{(I-\Delta_K)E_T/I\}$ where $\Delta_k = \sum_{l=1}^{k}\hat{\nu}(\gamma_l)$ and $\Delta_K = \sum_{l=1}^{K}\hat{\nu}(\gamma_l)$.

For the proposed detector, one needs to perform $(I-K+1)M$ squared Euclidean distance calculations in (4.9). Hence, the proposed algorithm makes $N(I-K+1)M$ squared Euclidean distance calculations in total, which is considerably smaller than that of the optimum ML detector, i.e., $2^{\lfloor\log_2\binom{N}{K}\rfloor+\lfloor\log_2\binom{I-1}{K-1}\rfloor}M^K$. By assuming $K=\alpha N$, $I=\beta N$, $0<\alpha\leq 1$, and $\alpha\leq\beta$ and using the Stirling's approximation,[9] the ML and

[9] $N! \sim \sqrt{2\pi N}(N/e)^N$ when $N\to\infty$.

LC-ML detectors can be executed in $O((2^{H(\alpha)+\beta H(\alpha/\beta)}M^{\alpha})^{N})$ and $O(N^2M)$ operations, respectively, where

$$H(r) = -r\log_2 r - (1-r)\log_2(1-r) \tag{4.11}$$

is the binary entropy function.

Remark. IM and ICM have the same mapping and detection complexities for the same SE when we implement look-up tables during the transmission and use the optimum ML detectors at the receiver. This is because the size of the look-up tables and the number of squared Euclidean distance calculations are the same for both of the techniques in this case. However, the optimum ML detectors of these schemes become impractical for high SEs due to their excessive complexity. To overcome the high complexity of the optimum ML detectors, they employ low-complexity element-wise detectors. The low-complexity detector of IM, called LLR detector, performs M' squared Euclidean distance calculations per channel where M' is the size of the constellation on the activated channels [2]. On the other hand, the LC-ML detector of ICM performs $(I-K+1)M$ squared Euclidean distance calculations per channel. It is straightforward to show that the size of the constellation, M', should satisfy $M' = M\left[\binom{I-1}{K-1}\right]^{1/K}$ for IM to achieve the same SE as that of ICM when IM and ICM have the same block size, N, and the number of activated channels, K. Hence, the LLR detector of IM and LC-ML detector of ICM perform $M\left[\binom{I-1}{K-1}\right]^{1/K}$ and $(I-K+1)M$ squared Euclidean distance calculations per channel, respectively. For example, when $N = 4$, $K = 3$, $I = 5$, and $M = 8$, the LLR detector of IM and LC-ML detector of ICM perform 15 and 24 squared Euclidean distance calculations per channel, respectively. Thus, in this case, the detection complexity of ICM is slightly higher than that of IM.

Although we presented an element-wise low complexity detection algorithm for the proposed ICM concept, we still need to use look-up tables to map the signal

vectors to the information bits. The receiver complexity would still be high when the sizes of the look-up tables are large. Thus, it is essential to perform the mapping without implementing a look-up table. Fortunately, the structure of the index and composition encoding enables us to conduct such a mapping without implementing a look-up table.

We already know from [2] that the mapping of the information bits to the indices of the activated channels can be done through a combinatorial method. Similarly, one can map the information bits to energies of the codeword elements without using a look-up table. More explicitly, a weak composition of an integer can be represented by using dashes and vertical bars similar to a MacMahon graph of a composition [117]. To represent weak N-compositions of I, I dashes are used to signify the positive parts of the integer, whereas $N-1$ bars are used to indicate the addition operations, or in other words, the separations between the parts. To obtain the dash and bar representation of a composition, we write the dashes first. Then, we place the bars in a way that the ith space separated by these bars has μ_i dashes where μ_i is the ith part of the composition. For example, let us have a look at the weak 3-compositions of $I = 3$. In that case, one needs $I = 3$ dashes, i.e., $-\ -\ -$, and $N - 1 = 2$ bars since $I = 3$ can be composed of three positive parts at most, and there will be $N = 3$ parts separated by two bars. The weak compositions $3 = 1 + 1 + 1$ and $3 = 0 + 0 + 3$ can be represented as "$-|-|-$" and "$||--$", respectively. Hence, each weak composition can be represented by a combination of bars and dashes. After obtaining the dash and bar representation of a weak composition, the locations[10] of bars or dashes in the representation can be mapped to unique decimal numbers ranging from 0 to $\binom{I+N-1}{N-1} - 1$ by using the combinatorial method in [2]. Finally, one can convert the decimal numbers into binary bits.

[10]Note that there are $I+N-1$ locations in the dash and bar representations due to I dashes and $N-1$ bars. For "$-|-|-$" and "$||---$", the locations of the bars can be written as $\{2,4\}$ and $\{1,2\}$, respectively.

On the other hand, one can execute the same approach for mapping the strict compositions to binary bits since there is a one-to-one correspondence between compositions and weak compositions. This correspondence arises because we capture weak compositions of I with N parts by subtracting one from each component of strict compositions of $I + N$ with N summands.

4.4 Codebook Selection

The ICM concept can support a wide range of SE since it can alter several parameters such as N, I, K, and M to achieve a target SE. However, it may not be possible to find proper system parameters for the proposed WCM and CM techniques to attain the same SE as the conventional modulation, IM, and ICM techniques. That is particularly important to conduct fair comparisons among the proposed and benchmark schemes. Thus, we provide a codebook selection algorithm to yield a target SE with the WCM and CM methods.

As mentioned in the previous chapters, the optimum codebook selection criterion in a fading channel at high SNR is based on maximizing the minimum Hamming and product distances among codewords. This is equivalent to the well-known rank-determinant criterion of [27]. Hence, in the proposed schemes, the error performance at high SNR is limited by the codewords whose Hamming distance is minimum, i.e., the codewords whose difference matrix has the minimum rank. Note that the difference matrix can be defined as $(\mathbf{X}_i - \mathbf{X}_j)$ for the codeword pairs $\mathbf{x}_i$ and $\mathbf{x}_j$ where $\mathbf{X}_i = \text{diag}(\mathbf{x}_i)$ and $\mathbf{X}_j = \text{diag}(\mathbf{x}_j)$. For the proposed schemes, such a minimum distance is observed from unit rank codeword pairs related to the conventional modulation symbols. To alleviate the effect of these pairs in the codebook, we propose a practical algorithm that culls these pairs from the codebook until a target SE, η, is reached. The proposed algorithm is given in Algorithm 4. The algorithm starts by

Algorithm 4: Exclusion of minimum rank codeword pairs

Step 1: Form the codebook $\mathcal{C}$ where $\mathcal{C} \coloneqq \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{L_c}\}$;
Step 2: Form $L_c \times L_c$ matrix $\mathbf{Z}$ where $z_{ij} = \text{rank}(\mathbf{X}_i - \mathbf{X}_j)$, $i, j \in \{1, 2, \ldots, L_c\}$;
Step 3: Excluding minimum rank pairs;
while $L_c \neq 2^\eta, \forall \boldsymbol{x}_i \in \mathcal{C}$ **do**
  $\mathbf{v} = \text{nonzeros}(\mathbf{Z})$, $z_{min} = \min \mathbf{v}$;
  $\mathbf{r}_i = \text{find}(\mathbf{Z}(:, i) == z_{min})$, $r_i = |\mathbf{r}_i|$;
  $\breve{i} = \arg\max_i r_i$;
  $\mathcal{C} \leftarrow \mathcal{C} - \{\mathbf{x}_{\breve{i}}\}$;
  $L_c \leftarrow L_c - 1$;
  $\mathbf{Z}(:, \breve{i}) = [\ ]$, $\mathbf{Z}(\breve{i}, :) = [\ ]$;
end
return $\mathcal{C}$;

forming the codebook $\mathcal{C} \coloneqq \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{L_c}\}$ according to the WCM or CM methods where $\mathbf{x}_i$ is the ith, $i \in \{1, 2, \ldots, L_c\}$, codeword in the codebook and L_c is the size of the codebook. In the second step, the algorithm forms $L_c \times L_c$ matrix $\mathbf{Z}$ whose ith row and jth column element, z_{ij}, is the rank of the difference matrix $(\mathbf{X}_i - \mathbf{X}_j)$, thus the function "rank(.)" is used to measure the rank of the difference matrices. Then, in the third step, the algorithm culls the codewords that have the maximum number of minimum nonzero rank pairs one-by-one until a target SE, η, is reached, i.e., $L_c = 2^\eta$. In this step, the function "nonzeros(.)" is used to find nonzero elements of the matrix $\mathbf{Z}$ and assign these elements to a vector, $\mathbf{v}$, whereas the function "find(.)" is used to find row indices of elements in the ith column of $\mathbf{Z}$ that equal the minimum value of $\mathbf{v}$, i.e., z_{min}. Moreover, $|\mathbf{r}_i|$ stands for the cardinality of the vector $\mathbf{r}_i$. Then, we determine the codeword that has the maximum number of minimum rank pairs by $\breve{i} = \arg\max_i r_i$ and remove it from the codebook. Finally, we update the codebook size by removing one and adjust the matrix $\mathbf{Z}$ by eliminating the $\breve{i}$th row and column of $\mathbf{Z}$, i.e., $\mathbf{Z}(:, \breve{i}) = [\ ], \mathbf{Z}(\breve{i}, :) = [\ ]$.

4.5 Performance Analysis

4.5.1 Bit Error Rate

Once again, an upper-bound on the average BER is given by the well-known union bound as follows

$$P_b \leq \frac{1}{f2^f} \sum_{i=1}^{2^f} \sum_{j=1}^{2^f} P(\mathbf{X}_i \rightarrow \mathbf{X}_j) D(\mathbf{X}_i \rightarrow \mathbf{X}_j) \tag{4.12}$$

where $P(\mathbf{X}_i \rightarrow \mathbf{X}_j)$ is the PEP regarding the erroneous detection of $\mathbf{X}_i$ as $\mathbf{X}_j$ where $i \neq j$, $i, j \in \{1, \ldots, L_c\}$, $\mathbf{X}_i = \text{diag}(\mathbf{x}_i)$, $\mathbf{X}_j = \text{diag}(\mathbf{x}_j)$, and $D(\mathbf{X}_i \rightarrow \mathbf{X}_j)$ is the number of bits in error for the corresponding pairwise error event. One can use the same PEP expression as in (3.19) and substitute the codewords of the proposed WCM, CM, and ICM methods to obtain an upper-bound on the average BER.

Remark. The minimum Hamming distance between the sets that keep the energies of the codeword elements is two, just like the index symbols of IM. However, the proposed schemes can also send conventional modulation symbols that limit their diversity gain to one.

As a special case of ICM, one can use only the indices and compositions to embed information and obtain an additional diversity gain.[11] However, in this case, the transmitted symbol vectors should be carefully chosen to get the optimum error performance. One simple yet efficient strategy is choosing the transmitted symbols in a way that the angular difference between the different energy symbols is maximized. Hence, in this case, one can choose the signal on the kth activated channel as $x_k^p = \sqrt{\nu_k E_T / I} \exp\left(\frac{2j\nu_k \pi}{I-K+1}\right)$. As will be shown in the Numerical Results section, the error performance of this special case is promising for low SEs.

[11] For this special case, the SE is $\eta = \frac{\lfloor \log_2 \binom{N}{K} \rfloor + \lfloor \log_2 \binom{I-1}{K-1} \rfloor}{N}$.

4.5.2 Spectral Efficiency

For a fixed codeword length N, it is straightforward to show that a higher value of I results in a higher value of SE for the WCM and CM concepts. Moreover, when I is fixed, it is easy to see that the optimum value of N should satisfy $N \in \left\{\lfloor \frac{I-1}{2} \rfloor + 1, \lceil \frac{I-1}{2} \rceil + 1\right\}$ for the CM method. On the other hand, when I and N are fixed, there exists an optimum value of K that maximizes the SE of the ICM technique. By relaxing the constraints on the number of bits mapped to the channel and energy indices, i.e., removing the floor operations from the SE equation, defining the scaling factors α and β where $0 < \alpha \leq 1$ and $\alpha \leq \beta$, and rewriting K and I as $K = \alpha N$ and $I = \beta N$, respectively, one can formulate the SE maximization problem as

$$\begin{aligned} \underset{\alpha}{\text{maximize}} \quad & \eta = \frac{\log_2 \binom{N}{\alpha N} + \log_2 \binom{\beta N - 1}{\alpha N - 1} + \alpha N \log_2 M}{N} \\ \text{subject to} \quad & 0 < \alpha \leq 1, \\ & \alpha \leq \beta. \end{aligned} \tag{4.13}$$

Proposition 1. The value of α that satisfies (4.13) is

$$\alpha^* = \frac{M(\beta+1) - \sqrt{M^2(\beta-1)^2 + 4M\beta}}{2(M-1)}. \tag{4.14}$$

Thus, one can pick the optimum number of activated channels in an ICM block as

$$K^* \in \left\{\lfloor \alpha^* N \rfloor, \lceil \alpha^* N \rceil\right\}. \tag{4.15}$$

Proof. See Appendix 4.A.1. □

Proposition 2. As $N \to \infty$, the asymptotic SE of WCM, CM, and ICM can be given as

$$\eta_{WCM} \sim (\beta+1) H(1/(\beta+1)) + \lambda\beta, \tag{4.16}$$

$$\eta_{CM} \sim (\beta)H(1/\beta) + \log_2 M, \tag{4.17}$$

$$\eta_{ICM} \sim H(\alpha) + \log_2(\beta^\beta/\alpha^\alpha) - (\beta-\alpha)\log_2(\beta-\alpha) + \alpha \log_2 M. \tag{4.18}$$

Proof. The proof is simply obtained by applying the Stirling's formula to binomial coefficients in the SE expressions and ignoring the expressions that approach zero as $N \to \infty$. □

4.5.3 Mutual Information

In this subsection, we derive mutual information expressions for CM.[12] Since it is not easy to attain the capacity-achieving input distributions for the CM techniques, we assume that the codeword elements have complex Gaussian distribution as in the IM studies. To perform CM encoding over the parallel Gaussian channels, we take a unique approach and select the variance of each Gaussian channel according to the summands of compositions. We further assume that the compositions are randomly drawn from the set of possible compositions and calculate the optimum composition probabilities that maximize the mutual information.

Let us drop the block indices once again and collect the N received symbols in the vector $\mathbf{y} \coloneqq [y(1), y(2), \ldots, y(N)]$ for the transmit vector $\mathbf{x} \coloneqq [x_1, x_2, \ldots, x_N]$. In this case, one can write the mutual information as

$$\mathrm{I}(\mathbf{x};\mathbf{y}) = \mathbb{h}(\mathbf{y}) - \mathbb{h}(\mathbf{y}|\mathbf{x}) = \mathbb{h}(\mathbf{y}) - \mathbb{h}(\mathbf{w}) \tag{4.19}$$

where $\mathbb{h}(.)$ is the differential entropy operator, $\mathbf{w} \coloneqq [w(1), w(2), \ldots, w(N)]$, and $w(n) \sim \mathcal{CN}(0, N_0)$ is the additive Gaussian noise related to the nth codeword element.

On the other hand, we denote the input vector associated with the selected composition as $\mathbf{x}_c$. Such a vector is randomly distributed over all possible energy vectors

[12] Here, we conduct the mutual information analysis for only CM since we find it straightforward to extend the study in this section for WCM and ICM.

with probabilities $p_l \coloneqq \Pr(\mathbf{x}_c = \xi_l)$ where ξ_l is the energy vector related to the lth composition, and $l = 1, 2, \ldots, L_{CM}$. We further assume that x_n is distributed with complex Gaussian distribution, i.e., $x_n \sim \mathcal{CN}(0, E_{l,n})$ where $E_{l,n}$ is the nth element of ξ_l.

4.5.3.1 Exact Mutual Information

Using (4.19), we can write

$$\mathrm{I}(\mathbf{x};\mathbf{y}) = -\int f_{\mathbf{y}}(\mathbf{y}) \log_2 f_{\mathbf{y}}(\mathbf{y}) d\mathbf{y} - N \log_2(\pi e N_0). \tag{4.20}$$

By conditioning over each energy vector, ξ_l, the PDF of $\mathbf{y}$ can be written as

$$f_{\mathbf{y}}(\mathbf{y}) = \sum_{l=1}^{L_{CM}} p_l f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l) \tag{4.21}$$

where

$$f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l) = \prod_{n=1}^{N} \frac{e^{\frac{-|y(n)|^2}{\sigma_{l,n}^2}}}{\pi \sigma_{l,n}^2} \tag{4.22}$$

with $\sigma_{l,n}^2 = E_{l,n}|h(n)|^2 + N_0$. (4.22) follows from the independence among the channels and the fact that each channel has a complex normal distribution. To the best of the author's knowledge, (4.20) does not yield a closed-form solution. Moreover, it is not easy to calculate (4.20) numerically due to the N-fold integral.

On the other hand, the input space $\mathbf{x}$ consists of symbol space $\mathbf{x}_s$ and composition space $\mathbf{x}_c$. Considering this fact and using the well-known chain rule [118], one can write the mutual information as

$$\mathrm{I}(\mathbf{x};\mathbf{y}) = \mathrm{I}(\mathbf{x}_s, \mathbf{x}_c; \mathbf{y}) = \mathrm{I}(\mathbf{x}_s; \mathbf{y}|\mathbf{x}_c) + \mathrm{I}(\mathbf{x}_c; \mathbf{y}). \tag{4.23}$$

Averaging over $\mathbf{x}_c$, the first term in the right-hand side of (4.23) can be obtained as

$$\mathrm{I}(\mathbf{x}_s;\mathbf{y}|\mathbf{x}_c) = \sum_{l=1}^{L_{CM}} \Pr(\mathbf{x}_c = \xi_l)\mathrm{I}(\mathbf{x}_s;\mathbf{y}|\mathbf{x}_c = \xi_l) \quad = \sum_{l=1}^{L_{CM}} p_l \sum_{n=1}^{N} \log_2\left(1 + \frac{E_{l,n}|h(n)|^2}{N_0}\right). \tag{4.24}$$

Moreover, for the second term in the right-hand side of (4.23), we have the following relation

$$\begin{aligned} \mathrm{I}(\mathbf{x}_c;\mathbf{y}) &= \mathbb{H}(\mathbf{x}_c) - \mathbb{H}(\mathbf{x}_c|\mathbf{y}) \\ &= -\sum_{l=1}^{L_{CM}} p_l \log_2 p_l - \mathbb{H}(\mathbf{x}_c|\mathbf{y}). \end{aligned} \tag{4.25}$$

By using the definition of conditional entropy function, one can write

$$\begin{aligned} \mathbb{H}(\mathbf{x}_c|\mathbf{y}) &= \sum_{l=1}^{L_{CM}} \int_{\mathbf{y}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)\Pr(\mathbf{x}_c = \xi_l)\log_2\left(\frac{1}{\Pr(\mathbf{x}_c = \xi_l|\mathbf{y})}\right)d\mathbf{y} \\ &= \sum_{l=1}^{L_{CM}} p_l \int_{\mathbf{y}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)\log_2\left(\frac{\sum_{\hat{l}=1}^{L_{CM}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_{\hat{l}})}{f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)}\right)d\mathbf{y} \\ &= \sum_{l=1}^{L_{CM}} p_l \int_{\mathbf{y}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)\log_2\left(\frac{1}{f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)}\right)d\mathbf{y} \\ &\quad + \sum_{l=1}^{L_{CM}} p_l \int_{\mathbf{y}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)\log_2\left(\sum_{\hat{l}=1}^{L_{CM}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_{\hat{l}})\right)d\mathbf{y}. \end{aligned} \tag{4.26}$$

The N-fold integral in the first term of the last equation in (4.26) leads to the N-fold sum of the differential entropy related to the complex Gaussian random variables, i.e.,

$$-\sum_{l=1}^{L_{CM}} p_l \int_{\mathbf{y}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)\log_2 f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)d\mathbf{y} = \sum_{l=1}^{L_{CM}} p_l \sum_{n=1}^{N} \log_2(\pi e\sigma_{l,n}^2). \tag{4.27}$$

However, to the best of our knowledge, there is no closed-form solution for the second

term of the last equation in (4.26). Thus, by substituting the above expressions into (4.23), the mutual information of CM can be given by

$$\mathrm{I}(\mathbf{x};\mathbf{y}) = -N\log_2(\pi e N_0) - \sum_{l=1}^{L_{CM}} p_l \log_2 p_l - \sum_{l=1}^{L_{CM}} p_l \mathrm{E}_{(\mathbf{y}|\mathbf{x}_c=\xi_l)}\left[\log_2\left(\sum_{\hat{l}=1}^{L_{CM}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c=\xi_{\hat{l}})\right)\right]. \tag{4.28}$$

On the other hand, the corresponding optimization problem can be formulated as

$$\begin{aligned} \underset{\mathbf{p}}{\text{maximize}} \quad & \mathrm{I}(\mathbf{x};\mathbf{y}) \\ \text{subject to} \quad & \sum_{l=1}^{L_{CM}} p_l = 1, \\ & 0 \le p_l \le 1 \end{aligned} \tag{4.29}$$

where $\mathbf{p} \coloneqq (p_1, p_2, \ldots, p_{L_{CM}})$. Note that the exact mutual information is not a closed-form expression. Thus, it is not easy to find the optimum composition probabilities for the problem in (4.29). Although one can employ numerical optimization due to the convexity of the problem, high-dimensional numerical integration makes it impractical. In this regard, we provide closed-form expressions related to the mutual information through bounding and observing asymptotic behavior in the rest of this section.

4.5.3.2 A Lower-Bound

By using Jensen's inequality and considering the concavity of the logarithm function, the last term in (4.26) can be lower-bounded by

$$-\sum_{l=1}^{L_{CM}} p_l \mathrm{E}_{(\mathbf{y}|\mathbf{x}_c=\xi_l)}\left[\sum_{\hat{l}=1}^{L_{CM}} f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c=\xi_{\hat{l}})\right] \ge -\sum_{l=1}^{L_{CM}} p_l \log_2 \sum_{\hat{l}=1}^{L_{CM}} \frac{1}{\pi^N \det(\Xi_l + \Xi_{\hat{l}})} \tag{4.30}$$

where $\Xi_l = \text{diag}(\Sigma_l)$ is a diagonal matrix and $\Sigma_l \coloneqq (\sigma^2_{l,1}, \sigma^2_{l,2}, \ldots, \sigma^2_{l,N})$. By exploiting (4.30) in (4.28), we obtain the following lower-bound on the mutual information

$$\text{I}(\mathbf{x};\mathbf{y}) \geq -N\log_2(eN_0) - \sum_{l=1}^{L_{CM}} p_l \log_2 p_l - \sum_{l=1}^{L_{CM}} p_l \log_2 \sum_{\hat{l}=1}^{L_{CM}} \frac{1}{\det(\Xi_l + \Xi_{\hat{l}})}. \tag{4.31}$$

Proposition 3. The probabilities, p_l, $l = 1, \ldots, L_{CM}$, that maximize the lower-bound are given by

$$p_l^\star = \frac{\omega_l}{\sum_{l=1}^{L_{CM}} \omega_l} \tag{4.32}$$

where $\omega_l = \left[\sum_{\hat{l}=1}^{L_{CM}} \frac{1}{\det(\Xi_l+\Xi_{\hat{l}})}\right]^{-1}$.

Proof. See Appendix 4.A.2. □

4.5.3.3 An Upper-Bound

For ease of representation, we invoke a slight change of notation and define $f_l(\mathbf{y}) = f_{\mathbf{y}|\mathbf{x}_c}(\mathbf{y}|\mathbf{x}_c = \xi_l)$ and $f_\Sigma(\mathbf{y}) = \sum_{\hat{l}=1}^{L_{CM}} f_{\hat{l}}(\mathbf{y})$. Using these equations, one can rewrite the conditional entropy $\mathbb{H}(\mathbf{x}_c|\mathbf{y})$ in the following form

$$\mathbb{H}(\mathbf{x}_c|\mathbf{y}) = -\sum_{l=1}^{L_{CM}} p_l \mathbb{D}(f_l(\mathbf{y})||f_\Sigma(\mathbf{y})) \tag{4.33}$$

where $\mathbb{D}(f_l(\mathbf{y})||f_\Sigma(\mathbf{y})) = \int_{\mathbf{y}} f_l(\mathbf{y}) \log_2 \frac{f_l(\mathbf{y})}{f_\Sigma(\mathbf{y})} d\mathbf{y}$ is the relative entropy or Kullback-Leibler (KL) distance between the PDFs $f_l(\mathbf{y})$ and $f_\Sigma(\mathbf{y})$. By using variational approximation, the KL distance can be upper-bounded by [90, 119]

$$\mathbb{D}(f_l(\mathbf{y})||f_\Sigma(\mathbf{y})) \leq -\log_2 \sum_{\hat{l}=1}^{L_{CM}} e^{-\mathbb{D}(f_l(\mathbf{y})||f_{\hat{l}}(\mathbf{y}))}. \tag{4.34}$$

Since both $f_l(\mathbf{y})$ and $f_{\hat{l}}(\mathbf{y})$ are multivariate normal distribution functions, i.e., $(\mathbf{y}|\mathbf{x}_c = \xi_l) \sim \mathcal{N}(0, \Xi_l)$, the KL distance between these density functions is given by

$$\mathbb{D}(f_l(\mathbf{y})||f_{\hat{l}}(\mathbf{y})) = \ln \frac{\det(\Xi_{\hat{l}})}{\det(\Xi_l)} + \mathbf{Tr}(\Xi_{\hat{l}}^{-1}\Xi_l) - N \tag{4.35}$$

where $\mathbf{Tr}(\mathbf{A})$ is the trace of matrix $\mathbf{A}$. By using the upper-bound in (4.34), one can write

$$\mathbb{H}(\mathbf{x}_c|\mathbf{y}) \geq \sum_{l=1}^{L_{CM}} p_l \log_2 \sum_{\hat{l}=1}^{L_{CM}} e^{-\mathbb{D}(f_l(\mathbf{y})||f_{\hat{l}}(\mathbf{y}))}. \tag{4.36}$$

Moreover, by using this inequality, we obtain the following upper-bound

$$\mathrm{I}(\mathbf{x};\mathbf{y}) \leq \sum_{l=1}^{L_{CM}} p_l \sum_{n=1}^{N} \log_2\left(1+\frac{E_{l,n}|h(n)|^2}{N_0}\right) - \sum_{l=1}^{L_{CM}} p_l \log_2 p_l - \sum_{l=1}^{L_{CM}} p_l \log_2 \sum_{\hat{l}=1}^{L_{CM}} e^{-\mathbb{D}(f_l(\mathbf{y})||f_{\hat{l}}(\mathbf{y}))}. \tag{4.37}$$

Proposition 4. The probabilities, p_l, $l = 1, \ldots, L_{CM}$, that maximizes the upper-bound in (4.37) are given by

$$p_l^{\star} = \frac{\kappa_l}{\sum_{l=1}^{L_{CM}} \kappa_l}. \tag{4.38}$$

where $\kappa_l = \left[\sum_{\hat{l}=1}^{L_{CM}} e^{-D(f_l(\mathbf{y})|f_{\hat{l}}(\mathbf{y}))}\right]^{-1}$.

Proof. Proof can be obtained by following the same steps in the proof of Prop. 3. □

4.5.3.4 Closed-Form Asymptotics

Let us write the mutual information explicitly as follows

$$\mathrm{I}(\mathbf{x};\mathbf{y}) = -\sum_{l=1}^{L_{CM}} p_l \int \prod_{n=1}^{N} dy(n)dy^*(n)\frac{e^{-\frac{|y(n)|^2}{\sigma_{l,n}^2}}}{\pi\sigma_{l,n}^2}\log_2\left[\sum_{\hat{l}=1}^{L_{CM}} p_{\hat{l}}\prod_{\hat{n}=1}^{N}\frac{e^{-\frac{|y(\hat{n})|^2}{\sigma_{\hat{l},\hat{n}}^2}}}{\pi\sigma_{\hat{l},\hat{n}}^2}\right] - N\log_2(\pi e N_0). \tag{4.39}$$

For the integral related to the lth term, we exploit the change of variables through $y(n) \to \sigma_{l,n}y(n)$ and $y^*(n) \to \sigma_{l,n}y^*(n)$ and rewrite (4.39) as

$$\begin{aligned}\mathrm{I}(\mathbf{x};\mathbf{y}) = &- N\log_2(\pi e N_0) - \sum_{l=1}^{L_{CM}} p_l \int \prod_{n=1}^{N} dy(n)dy^*(n)\frac{e^{-|y(n)|^2}}{\pi} \\ &\times \log_2\left[\sum_{\hat{l}=1}^{L_{CM}}\frac{p_{\hat{l}}}{\prod_{\hat{n}=1}^{N}\pi\sigma_{\hat{l},\hat{n}}^2}\prod_{\hat{n}\in\Gamma_{l,\hat{l}}} e^{-|y(\hat{n})|^2}\prod_{\hat{n}\notin\Gamma_{l,\hat{l}}} e^{-\frac{\sigma_{l,\hat{n}}^2|y(\hat{n})|^2}{\sigma_{\hat{l},\hat{n}}^2}}\right]\end{aligned} \tag{4.40}$$

where $\Gamma_{l,\hat{l}}$ is the set of indices related to the elements in ξ_l and $\xi_{\hat{l}}$ that are equal. By factoring out $\frac{1}{(\pi N_0)^N}\prod_{\hat{n}=1}^{N} e^{-|y(\hat{n})|^2}$ in the argument of logarithm and evaluating the Gaussian integral, we obtain

$$\begin{aligned}\mathrm{I}(\mathbf{x};\mathbf{y}) = &-\sum_{l=1}^{L_{CM}} p_l \int \prod_{n=1}^{N} dy(n)dy^*(n)\frac{e^{-|y(n)|^2}}{\pi} \\ &\times \log_2\left[\sum_{\hat{l}=1}^{L_{CM}}\frac{p_{\hat{l}}}{\prod_{\hat{n}=1}^{N}\left(E_{\hat{l},\hat{n}}|h(\hat{n})|^2/N_0+1\right)}\prod_{\hat{n}\notin\Gamma_{l,\hat{l}}} e^{\left(\frac{E_{\hat{l},\hat{n}}-E_{l,\hat{n}}}{\sigma_{\hat{l},\hat{n}}^2}\right)|h(\hat{n})|^2|y(\hat{n})|^2}\right].\end{aligned} \tag{4.41}$$

Although (4.41) does not yield a closed-form expression at high SNR, i.e., $N_0 \to 0$, one can observe the following asymptotic expression from (4.41) at low SNR, i.e., $N_0 \to \infty$,

$$\mathrm{I}(\mathbf{x};\mathbf{y}) \sim -\log_2\left[\sum_{l=1}^{L_{CM}}\frac{p_l}{\prod_{n=1}^{N}\left(E_{l,n}|h(n)|^2/N_0+1\right)}\right]. \tag{4.42}$$

Proposition 5. By denoting $l' = \arg\max_l \sum_{n=1}^{N} \log_2 \left(\frac{E_{l,n}|h(n)|^2}{N_0} + 1 \right)$, the probabilities

$$p_l = \begin{cases} 1, & \text{if } l = l', \\ 0, & \text{otherwise,} \end{cases} \tag{4.43}$$

maximizes the asymptotic mutual information in (4.42). Moreover, the asymptotic relation obtained by choosing p_l as in (4.43) is equivalent to the well-known capacity of N parallel channels.

Proof. See Appendix 4.A.3. □

Remark. In this section, we derived the optimum composition probabilities that maximize the closed-form expressions related to the mutual information. However, these probabilities are usually not feasible since we map uniformly distributed binary bits to the compositions. That introduces a constraint on the composition probabilities. Although a fixed-length encoding scheme fails to cover all of the compositions since the number of compositions is usually not a power of two, one can exploit binary-tree encoding as in [80] and utilize all of the compositions by employing Huffman encoding in a reverse manner. In this case, the composition probabilities should follow a dyadic distribution, i.e., $p_l = 2^{-\beta_l}$ where $\beta_l \in \mathbb{N}$, $\sum_{l=1}^{L_{CM}} 2^{-\beta_l} = 1$, and $\mathbb{N}$ is the set of natural numbers. Therefore, it is necessary to map the optimum composition probabilities to the dyadic distributions. One can undertake this task by employing Algorithm 2 in [80].

4.6 Numerical Results

In this section, we provide numerical BER (as a function of SNR per channel), SE, and mutual information results. In BER figures, "WCM (N, I, λ), Alg. 4" stands for the proposed WCM scheme that employs Algorithm 4 and uses λ to alter the SE

as well as weak compositions of I with N parts to decide the energies of N codeword elements, whereas "CM (N, I, M)" stands for the proposed CM method that determines the energies of N codeword elements according to the composition of an integer I with N parts and carries M-PSK symbols on each element. Moreover, "ICM (N, K, I, M)" is the proposed ICM mechanism having K activated channels out of N in each block, choosing the energies associated with the activated channels according to the compositions of I into K parts, and employing M-PSK modulation on them, whereas "ICM (N, K, I)" stands for the special case of ICM that embeds information into only indices of codeword elements and compositions of an integer I with K parts. "IM (N, K, M)" signifies the conventional IM technique in which K out of N codeword elements are activated to send M-PSK symbols in each block. For the simulated schemes in this section, we pick $E_T = N$, thus the average energy per channel is assumed to be one.

4.6.1 Bit Error Rate

Fig. 4.6 compares the BER performance of WCM $(4, 4, 1)$, Alg. 4 and CM $(4, 7, 2)$ with IM $(4, 3, 4)$ and QPSK. Here, the SEs of all schemes are equal to 2 bpcu. The WCM $(4, 4, 1)$ concept is capable of achieving $\frac{\lfloor \log_2 \binom{I+N-1}{N-1} \rfloor + \lambda I}{N} = \frac{\lfloor \log_2 \binom{7}{3} \rfloor + 4}{4} = 2.25$ bpcu; however, we cull half of the WCM $(4, 4, 1)$ codewords with Algorithm 4 to achieve the same SE as those of the remaining concepts. Hence, the remaining codebook, WCM $(4, 4, 1)$, Alg. 4, exhibits an SE equal to 2 bpcu. As seen from the figure, CM $(4, 7, 2)$ can achieve a better BER performance than QPSK by providing almost the same BER results as IM $(4, 3, 4)$ at high SNR. On the other hand, WCM $(4, 4, 1)$, Alg. 4 considerably outperforms all other schemes, especially at high SNR. Here, the superiority of the WCM and CM concepts comes from their ability to produce more codewords whose minimum Hamming distance is two. That enables the proposed concepts to achieve the same SE as the benchmark methods by employing a lower

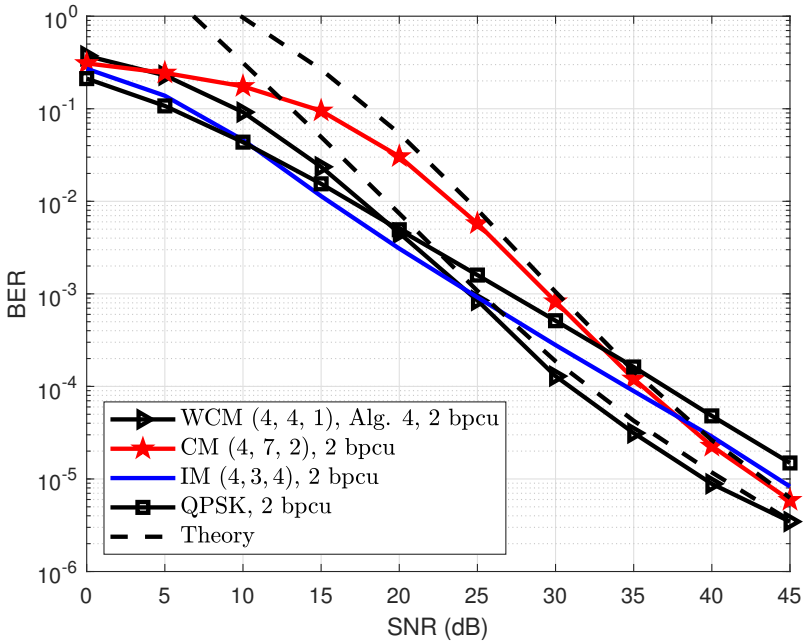


Figure 4.6. BER comparison of WCM $(4, 4, 1)$, Alg. 4 and CM $(4, 7, 2)$ with IM $(4, 3, 4)$ and QPSK.

order modulation and attaining a higher MED between the modulation symbols. In Fig. 4.6, we also compare the theoretical upper-bound results obtained by (4.12) with the computer simulation results. The theoretical results match with the simulation results at especially high SNR.

In Fig. 4.7, we compare the BER performance of WCM $(4, 6, 1)$, Alg. 4, CM $(4, 6, 4)$, and CM $(4, 12, 2)$ with IM $(4, 3, 8)$. The SEs of all systems are 2.75 bpcu.[13] As seen from the figure, the proposed schemes outperform the IM method, especially at high SNR. Moreover, WCM $(4, 6, 1)$, Alg. 4, CM $(4, 6, 4)$, and CM $(4, 12, 2)$ provide almost 3.2 dB, 3.7 dB, and 4.5 dB SNR gains, respectively, compared to IM $(4, 3, 8)$ at a BER value of 10^{-5}.

In Fig. 4.8, we compare the BER performance of ICM with that of WCM, CM, IM, and BPSK. Except for BPSK, all systems achieve the same SE, which is 1.5 bpcu.

[13]Here, a block of WCM $(4, 6, 1)$ is capable of transmitting $\lfloor \log_2 \binom{6+4-1}{4-1} \rfloor + 6 = 12$ bits; however, we cull half of the WCM $(4, 6, 1)$ codebook by using Algorithm 4 to achieve the same SE as a block of IM $(4, 3, 8)$ that is capable of sending 11 bits.

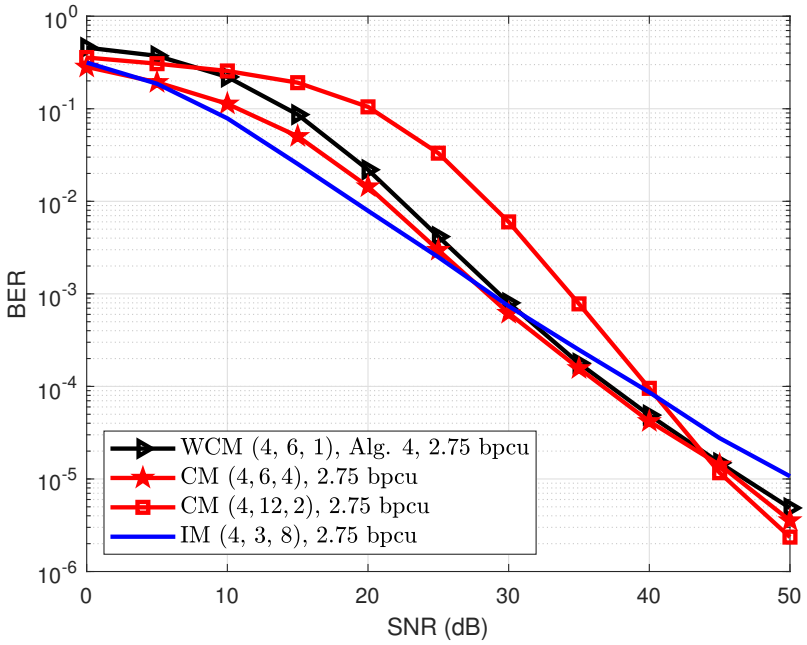


Figure 4.7. BER comparison of WCM $(4, 6, 1)$, Alg. 4, CM $(4, 6, 4)$, and CM $(4, 12, 2)$ with IM $(4, 3, 8)$.

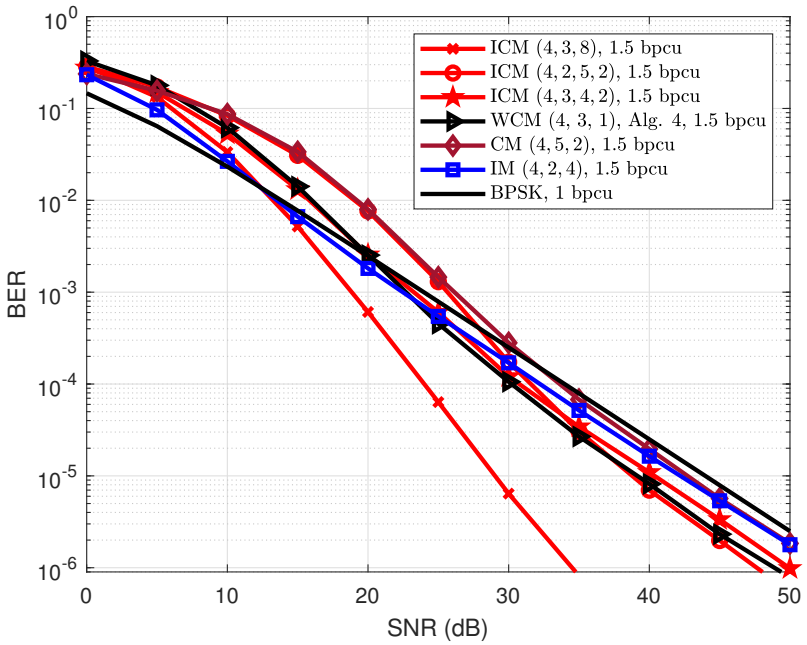


Figure 4.8. BER comparison of ICM with WCM, CM, IM, and M-PSK when $N = 4$, $K \in \{2, 3\}$, $I \in \{3, 4, 5, 8\}$, $M \in \{2, 4\}$, and $\lambda = 1$.

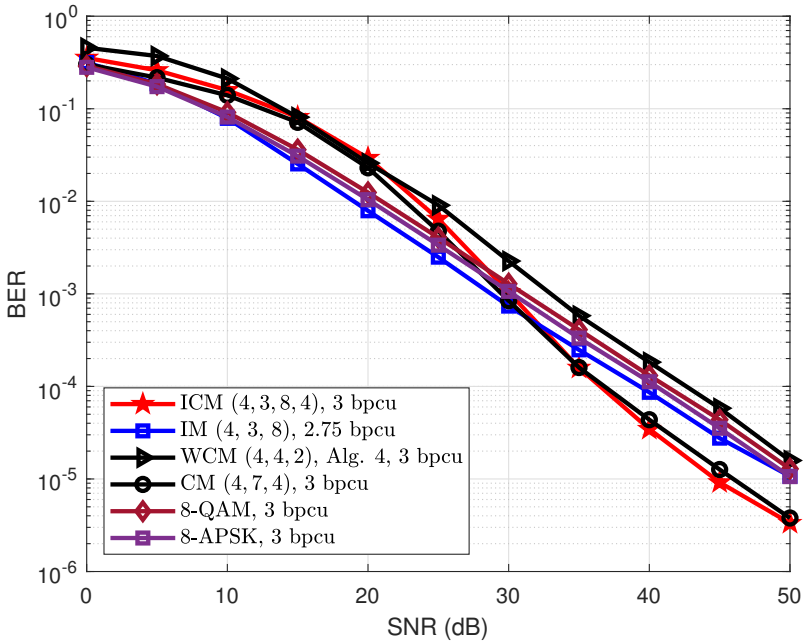


Figure 4.9. BER comparison of ICM with IM, WCM, CM, QAM, and APSK when $N = 4$, $K = 3$, $I \in \{4, 7, 8\}$, $M \in \{4, 8\}$, and $\lambda = 2$.

As seen from the figure, ICM $(4, 3, 8)$ considerably outperforms all other techniques by introducing an additional diversity gain. ICM $(4, 3, 8)$ is a special case of ICM in which we use only the indices and compositions to embed information. Since the minimum Hamming distance among the associated codewords is two, ICM $(4, 3, 8)$ achieves a modulation diversity of two.[14] Moreover, the BER performance of ICM $(4, 3, 4, 2)$ is very close to that of WCM $(4, 3, 1)$, Alg. 4, and these methods outperform the CM, IM, and BPSK concepts, especially at high SNR, since they produce more symbols whose minimum Hamming distance is two. However, since the MED for the overall codebook of ICM and WCM is less than that of IM, they are outperformed by IM at low SNR. Due to the same reason, ICM $(4, 2, 5, 2)$ is outperformed by all other schemes except for CM $(4, 5, 2)$ at low SNR; however, it outperforms all other methods except for ICM $(4, 3, 8)$ at high SNR.

[14]Here, since we consider 4-dimensional codewords, i.e., $N = 4$, the maximum value of modulation diversity is four. As mentioned in Section 2.3.3, one can obtain the maximum modulation diversity by using complex algebraic number fields or rotating the multidimensional codewords.

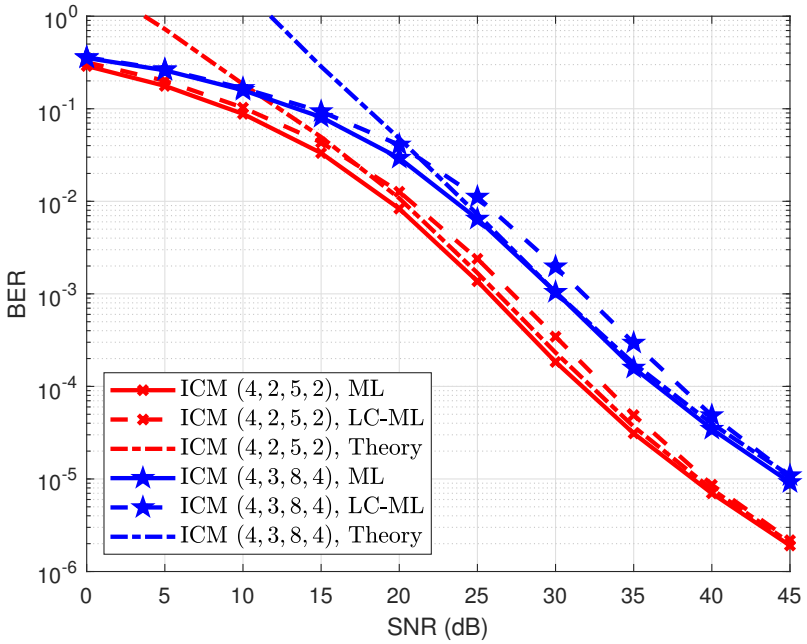
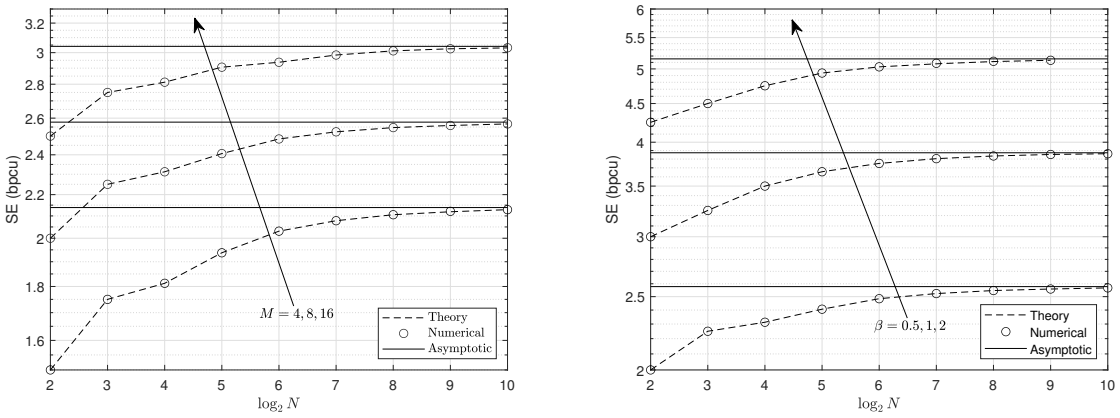


Figure 4.10. BER comparison of the proposed LC-ML detector with the optimum ML detector for ICM.

Fig. 4.9 compares the BER performance of ICM with that of IM, WCM, CM, QAM, and amplitude and phase shift keying (APSK).[15] As observed from the figure, ICM $(4, 3, 8, 4)$ and CM $(4, 7, 4)$ exhibit very similar BER performance, and they outperform all other techniques, especially at high SNR. Compared to Fig. 4.8, the effectiveness of the WCM scheme against the other schemes substantially degrades as it requires deploying higher-order modulation to achieve 3 bpcu. However, the ICM scheme preserves its effectiveness in both data rates. It is also worth mentioning that ICM $(4, 3, 8, 4)$ considerably outperforms the 8-QAM and 8-APSK techniques that use the energies of the constellation symbols for encoding as in the proposed encoding mechanisms.

Fig. 4.10 demonstrates the BER performance of the proposed LC-ML detector compared to the optimum ML detector for the ICM concept when $N = 4$, $K \in \{2, 3\}$,

[15]For the curve related to 8-APSK, we use the optimum 8-ary APSK design in [120, 121]. In this design, the constellation consists of two circles (inner and outer) with four signals on each circle. The radii of the inner and outer circles are $r_1 = 1/\sqrt{2} \simeq 0.707$ and $r_2 = 1.366$, respectively.



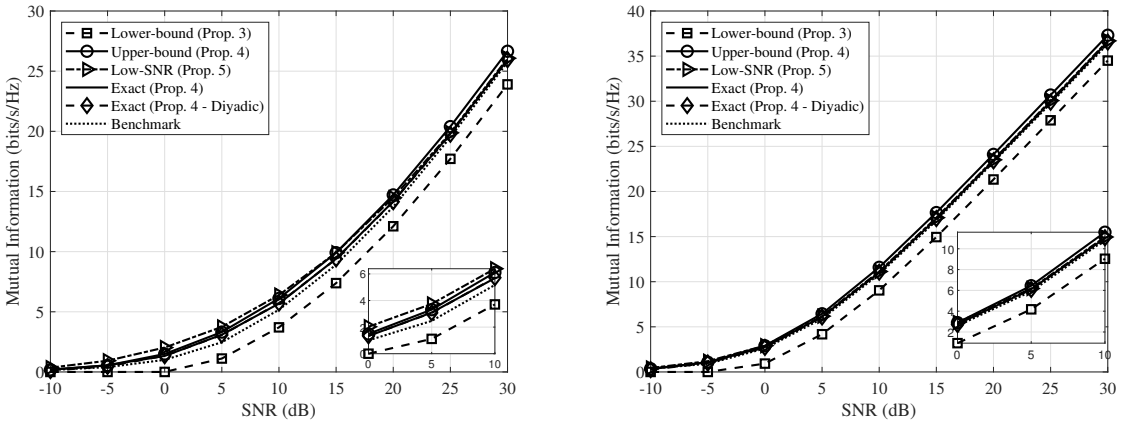
(a) $\beta = 0.5$ and $M \in \{4, 8, 16\}$ (b) $\beta \in \{0.5, 1, 2\}$ and $M = 8$

Figure 4.11. Comparison of the theoretical, numerical, and asymptotic SE results for ICM. Here, the curves with extensions "Theory" in the legend are obtained by substituting (4.15) into (4.6), whereas those with extensions "Numerical" in the legend are obtained by substituting all possible values of K into (4.6) and choosing the maximum SE. On the other hand, the curves labeled "Asymptotic" in the legend are obtained by substituting (4.14) into (4.18).

$I \in \{5, 8\}$, and $M \in \{2, 4\}$. As seen from the figure, the performance of the proposed LC-ML detector is very close to that of the optimum ML detector, especially at low and high SNR. On the other hand, the curves with extensions "Theory" in the legend relate to the theoretical upper-bound results for the proposed ICM scheme. The upper-bound curves are also consistent with the computer simulations results.

4.6.2 Spectral Efficiency

In Fig. 4.11, we compare theoretical, numerical, and asymptotic SE results of ICM for two cases: (1) $\beta = 0.5$ and $M \in \{4, 8, 16\}$ in Fig. 4.11a; (2) $\beta \in \{0.5, 1, 2\}$ and $M = 8$ in Fig. 4.11b. Fig. 4.11 verifies the validity of the SE expressions in Section 4.5.2 since the theoretical and numerical SE results match perfectly, and they approach the asymptotic SE as N increases. On the other hand, the SE is considerably improved when we increase N. By looking at Fig. 4.11a when $N = 4$, the ICM schemes can achieve only 70.14%, 77.60%, and 82.20% of the asymptotic SE for $M = 2, 4$, and 8,



(a) $\delta = 0.2$ (b) $\delta = 0.7$

Figure 4.12. Mutual information vs. SNR for $N = 4$, $I = 9$, $\delta = 0.2$, and $\delta = 0.7$ where $|h(n)|^2 = \delta^n$ for $n = 1, 2, \ldots, N$.

respectively, whereas these percentages increase to 81.82%, 87.30%, and 90.41% for $M = 2, 4$, and 8, respectively, when $N = 8$. We observe a similar trend from Fig. 4.11b.

4.6.3 Mutual Information

In Fig. 4.12, we compare the mutual information expressions derived in Section 4.5.3. For the lower-bound, upper-bound, and low-SNR expressions, we consider the respective optimum probabilities in Propositions 3, 4, and 5, respectively. We obtain the curves with extensions "Exact (Prop. 4)" by substituting the optimum upper-bound probabilities of Prop. 4 into the exact mutual information expression. Moreover, we utilize binary-tree encoding and map the probabilities in Prop. 4 to dyadic probabilities according to Algorithm 2 in [80] for the curves with extensions "Exact (Prop. 4 - Dyadic)". Also, the "Benchmark" curves are related to the exact mutual information of CM exploiting fixed-length encoding and having uniform composition probabilities.[16] As seen from Figs. 4.12a and 4.12b, except for the lower-bound, the derived

[16]Here, we evaluate the N-dimensional integral in the exact mutual information expression using a Monte Carlo estimator.

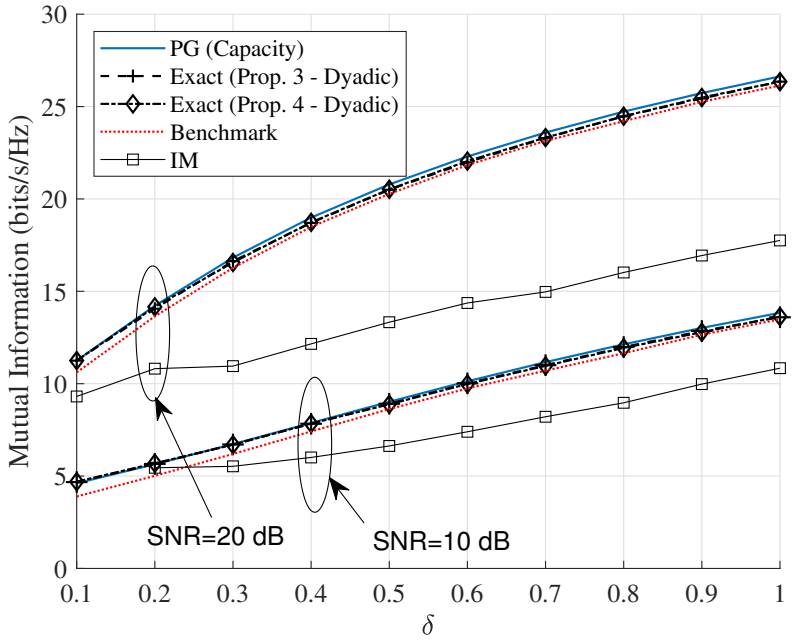


Figure 4.13. Comparison of mutual information for CM, parallel Gaussian channels, and IM when $N = 4$.

expressions are tight as they are well aligned with the exact results at all SNR. On the other hand, exploiting binary-tree encoding as in [80] improves the mutual information notably compared to the benchmark at medium-to-high SNR when the channel conditions are more variable, i.e., $\delta = 0.2$. However, the advantage offered by binary-tree encoding vanishes when the channel conditions are less variable, i.e., $\delta = 0.7$.

In Fig. 4.13, we compare CM with parallel Gaussian channels and IM in mutual information. Here, the mutual information results are provided as a function of δ for $N = 4$ and SNR $\in \{10 \text{ dB}, 20 \text{ dB}\}$. The curves labeled "PG (Capacity)" are related to the capacity of parallel Gaussian channels, whereas the ones labeled "Exact (Prop. 3 - Dyadic)" and "Exact (Prop. 4 - Dyadic)" give the exact mutual information of CM over the equivalent dyadic versions of the probabilities in Propositions 3 and 4, respectively. We obtain the curves labeled "Benchmark" in the same way as in the previous figure. For the curves related to the CM concept, including the benchmark

curves, we have $N = 4$ and $I = 11$. Moreover, "IM" stands for the IM technique having $K = 2$ activated channels out of $N = 4$, and the corresponding curves are obtained by exploiting Prop. 6 and Algorithm 2 in [80]. As seen from the figure, utilizing the optimum probabilities in Propositions 3 and 4 results in almost the same mutual information for all values of δ and SNR. The resulting mutual information values are noticeably higher than those of the benchmark CM scheme when especially $\delta \leq 0.4$ and SNR is 10 dB. More importantly, the mutual information of the CM technique is considerably higher than that of the IM method and only slightly less than the capacity of parallel Gaussian channels.

4.7 Summary

Unlike the previous chapter that uncovers the novel application of set partitioning in encoding, this chapter discovered novel encoding methods based on using compositions. In this regard, we proposed novel modulation concepts, which we call WCM, CM, and ICM. These concepts utilize strict or weak compositions of a positive integer to embed additional information into the energies of the activated codeword elements. This approach is completely different from the SPM encoding since we use disjoint constellations to form the SPM codewords, whereas we use the transmit energy levels to construct the CM symbols. For the proposed modulation concepts, we represented practical implementations in parallel channels. We analyzed the performance in terms of BER, SE, and mutual information.[17] We showed through computer simulations and theoretical calculations that the proposed schemes can provide noteworthy improvements compared to conventional modulation and IM. Our findings show also that the use of compositions on top of the classical and IM encoding has the potential to improve the performance.

[17]Unlike the other concepts of this thesis, we were also able to investigate the mutual information of the CM concepts thanks to the composition encoding through the energy levels.

4.A Appendices

4.A.1 Proof of Proposition 1

For $0 < \alpha_1 < \alpha_2 \leq 1$, one can write

$$\begin{aligned}\frac{\Delta\eta}{\Delta\alpha} = &\frac{\log_2 \binom{N}{\alpha_2 N} + \log_2 \binom{\beta N-1}{\alpha_2 N-1} + \alpha_2 N \log_2 M}{N(\alpha_2 - \alpha_1)} \\ &- \frac{\log_2 \binom{N}{\alpha_1 N} + \log_2 \binom{\beta N-1}{\alpha_1 N-1} + \alpha_1 N \log_2 M}{N(\alpha_2 - \alpha_1)}.\end{aligned} \tag{4.44}$$

Since $K \in \mathbb{Z}^+$, the minimum value of $\Delta\alpha = \alpha_2 - \alpha_1$ is $\Delta\alpha = \frac{1}{N}$. It is clear that the optimal value of α can be obtained when $\frac{\Delta\eta}{\Delta\alpha} = 0$. Therefore, by setting right-hand side of the equation equal to zero and substituting $\alpha_2 = \alpha_1 + \frac{1}{N}$ into it, we obtain the following equality

$$\log_2 \left(\frac{N - \alpha_1 N}{\alpha_1 N}\right) + \log_2 \left(\frac{\beta N - \alpha_1 N}{\alpha_1 N}\right) + \log_2 M = 0. \tag{4.45}$$

By arranging this equation, one can obtain the following classical quadratic program

$$(M-1)N^2\alpha_1^2 - M(1+\beta)N^2\alpha_1 + M\beta N^2 = 0. \tag{4.46}$$

This classical program yields the following two solutions

$$\alpha^* = \frac{M(1+\beta) \pm \sqrt{M^2(1+\beta)^2 - 4M(M-1)\beta}}{2(M-1)}. \tag{4.47}$$

However, one can write

$$\begin{aligned}\alpha^* &= \frac{M(1+\beta) + \sqrt{M^2(1+\beta)^2 - 4M(M-1)\beta}}{2(M-1)} \\ &= \frac{M(1+\beta) + \sqrt{M^2(\beta-1)^2 + 4M\beta}}{2(M-1)}\end{aligned} \tag{4.48}$$

$$> \frac{M(\beta+1)+M|\beta-1|}{2M} = \frac{(\beta+1)+|\beta-1|}{2}.$$

We observe from the last line of (4.48) that $\alpha^* > 1$ when $0 < \beta < 1$ and $\alpha^* > \beta$ when $\beta > 1$. Thus, (4.48) is not a valid solution for the problem in (4.13) as it does not satisfy the constraints. On the other hand, for the other solution, the following inequality holds

$$\begin{aligned} \alpha^* &= \frac{M(1+\beta)-\sqrt{M^2(\beta-1)^2+4M\beta}}{2(M-1)} \\ &< \frac{M(1+\beta)-|M(\beta-1)-2\beta|}{2(M-1)} = \beta \end{aligned} \tag{4.49}$$

when $0 < \beta < 1$.[18] Moreover, when $\beta \geq 1$, we have

$$\begin{aligned} \alpha^* &= \frac{M(1+\beta)-\sqrt{M^2(\beta-1)^2+4M\beta}}{2(M-1)} \\ &< \frac{M(1+\beta)-|M(\beta-1)+2|}{2(M-1)} = 1. \end{aligned} \tag{4.50}$$

Therefore, this is a valid solution of (4.13).

4.A.2 Proof of Proposition 3

The Lagrange function of the lower-bound maximization problem can be formulated as

$$\mathcal{L}(p_l,\mu) = -N\log_2(eN_0) - \sum_{l=1}^{L_{CM}} p_l \left[\log_2 p_l + \log_2 \sum_{\hat{l}=1}^{L_{CM}} \frac{1}{\det(\Xi_l+\Xi_{\hat{l}})} \right] - \mu\left(\sum_{l=1}^{L_{CM}} p_l - 1 \right). \tag{4.51}$$

[18] One can check that $|M(\beta-1)-2\beta|^2 < M^2(\beta-1)^2+4M\beta$ when $0 < \beta < 1$.

By taking the partial derivative with respect to p_l, we obtain the following set of equations

$$-\log_2 p_l - 1/\ln 2 - \log_2 \sum_{\hat{l}=1}^{L_{CM}} \frac{1}{\det(\Xi_l + \Xi_{\hat{l}})} - \mu = 0. \tag{4.52}$$

By solving (4.52), we have

$$p_l = \frac{2^{-\mu}\omega_l}{e} \tag{4.53}$$

where $\omega_l = \left[\sum_{\hat{l}=1}^{L_{CM}} \frac{1}{\det(\Xi_l+\Xi_{\hat{l}})}\right]^{-1}$. Since we have $\sum_{l=1}^{L_{CM}} p_l = 1$, the optimum value of p_l can be expressed as

$$p_l^{\star} = \frac{\omega_l}{\sum_{l=1}^{L_{CM}} \omega_l}. \tag{4.54}$$

4.A.3 Proof of Proposition 5

The following inequality holds for the asymptotic mutual information in (4.42)

$$\begin{aligned} \mathrm{I}(\mathbf{x};\mathbf{y}) &\sim -\log_2 \left[\sum_{l=1}^{L_{CM}} \frac{p_l}{\prod_{n=1}^{N}\left(\frac{E_{l,n}|h(n)|^2}{N_0}+1\right)}\right] \\ &\leq -\log_2 \left[\frac{\sum_{l=1}^{L_{CM}} p_l}{\prod_{n=1}^{N}\left(\frac{E_{l',n}|h(n)|^2}{N_0}+1\right)}\right] = \log_2 \left[\prod_{n=1}^{N}\left(\frac{E_{l',n}|h(n)|^2}{N_0}+1\right)\right] \\ &= \sum_{n=1}^{N} \log_2 \left[\frac{E_{l',n}|h(n)|^2}{N_0}+1\right]. \end{aligned} \tag{4.55}$$

The inequality is achieved when we choose the probabilities according to (4.43). Notice that the resulting mutual information is equivalent to the well-known capacity of N parallel channels.

Chapter 5

Encoding Using an MDS Code

We showed in the previous chapters that the combinatorial modulation techniques have beneficial properties thanks to the presence of index symbols, i.e., the codewords formed by combinatorial tools such as permutation, set partition, and composition. As we also presented, index symbols have a higher minimum Hamming distance between their pairs than conventional modulation symbols.[1] Although the overall diversity gain is unity for the combinatorial methods, encoding in both index and classical signal domains builds an attractive structure where we can transmit some bits more reliably, i.e., unequal error protection. Moreover, the presence of index symbols enables a higher MED in the signal domain. For instance, classical IM activates only a fraction of codeword elements and distributes the total energy among the activated ones. DM-IM and MM-IM use disjoint constellations on all codeword elements to form the index symbols,[2] and these techniques achieve the same SE as the conventional modulation methods by employing lower order modulation in the classical signal domain.

On the other hand, in Chapters 3 and 4, we aimed to exploit the potential of

[1]The minimum Hamming distance between classical modulation symbols is limited to one, whereas this is two for index symbols.

[2]Note that one can obtain disjoint constellations by set partitioning of well-known two-dimensional lattices as in trellis-coded modulation.

the index domain more and proposed novel modulation concepts that are capable of producing more index symbols and outperforming the existing combinatorial and classical methods Nevertheless, the existing and proposed combinatorial methods cannot achieve the full potential of the index domain since they do not produce the highest number of possible index codewords, i.e., the highest number of symbols whose minimum Hamming distance is two.

Against this background, we propose novel encoding methods based on a simple MDS code which is capable of exploiting the full potential of the index domain. We use the MDS code to obtain the highest number of codewords whose minimum Hamming distance is two just like the index symbols of the IM benchmarks. Moreover, we aim to show that one can use a simple MDS code and obtain practical, flexible, and low-complexity modulation concepts that are capable of outperforming combinatorial techniques. In this regard, we exploit amplitude and phase levels as well as in-phase and quadrature components of codeword elements to form N-tuples of the MDS code. Our novel contributions can be summarized as follows:

- We propose novel modulation concepts that we call MDS amplitude and phase modulation (MDS-APM) and MDS in-phase and quadrature modulation (MDS-IQM). We exploit the amplitude and phase levels to form the tuples of a simple MDS code for MDS-APM, and in-phase and quadrature components for MDS-IQM.

- We depict parallel channel implementations of the proposed schemes in which we utilize well-known PSK, QAM, and PAM constellations to construct the MDS structure. We further show that the codebooks of these implementations are capable of encompassing those of the MM-IM scheme as a special case, making the proposed methods important benchmarks for the IM techniques. Thus, they are more general classes of channel encoding which provide high SE and satisfactory error performance.

- For the proposed schemes, we present a sub-optimal ML detection mechanism that is capable of achieving near-optimal performance.
- We show that the proposed concepts can be more beneficial than the IM techniques in the presence of Gray Coding and channel encoding. In this regard, we present that element-wise low-complexity implementations are possible for Gray coding and LLR calculations of the MDS modulation techniques.
- We investigate the SE, BER, and minimum codeword distances of the proposed techniques and provide comprehensive numerical results showing the effectiveness against the classical and combinatorial techniques.
- Our numerical and analytical findings indicate that the proposed methods can outperform existing benchmarks in terms of detection complexity, minimum codeword distance, achievable rate, and error performance.
- The studies in the IM literature focus mainly on showing the effectiveness of the IM techniques against the conventional modulation techniques such as QAM and PSK. However, the actual potential of the classical methods is vastly ignored in these studies as they do not consider well-known encoding techniques when benchmarking the IM techniques. In this regard, this chapter shows that one can achieve better error performance than the IM and related concepts while preserving an encoding and decoding structure as simple as these concepts by using a conventional technique called MDS code.

5.1 Maximum Distance Separable Coded Modulation

PM can outperform conventional modulation techniques such as QAM and PSK since the minimum Hamming distance between the symbols of PM is two [34], whereas such

a distance is one for conventional modulation symbols. However, the potential of the PM concepts is limited since they cannot produce the highest number of codewords whose minimum Hamming distance is two. In this regard, we noticed that it is possible to achieve such a number and outperform the combinatorial concepts in terms of SE and error performance by using a simple MDS code in the modulation design.

The highest number of symbols that achieve the same minimum Hamming distance as index symbols of the combinatorial methods is bounded by the Singleton bound [122]. With $B_Q(L,N)$ denoting the maximum number of possible codewords in a Q-ary block code of length N and the minimum Hamming distance L, the Singleton bound states that

$$B_Q(L,N) \leq Q^{N-L+1}. \tag{5.1}$$

Since $L=2$ for index symbols, the Singleton bound becomes $B_Q(2,N) \leq Q^{N-1}$. This bound can be achieved by a simple MDS code [122].

The MDS code constructs the first $N-1$ elements, $I_\tau, \tau \in \{1,\ldots,N-1\}$, by using the integers $1,\ldots,Q$ as symbols, i.e., $I_\tau \in \{1,\ldots,Q\}$, and the last symbol, I_N, is selected from the same integers by letting the code be those N-tuples summing to zero under modulo-Q arithmetic, i.e., $(I_1 + I_2 + \ldots + I_N) \bmod Q = 0$ [122]. Therefore, by using the MDS code, one can form Q^{N-1} N-tuples. It is important to note that the minimum Hamming distance between the N-tuples of this MDS code is two, and Q^{N-1} is the maximum number of Q-ary N-tuples that satisfy this condition. Now, we define modulation concepts based on this MDS code.

5.1.1 Maximum Distance Separable Amplitude and Phase Modulation

The MDS-APM codebook consists of L_1 codewords and each codeword is an N-tuple of complex numbers which can be regarded as $\mathbf{x}_{\zeta,v} \coloneqq [x_{\zeta,v}(1), x_{\zeta,v}(2), \ldots, x_{\zeta,v}(N)]$ in an Euclidean space $\mathcal{X}$ of $2N$ dimensions where $x_{\zeta,v}(n) = |x_{\zeta,n}|e^{j\varphi_{v,n}} \in \mathbb{C}$ and $n = 1, 2, \ldots, N$. Here, we use the N-tuples of the MDS code to determine indices related to the amplitudes and phases of the codeword elements, i.e., $|x_{\zeta,n}|$ and $\varphi_{v,n}$, respectively. Hence, the values of the codebook elements are chosen according to the integers that form N-tuples of the MDS code. More explicitly, the integers $1, 2, \ldots, Z$ that form an N-tuple regarding the MDS code are used to select one of Z amplitude levels, $\sqrt{A_z}$, $z = 1, \ldots, Z$, for codeword elements in a way that $|x_{\zeta,n}| = \sqrt{\frac{Z}{E}A_{I_{\zeta,n}}}$ where $E = A_1 + A_2 + \ldots + A_Z$ and the subscript $I_{\zeta,n} \in \{1, 2, \ldots, Z\}$ is the nth integer in the ζth N-tuple $\mathcal{I}_\zeta$ related to the MDS code where $\mathcal{I}_\zeta \coloneqq (I_{\zeta,1}, I_{\zeta,2}, \ldots, I_{\zeta,N})$.[3] Since the MDS code has Z^{N-1} N-tuples, we have Z^{N-1} codewords related to amplitude levels, i.e., $\zeta = 1, 2, \ldots, Z^{N-1}$. Similarly, the integers $1, 2, \ldots, J$ that form an N-tuple regarding the MDS code are used to select one of J phases, $\phi_j \in [0, 2\pi]$, $j = 1, \ldots, J$, for codeword elements in a way that $\varphi_{v,n} = \phi_{I_{v,n}}$ where the subscript $I_{v,n} \in \{1, 2, \ldots, J\}$ is the nth integer in the vth N-tuple $\mathcal{I}_v$ related to the MDS code where $\mathcal{I}_v \coloneqq (I_{v,1}, I_{v,2}, \ldots, I_{v,N})$. Since the MDS code produces J^{N-1} N-tuples, we have J^{N-1} codewords related to these tuples, i.e., $v = 1, 2, \ldots, J^{N-1}$. By letting $Q_1 \coloneqq ZJ$, the overall codebook size of MDS-APM is given by $L_1 = Q_1^{N-1}$.

In Table 5.1, we show an example of how we form the amplitude and phase vectors for an MDS-APM system when $N = 3$, $Z = 2$, $J = 2$, $A_1 = 1$, $A_2 = 2$, $\phi_1 = 0$, and $\phi_2 = \pi$. We show the associated constellation diagram in Fig. 5.1. If we have a look

[3] It is important to note that the term $\frac{Z}{E}$ in $|x_{\zeta,n}|$ is for making the average energy of the codeword elements equal to one. One can also check that each codeword element will have the same average energy since each amplitude level will be observed the same number of times in each codeword due to the structure of the MDS code.

Table 5.1. Generating amplitude and phase vectors for MDS-APM when $Z = J = 2$ and $N = 3$.

$(N-1)$-Tuple	N-Tuple MDS Code	Amplitude Vectors	Phase Vectors
$(1,1)$	$\mathcal{I}_1 = (1,1,2)$	$(\sqrt{2/3}, \sqrt{2/3}, \sqrt{4/3})$	$(1,1,e^{j\pi})$
$(1,2)$	$\mathcal{I}_2 = (1,2,1)$	$(\sqrt{2/3}, \sqrt{4/3}, \sqrt{2/3})$	$(1,e^{j\pi},1)$
$(2,1)$	$\mathcal{I}_3 = (2,1,1)$	$(\sqrt{4/3}, \sqrt{2/3}, \sqrt{2/3})$	$(e^{j\pi},1,1)$
$(2,2)$	$\mathcal{I}_4 = (2,2,2)$	$(\sqrt{4/3}, \sqrt{4/3}, \sqrt{4/3})$	$(e^{j\pi},e^{j\pi},e^{j\pi})$

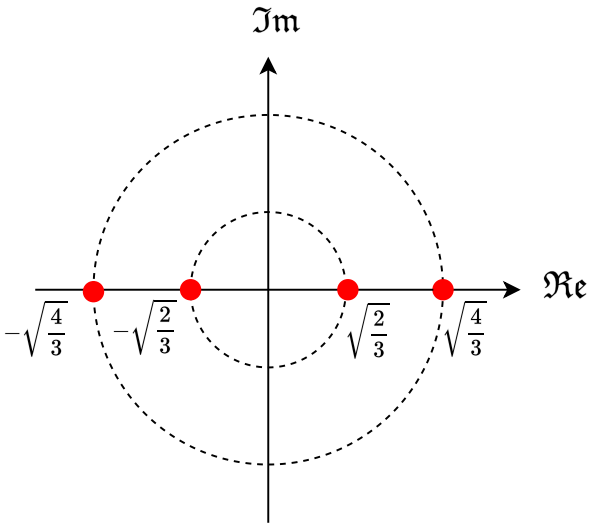


Figure 5.1. Constellation diagram of MDS-APM when $Z = J = 2$ and $N = 3$.

at the amplitude vector related to the second N-tuple $\mathcal{I}_2 = (I_{2,1}, I_{2,2}, I_{2,3}) = (1,2,1)$ in the table, we see that the amplitudes are obtained as $|x_{2,1}| = \sqrt{\frac{Z}{E}A_{I_{2,1}}} = \sqrt{\frac{2}{3}}$, $|x_{2,2}| = \sqrt{\frac{Z}{E}A_{I_{2,2}}} = \sqrt{\frac{4}{3}}$, and $|x_{2,3}| = \sqrt{\frac{Z}{E}A_{I_{2,3}}} = \sqrt{\frac{2}{3}}$ since $A_1 = 1$, $A_2 = 2$ and $Z/E = 2/3$. Moreover, elements of the phase vector related to the same N-tuple are given by $\varphi_{2,1} = \phi_1$, $\varphi_{2,2} = \phi_2$, and $\varphi_{2,3} = \phi_1$ where $\phi_1 = 0$ and $\phi_2 = \pi$. To generate the MDS-APM codewords, we perform element-wise multiplication between the elements of amplitude and phase vectors as indicated in Fig. 5.1. Thus, in the case where we use the same N-tuple $\mathcal{I}_2$ to generate amplitude and phase vectors, the resulting MDS-APM codeword is given by $\mathbf{x}_{2,2} = (\sqrt{2/3}, \sqrt{4/3}e^{j\pi}, \sqrt{2/3})$. Moreover, if we use $\mathcal{I}_1$ and $\mathcal{I}_4$ to construct amplitude and phase vectors, respectively, the corresponding MDS-APM codeword is $\mathbf{x}_{1,4} = (\sqrt{2/3}e^{j\pi}, \sqrt{2/3}e^{j\pi}, \sqrt{4/3}e^{j\pi})$.

In MDS-APM, the mapping of information bits to the amplitude and phase vectors can be performed by using look-up tables. However, once the transmitted signal vector is received at the receiver, the optimum ML detector performs comparisons between the received signal vectors and the signal vectors available in the look-up tables. Thus, the receiver complexity becomes excessive when the codebook size is large. As will be shown in the next section, an element-wise low complexity detection is possible for the proposed concepts; however, we still need to use look-up tables to map the signal vectors to the information bits, and the detection complexity would still be high when the size of the look-up tables is large. To prevent such a high detection complexity, it is crucial to perform the mapping without implementing a look-up table. The MDS structure enables us to perform such a mapping without implementing a look-up table. The bits-to-symbols mapping for the MDS-APM scheme exhibits a very simple structure since Z amplitude levels and J phases can be observed repeatedly on $N-1$ locations and the corresponding $(N-1)$-tuples, which generate the N-tuple MDS code, can easily be converted into decimal numbers according to bases Z and J, respectively. Such decimal numbers related to lexicographically ordered $(N-1)$-tuples are natural numbers that are starting from zero and strictly increasing one by one for such $(N-1)$-tuples just like the lexicographically ordered fixed-length bit sequences. Thus, after mapping each bit sequence to the decimal numbers, we can easily obtain the corresponding $(N-1)$-tuples by using base-Z and base-J representations of such decimal numbers with $N-1$ digits. To obtain these $(N-1)$-tuples, we find the integers α_i, $i=1,\ldots,N-1$ that satisfy $D=\alpha_1 Z^0+\alpha_2 Z^1+\ldots+\alpha_{N-1}Z^{N-2}$ where $0\leq\alpha_i\leq Z-1$ and D is the decimal number corresponding to the bit sequence.[4] Then, the regarding $(N-1)$-tuple is obtained by adding one to α_i, i.e., $(\alpha_1+1,\alpha_2+1,\ldots,\alpha_{N-1}+1)$, since the integers used in the MDS code start from one. For the proposed structure, the receiver is also capable of demapping the detected indices to the bit sequences by

[4]Note that this procedure is equivalent to changing the base of the decimal number. Note also that the same procedure can easily be applied to phase vectors by substituting J into Z.

Table 5.2. Bits-to-index symbols mapping for MDS-APM when $Z = 3$ and $N = 3$.

Bit Sequence	Decimal Number	$(\alpha_1, \alpha_2, \ldots, \alpha_{N-1})$	$(N-1)$-Tuple	N-Tuple
[0 0 0]	0	(0, 0)	(1, 1)	(1, 1, 1)
[0 0 1]	1	(1, 0)	(2, 1)	(2, 1, 3)
[0 1 0]	2	(2, 0)	(3, 1)	(3, 1, 2)
[0 1 1]	3	(0, 1)	(1, 2)	(1, 2, 3)
[1 0 0]	4	(1, 1)	(2, 2)	(2, 2, 2)
[1 0 1]	5	(2, 1)	(3, 2)	(3, 2, 1)
[1 1 0]	6	(0, 2)	(1, 3)	(1, 3, 2)
[1 1 1]	7	(1, 2)	(2, 3)	(2, 3, 1)

looking at the first $N-1$ indices denoted by $\hat{I}_{\zeta,1}, \hat{I}_{\zeta,2}, \ldots, \hat{I}_{\zeta,N-1}$. In this case, we subtract one from these indices and obtain the decoded decimal number as $\hat{D} = (\hat{I}_{\zeta,1} - 1)Z^0 + (\hat{I}_{\zeta,2} - 1)Z^1 + \ldots + (\hat{I}_{\zeta,N-1} - 1)Z^{N-2}$. A simple bits-to-index symbols mapping example is shown in Table 5.2 for $Z = 3$ and $N = 3$.

5.1.2 Maximum Distance Separable In-phase and Quadrature Modulation

The MDS-IQM codebook consists of L_2 codewords. Each codeword is an N-tuple of complex numbers which can be regarded as a vector $\mathbf{x}_{\rho,\delta} = [x_{\rho,\delta}(1), x_{\rho,\delta}(2), \ldots, x_{\rho,\delta}(N)]$ in an Euclidean space $\mathcal{X}$ of $2N$ dimensions where $x_{\rho,\delta}(n) = x^I_{\rho,n} + jx^Q_{\delta,n} \in \mathbb{C}$ and $n = 1, 2, \ldots, N$. Here, the N-tuples of the MDS code are used to specify indices related to the real numbers $x^I_{\rho,n}$ and $x^Q_{\delta,n}$. For the in-phase components of the codeword elements, we use the integers $1, 2, \ldots, R$ that form an N-tuple regarding the MDS code to select one of R real numbers, x_r, $r = 1, \ldots, R$, in a way that $x^I_{\rho,n} = x_{I_{\rho,n}}$. Here, the subscript $I_{\rho,n} \in \{1, 2, \ldots, R\}$ is the nth integer in the ρth N-tuple $\mathcal{I}_\rho$ related to the MDS code where $\mathcal{I}_\rho := (I_{\rho,1}, I_{\rho,2}, \ldots, I_{\rho,N})$. We also assume $\frac{E_T R}{2} = x_1^2 + x_2^2 + \ldots + x_R^2$ to make the average energy of the in-phase components equal to $E_T/2$ where E_T is the constraint on average energy of a codeword element. Moreover, we exploit the integers $1, 2, \ldots, T$ to form the quadrature components in the same way as the

Table 5.3. Generating in-phase and quadrature components for MDS-IQM when $R = T = 2$ and $N = 3$.

$(N-1)$-Tuple	N-Tuple MDS Code	In-phase Components	Quadrature Components
$(1,1)$	$\mathcal{I}_1 = (1,1,2)$	$(\sqrt{2}/2, \sqrt{2}/2, -\sqrt{2}/2)$	$(1/2, 1/2, -\sqrt{3}/2)$
$(1,2)$	$\mathcal{I}_2 = (1,2,1)$	$(\sqrt{2}/2, -\sqrt{2}/2, \sqrt{2}/2)$	$(1/2, -\sqrt{3}/2, 1/2)$
$(2,1)$	$\mathcal{I}_3 = (2,1,1)$	$(-\sqrt{2}/2, \sqrt{2}/2, \sqrt{2}/2)$	$(-\sqrt{3}/2, 1/2, 1/2)$
$(2,2)$	$\mathcal{I}_4 = (2,2,2)$	$(-\sqrt{2}/2, -\sqrt{2}/2, -\sqrt{2}/2)$	$(-\sqrt{3}/2, -\sqrt{3}/2, -\sqrt{3}/2)$

in-phase components. Thus, $x^{Q}_{\delta,n} = \breve{x}_{I_{\delta,n}}$ where $\breve{x}_t \in \mathbb{R}$ and $t = 1, 2, \ldots, T$. The assumption on the average element energy is also valid for the quadrature components, i.e., $\frac{E_T S}{2} = \breve{x}_1^2 + \breve{x}_2^2 + \ldots + \breve{x}_T^2$.[5] Therefore, the average energy is E_T for the codeword elements of MDS-IQM. Due to the structure of the MDS code, the MDS-IQM codebook has $L_2 = Q_2^{N-1}$ codewords where $Q_2 = RT$. Thus, $\rho \in \left\{1, 2, \ldots, R^{N-1}\right\}$ and $\delta \in \left\{1, 2, \ldots, T^{N-1}\right\}$.

In Table 5.3, we give an example of how we form the in-phase and quadrature components of an MDS-IQM scheme when $R = T = 2$, $N = 3$, $x_1 = \sqrt{2}/2$, $x_2 = -\sqrt{2}/2$, $\breve{x}_1 = 1/2$, and $\breve{x}_2 = -\sqrt{3}/2$. Fig. 5.2 illustrates the associated constellation points that we use to generate the MDS-IQM codewords. As seen from the figure, unlike MDS-APM, the MDS-IQM method forms its codewords by combining in-phase and quadrature components. Moreover, when we have a look at the third N-tuple of the MDS code in the table, i.e., $\mathcal{I}_3 = (2,1,1)$, we see that the elements of in-phase components are $x^{I}_{3,1} = x_2$, $x^{I}_{3,2} = x_1$, and $x^{I}_{3,3} = x_1$, whereas the elements of quadrature components are $x^{Q}_{3,1} = \breve{x}_2$, $x^{Q}_{3,2} = \breve{x}_1$, and $x^{Q}_{3,3} = \breve{x}_1$. One can easily observe such a structure from the regarding row of the table. Moreover, to form the codeword $\mathbf{x}_{2,4}$, for example, we combine the second and forth vectors related to in-phase and quadrature components and obtain it as $\mathbf{x}_{2,4} = (\sqrt{2}/2 - j\sqrt{3}/2, -\sqrt{2}/2 - j\sqrt{3}/2, \sqrt{2}/2 - j\sqrt{3}/2)$. The MDS-IQM concept has $(RT)^{N-1} = 16$ codewords. This could be verified by writing all 16 combinations related to the in-phase and quadrature components in the table.

[5]Note that the real numbers that are used to form the in-phase and quadrature components could be the same.

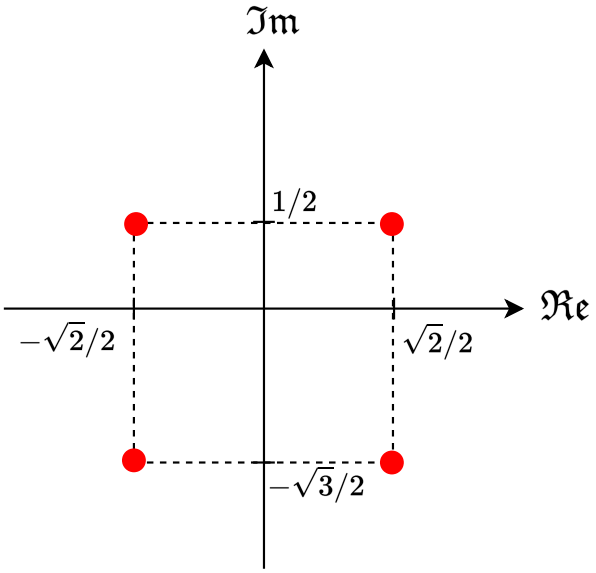


Figure 5.2. Constellation points that are used to generate the MDS-IQM codewords when $R = T = 2$ and $N = 3$.

Remark. The encoding mechanism of the proposed concepts is substantially different from that of PM in [9]. The proposed methods employ an MDS coding mechanism to encode information, whereas the PM concepts use combinatorial tools such as permutations and combinations for the same purpose. Moreover, when there are N different in-phase and quadrature elements for an N-element codeword, the MDS-IQM concept can produce $N^{2(N-1)}$ unique codewords. However, in this case, the PM concept can generate $(N!)^2$ different codewords at most. Since $N^{N-1} > N!$ for $N > 2$, the size of the MDS-IQM codebook is always greater than that of the PM codebook when $N > 2$. One can make the same inferences for the MDS-APM method.

5.2 Practical Model For Parallel Channels

This section represents practical applications of the MDS concepts for parallel channels. By these applications, we aim to indicate general methods of exploiting these concepts in an orthogonal signaling concept. We also aim to conduct fair comparisons among the MDS and combinatorial modulation schemes in terms of achievable rate,

minimum codeword distance, and error performance. Besides, we describe efficient low-complexity detection and bits-to-symbols/symbols-to-bits mapping algorithms for the proposed concepts.

5.2.1 Transmitter

We group the channels and information bits in the same way as the IM, SPM, and CM concepts to ensure low-complexity implementation. Therefore, we assume b input bits enter the transmitter, and a bit splitter splits these bits into $P = b/f$ blocks, each having f input bits. Similarly, the total number of channels N_T is divided into $P = N_T/N$ blocks, each having N channels. For each block of input bits, one of the proposed MDS techniques encodes f information bits into the regarding codewords. The remaining operations are particular for each system and will be explained later in this subsection. Note also that although we group the channels and information bits in the same way as the combinatorial modulation concepts, the proposed methods perform substantially different operations in each group than that of these concepts. Principally, the proposed methods employ an MDS coding mechanism to encode information into the signal vectors whereas, the IM concepts adopt a combinatorial approach based on the permutations and combinations of the channels. Therefore, the proposed implementations can achieve a higher SE and better error performance than the IM and conventional modulation techniques.

Although the models related to the MDS mechanisms employ substantially different encoding procedures in each group of bits and channels, they perform the same encoding procedure in their groups. Thus, we will focus on explaining the encoding mechanism for the pth group, $p = 1, 2, \ldots, P$, of each model in the rest of this section.

5.2.1.1 MDS-PSK/QAM

MDS-APM exploits the N-tuples of an MDS code to decide the amplitudes and phases of the codeword elements. One can also use such tuples to determine the disjoint sets of amplitudes and phases. These sets can be carved from well-known constellations similar to the MM-IM and SPM methods. We call the resulting concepts MDS-PSK and MDS-QAM when we draw the disjoint sets from the PSK and QAM constellations, respectively.

In each block of MDS-PSK/QAM, we use a simple MDS code along with Q disjoint M-ary constellations to construct the N-dimensional complex constellation points where $N \geq 2$. Moreover, we exploit the integers in an N-tuple to determine which element of the related constellation point will be chosen from which disjoint constellation among available disjoint constellations, $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_Q$ where $|\mathcal{M}_q| = |\mathcal{M}_{\hat{q}}| = M$, $\mathcal{M}_q \cap \mathcal{M}_{\hat{q}} = \emptyset$, $q, \hat{q} \in \{1, \ldots, Q\}$, and $q \neq \hat{q}$. Thus, the N-tuple $(I_1, I_2, \ldots, I_N)$ relates to the following set of disjoint constellations $\{\mathcal{M}_{I_1}, \ldots, \mathcal{M}_{I_N}\}$. This process produces Q^{N-1} constellation points as we have Q^{N-1} N-tuples. By using a fixed-length bit-to-symbol mapping scheme, one can map $\lfloor \log_2 Q^{N-1} \rfloor$ bits to these constellation points. After determining the set of disjoint constellations, each element of the constellation points is drawn from regarding disjoint constellation. Since the size of each disjoint constellation is M, i.e., $|M_q| = M, \forall q$, and a constellation point consists of N elements, one can form M^N different constellation points. Hence, the overall codebook size is $Q^{N-1}M^N$.

When implementing MDS-PSK/QAM in parallel channels, we divide f information bits into two parts, one of them having $f_1 = \lfloor \log_2 Q^{N-1} \rfloor$ bits and the other one having $f_2 = N \log_2 M$ bits with $f_1 + f_2 = f$. We use the first f_1 bits to determine a specific N-tuple of the MDS code. The elements of this N-tuple specify the indices of disjoint constellations. Assuming the ωth N-tuple, $\mathcal{I}_\omega = (I_{\omega,1}, I_{\omega,2}, \ldots, I_{\omega,N})$, is selected according to the f_1 bits, we carve the signal carried by the nth channel

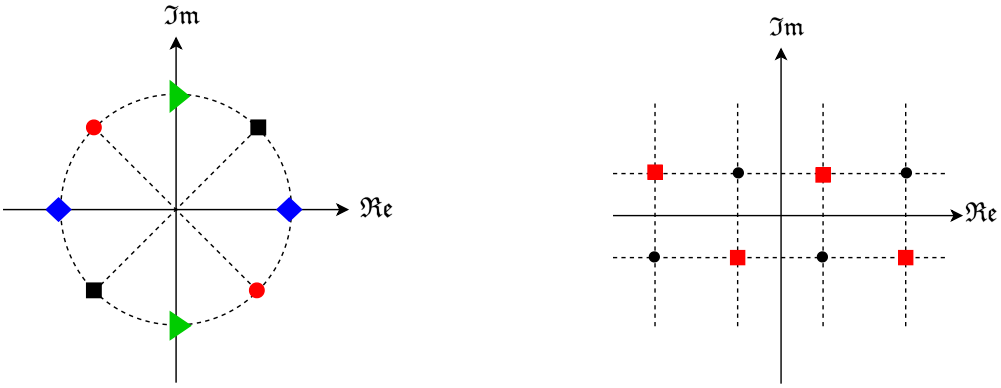
of the pth block, x_n^p, from the disjoint constellation $\mathcal{M}_{I_{\omega,n}}$, i.e., $x_n^p \in \mathcal{M}_{I_{\omega,n}}$ where $n = 1, 2, \ldots, N$. Once we decide the disjoint constellations according to the f_1 bits, the remaining $f_2 = N \log_2 M$ bits choose one of the M symbols contained in the regarding disjoint constellations. Therefore, the overall SE per channel is given by

$$\eta = \frac{\lfloor \log_2 Q^{N-1} \rfloor + N \log_2 M}{N}. \tag{5.2}$$

The symbol vector regarding the pth group can be written as $\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p]$ where $x_n^p \in \mathcal{M}_{I_{\omega,n}}$. After obtaining the symbol vectors for all blocks, a block creator forms the overall symbol as $\mathbf{x} = [x(1), x(2), \ldots, x(N_T)]^T = [\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$.

On the other hand, one can use different approaches to determine the disjoint constellations. As in [1], one can draw the disjoint constellations from well-known PSK and QAM constellations. The MDS scheme that uses PSK constellations to form the disjoint constellations is called MDS-PSK. In this case, the first disjoint constellation is set to a regular M-PSK constellation. Then, the remaining $Q-1$ constellations are obtained by rotating the initial PSK constellation by $\frac{2\pi\tilde{q}}{MQ}$ where $\tilde{q} = 1, 2, \ldots, Q-1$ [1]. In MDS-QAM, on the other hand, the well-known set partitioning technique in [100] is used to carve Q disjoint M-QAM constellations from a (QM)-QAM constellation. To better demonstrate how the disjoint constellations are constructed, we give an example of four disjoint BPSK constellations and two disjoint 4-QAM constellations in Fig. 5.3. Here, the constellation points with the same color and marker are related to the same disjoint constellation. One can observe from Fig. 5.3a that the second, third, and fourth BPSK constellations are obtained by rotating the first one with angles of $45°, 90°$, and $135°$, respectively . Fig. 5.3b shows how we attain two 4-QAM constellations by set partitioning the classical 8-QAM constellation.

Example 4. In Fig. 5.4, we show a practical application of the MDS mechanism in parallel channels. In this application, we perform the MDS encoding over the subcar-



(a) Four disjoint BPSK constellations. (b) Two disjoint 4-QAM constellations.

Figure 5.3. An example of disjoint PSK and QAM constellations.

riers of the well-known OFDM technique. The resulting technique is called *OFDM with MDS-PSK/QAM (OFDM-MDS-PSK/QAM)*. Here, we divide the incoming information bits and subcarriers of the OFDM method into P groups. Each group has N subcarriers and $f = f_1 + f_2$ information bits. We map the first f_1 bits to the N-element MDS codewords and corresponding disjoint PSK/QAM constellations. Then, in each group, we use the remaining f_2 bits to modulate the symbols on the subcarriers according to the disjoint PSK/QAM constellations determined by the first f_1 bits. After the MDS encoding, the same operations as conventional OFDM are applied in OFDM-MDS-PSK/QAM. Note that one can extend the application of the OFDM techniques to the remaining schemes of this section in a similar way described in this example.

5.2.1.2 MDS-APM

In MDS-APM, we split f information bits into three parts, namely f_1, f_2, and f_3 with $f = f_1 + f_2 + f_3$. The first $f_1 = \lfloor \log_2 Z^{N-1} \rfloor$ bits are used to determine the amplitudes of the signals carried by N channels, i.e., the energies of the channels, as described in Section 5.1.1. Thus, the integers $1, 2, \ldots, Z$ of the MDS code are used to select one of

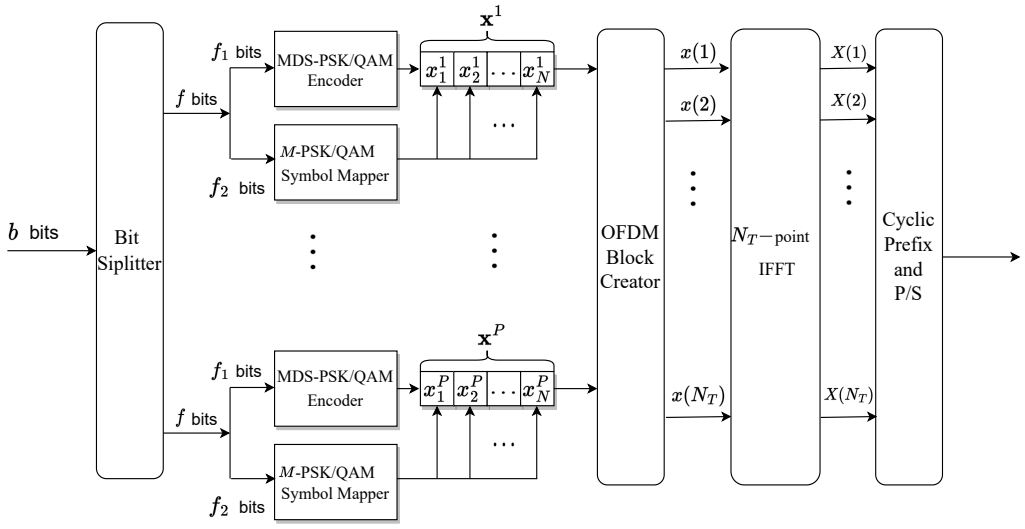


Figure 5.4. Transmitter structure of the OFDM-MDS-PSK/QAM scheme.

Z amplitude levels, $\sqrt{A_z}$, $z = 1, \ldots, Z$ to map the amplitude of the signal carried by a channel, i.e., $|x^p_{\zeta,n}| = \sqrt{\frac{Z}{E} A_{I_{\zeta,n}}}$ where $|x^p_{\zeta,n}|$ is related to the amplitude of the signal carried by the nth channel in the pth group, and the subscript $I_{\zeta,n}$ is the nth integer in the ζth, $\zeta = 1, \ldots, Z^{N-1}$, N-tuple of the MDS code, i.e., $\mathcal{I}_\zeta$. To form a practical MDS-APM scheme, we choose $A_z = z$, thus $|x^p_{\zeta,n}| = \sqrt{\frac{I_{\zeta,n}}{(Z+1)/2}}$.[6]

Once we map f_1 bits to the signal amplitudes, $f_2 = \lfloor \log_2 J^{N-1} \rfloor$ bits are used to select one of the J disjoint phase sets $\mathcal{M}_j$, $j \in \{1, 2, \ldots, J\}$, where each set has M different phases, i.e., $|\mathcal{M}_j| = M, \forall j$.[7] To maximize the angular difference between the constellation points, we obtain the disjoint phase sets $\mathcal{M}_j$ by rotating each constellation with the angle of $2(j-1)\pi/(JM)$, $j = 1, 2, \ldots, J$, as in [1,19]. Note that we start forming the disjoint phase sets by using the phases of a conventional M-PSK modulation, then we obtain the remaining sets by rotating the previous set by $2\pi/(JM)$. In this way, the MED between the constellation points is maximized and the best asymptotic BER performance is attained for MDS-APM. To decide the indices of the disjoint sets on the channels, we use the N-tuples of the MDS code that are formed by

[6] As will be shown in the Numerical Results Section, such a choice of amplitude levels enables MDS-APM to achieve a promising error performance. It also ensures unit average energy across the codeword elements.

[7] For two disjoint phase sets $\mathcal{M}_j$ and $\mathcal{M}_{\hat{j}}$, we have $\mathcal{M}_j \cap \mathcal{M}_{\hat{j}} = \emptyset$ for $j, \hat{j} \in \{1, 2, \ldots, J\}$ and $j \neq \hat{j}$.

the integers $1, 2, \ldots, J$. Then, the remaining $f_3 = N \log_2 M$ bits are used to determine one of the M phases related to the selected disjoint sets for each channel. These phases form the transmit symbol vectors along with the amplitudes determined by f_1 bits. It is important to note that the MDS-APM constellation is similar to that of star-QAM since MDS-APM embeds data into both amplitude levels and phases similar to star-QAM whose constellation consists of multiple concentric circles along with the different phase symbols on them. Thus, the amplitude levels in MDS-APM can be regarded as circles/rings as in star-QAM. To maximize the distance between the constellation points on the consecutive circles, we further rotate each circle by $\pi/(JM)$ degrees compared with the previous circle starting from the one after the innermost circle like the star-QAM scheme in [123]. Although $\pi/(JM)$ is not an optimal rotation angle when $Z \neq 2$, we observed from our simulation results that such a rotation angle enables substantially better error performance, at especially low SNR than the case where there is no rotation. The overall SE per channel for MDS-APM is then given by

$$\eta = \frac{\lfloor \log_2 Z^{N-1} \rfloor + \lfloor \log_2 J^{N-1} \rfloor + N \log_2 M}{N}. \tag{5.3}$$

To show the points that form the MDS-APM constellation, Fig. 5.5 illustrates an example of constellation diagram related to the MDS-APM scheme when $Z = 4$, $J = 2$, and $M = 2$. Here, we have $Z = 4$ amplitude levels (r_1, r_2, r_3, and r_4), $J = 2$ disjoint phase sets, and $M = 2$ phase symbols in each set. We denote the elements of one disjoint set by brown squares, whereas the other is shown by black circles. Moreover, the zth amplitude level, or in other words, the radius of the zth circle is denoted by r_z, $z = 1, 2, 3, 4$. Note also that we rotate the outer circles by $\theta = \frac{\pi}{JM} = \frac{\pi}{4}$ to increase the distance between the constellation points.

Remark. When there are JM phases available in the MDS-APM constellation, one

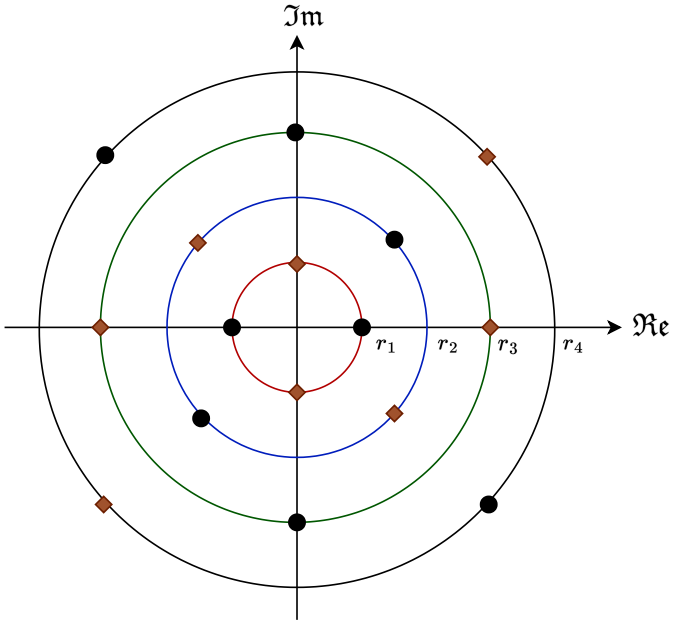


Figure 5.5. Constellation diagram of MDS-APM when $Z = 4$, $J = 2$, and $M = 2$.

can adjust the number of phase sets, J, and the size of each set, M, to meet the requirements of a specific communication setup. For a scenario that requires high SE, one should choose M as much as possible. However, for an error performance-sensitive scenario, J should be high as the corresponding codewords have better Hamming distance properties.

The mapping of f_1 bits to the channels' energies and f_2 bits to the phase sets can be performed without a look-up table implementation as discussed in Section 5.1.1. The symbol vector corresponding to the pth group can be written as $\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p]$ where $x_n^p = |x_{\zeta,n}^p| e^{j\varphi_{v,n}^p}$ and $\varphi_{v,n}^p \in \mathcal{M}_{I_{v,n}}$ where the subscripts $v \in \{1, 2, \ldots, J^{N-1}\}$ and $I_{v,n} \in \{1, 2, \ldots, J\}$. After obtaining the symbol vectors for all blocks, a block creator forms the overall symbol as $\mathbf{x} = [x(1), x(2), \ldots, x(N_T)]^T = [\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$.

5.2.1.3 MDS-IQM

In MDS-IQM, f information bits are divided into two parts, namely f_1 and f_2 bits with $f = f_1 + f_2$. We use f_1 and f_2 bits to determine the in-phase and quadrature

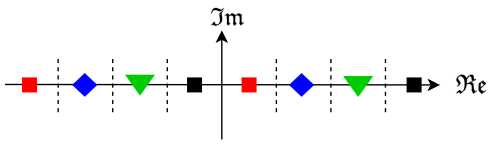


Figure 5.6. Four disjoint PAM constellations produced from the 8-PAM constellation [1].

components, respectively, of the signals carried by the channels. f_1 and f_2 bits are further divided into two parts where $f_1 = f_{11} + f_{12}$ and $f_2 = f_{21} + f_{22}$. Here, we use $f_{11} = \lfloor \log_2 R^{N-1} \rfloor$ bits to determine the indices related to R disjoint M-PAM constellations $\mathcal{P}_r$, $r \in \{1, 2, \ldots, R\}$. To form the disjoint PAM constellations, we use the mode selection strategy in [1, Sec. IV]. As shown in [1], such a mode selection strategy is practical and efficient in terms of error performance. Fig. 5.6 shows four disjoint PAM constellations attained by using the mode selection strategy in [1]. As seen from the figure, the disjoint constellations are generated from the 8-PAM constellation in a way that the minimum Euclidean distance among the points of the same disjoint constellation is maximized.

To obtain the indices of disjoint constellations, we use the N-tuples of the MDS code that is formed by the integers $1, 2, \ldots, R$. Assuming the ρth, $\rho \in \{1, \ldots, R^{N-1}\}$, N-tuple code is selected for transmission in the pth channel/bit group, we associate $I_{\rho,n}$, the nth integer in the ρth N-tuple, with in-phase component of the signal carried by the nth channel and select the in-phase component of the nth channel from $\mathcal{P}_{I_{\rho,n}}$. Once we decide the indices of disjoint constellations for all channels, i.e., $I_{\rho,1}, I_{\rho,2} \ldots, I_{\rho,N}$, $f_{12} = N \log_2 M$ bits are used to select one of M-PAM symbols from $\mathcal{P}_{I_{\rho,1}}, \mathcal{P}_{I_{\rho,2}}, \ldots, \mathcal{P}_{I_{\rho,N}}$ for the channels $1, 2, \ldots, N$, respectively. More specifically, the in-phase component of the signal carried by the nth channel, $\Re(x_n^p)$, is chosen from the $I_{\rho,n}$th PAM constellation, i.e., $\Re(x_n^p) \in \mathcal{P}_{I_{\rho,n}}$ and $I_{\rho,n} \in \{1, 2, \ldots, R\}$. Once the mapping of f_1 bits to the in-phase components is done, $f_2 = f_{21} + f_{22}$ bits are used to determine the quadrature components of the signals in the same way. Hence, we use

$f_{21} = \lfloor \log_2 T^{N-1} \rfloor$ bits to determine the indices related to T disjoint M-PAM constellations, $\mathcal{P}_t$, $t \in \{1, 2, \ldots, T\}$. The N-tuple MDS code related to the integers $1, 2, \ldots, T$ is used to determine the indices. Assuming the incoming f_{21} bits correspond to the δth, $\delta \in \{1, \ldots, T^{N-1}\}$, N-tuple of the MDS code, the quadrature component of the nth channel is associated with the $I_{\delta,n}$th PAM constellation, $\mathcal{P}_{I_{\delta,n}}$ where $I_{\delta,n} \in \{1, \ldots, T\}$. Once the indices are determined for all channels, $f_{22} = N \log_2 M$ bits are used to select one of M-PAM symbols for each channel.[8] Therefore, the quadrature component of the signal carried by the nth channel, $\Im(x_n^p)$, is chosen from the $I_{\delta,n}$th PAM constellation, i.e. $\Im(x_n^p) \in \mathcal{P}_{I_{\delta,n}}$ where $I_{\delta,n} \in \{1, 2, \ldots, T\}$. Thus, the overall SE per channel for MDS-IQM is given by

$$\eta = \frac{\lfloor \log_2 R^{N-1} \rfloor + \lfloor \log_2 T^{N-1} \rfloor + 2N \log_2 M}{N}. \tag{5.4}$$

The mapping of f_{11} and f_{21} bits to the disjoint PAM constellations can be performed without implementing a look-up table as discussed in Section 5.1.1. The symbol vector corresponding to the pth group can be written as $\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p]$ where $x_n^p = \Re(x_n^p) + j\Im(x_n^p)$. After obtaining the symbol vectors for all groups, a block creator forms the overall symbol vector as $\mathbf{x} = [x(1), x(2), \ldots, x(N_T)]^T = [\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^P]^T \in \mathbb{C}^{N_T \times 1}$.

Remark. When N is odd and we use the integers $1, 2, \ldots, N$ to form the N-tuples of the MDS code, the resulting N-tuples of the proposed concepts subsume the permutation indices formed by MM-IM. This can easily be observed from the fact that $(1 + 2 + \ldots + N) \bmod N = 0$ when N is odd. On the other hand, when N is even and we use the integers $1, 2, \ldots, N + 1$ to form the N-tuples of the MDS code, the resulting N-tuples subsume the permutation indices formed by MM-IM since

[8] Note that the PAM constellations that form the in-phase and quadrature components of MDS-IQM do not have to be the same in size. However, we assume each constellation consists of M points for simplicity.

$(1 + 2 + \ldots + N) \bmod (N + 1) = 0$ when N is even. For example, when $N = 3$, one can list the 3-tuples of the MDS code that use the integers $1, 2, 3$ as $(1, 1, 1)$, $(1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(2, 2, 2)$, $(2, 3, 1)$, $(3, 1, 2)$, $(3, 2, 1)$, and $(3, 3, 3)$. Notice that these 3-tuples subsume all permutations of the set $\{1, 2, 3\}$. Thus, the proposed concepts are more general concepts, which encompass the permutation indices generated by an MM-IM encoder. Moreover, the codebooks of the proposed systems are much more flexible than that of MM-IM, since the N-tuples of the MDS code could be generated by integers other than $1, 2, \ldots, N, N + 1$.

5.2.2 Receiver

At the receivers of the proposed methods, the received signal vector can be given as

$$\mathbf{y} \coloneqq [y(1), y(2), \ldots, y(N_T)]^T = \mathbf{X}\mathbf{h} + \mathbf{w} \tag{5.5}$$

where $\mathbf{X} = \text{diag}(\mathbf{x})$. Here, $\mathbf{h}$ and $\mathbf{w}$ are $N_T \times 1$ channel and noise vectors, respectively.

Since the encoding is performed independently for each group of the proposed concepts, decoding can also be performed independently at their receivers. Thus, we explain the decoding procedures for each group of these schemes in the following subsections.

5.2.2.1 MDS-PSK/QAM

The optimum ML detector decides the indices of disjoint constellations as well as the transmitted M-ary symbol vector as follows

$$(\hat{\mathcal{I}}_{\omega}^{p}, \hat{\mathbf{x}}^{p}) = \arg \min_{\mathcal{I}_{\omega}, \mathbf{x}^{p}} ||\mathbf{y}^{p} - \mathbf{X}^{p}\mathbf{h}^{p}||^2 \tag{5.6}$$

where $\hat{\mathcal{I}}_{\omega}^{p}$ and $\hat{\mathbf{x}}^{p}$ are the estimated constellation indices and M-ary symbol vector in the pth group. Moreover, $\mathbf{y}^{p} = [y((p-1)N + 1), \ldots, y(pN)]^T$, $\mathbf{X}^{p} = \text{diag}(\mathbf{x}^{p})$, and

$\mathbf{h}^p = [h((p-1)N+1), \ldots, h(pN)]$. Here, $\mathbf{x}^p = [x_1^p, x_2^p, \ldots, x_N^p]$ where $x_n^p \in \mathcal{M}_{I_{\omega,n}}$, $\forall n$.

The optimum ML detection complexity is of order $O(Q^{N-1}M^N)$ since we have Q^{N-1} and M^N MDS and M-ary modulation symbols, respectively, on N channels. Hence, such a detection mechanism is impractical when N, Q, and M are high. To overcome the excessive complexity of the optimum ML detector, we design a low-complexity suboptimal ML detector that operates on $N-1$ channels independently. Then, the disjoint constellation for the remaining channel is decided according to the disjoint constellations on these $N-1$ channels based on the MDS code. The proposed LC-ML detection hinges on the fact that each disjoint constellation can be used on each channel repeatedly, and the MDS code forms the index symbols in a way that the sum of corresponding N-tuples is zero under modulo-Q arithmetic. The proposed LC-ML detector for an MDS-PSK/QAM block can be explained as follows:[9]

1. Sort the channel gains of N channels in descending order.[10] In other words, $|h(\lambda_1)|^2 \geq \ldots \geq |h(\lambda_N)|^2$ where $\lambda_n \in \{1, \ldots, N\}$.

2. Determine the constellation $\mathcal{M}_{I_\lambda}$, $\lambda \in \{\lambda_1, \ldots, \lambda_{N-1}\}$, and the modulation symbol, $x_\lambda \in \mathcal{M}_{I_\lambda}$, on each channel by substituting all possible constellation symbols, which are drawn from the union of all disjoint constellations, $\mathcal{M} = \mathcal{M}_1 \cup \ldots \cup \mathcal{M}_Q$, into an ML detector except for the channel that has the smallest channel gain, i.e., λ_Nth channel. The detected disjoint constellation index and the modulation symbol for the λth channel can be written as

$$\left(\hat{I}_\lambda,\ \hat{x}_\lambda\right) = \arg \min_{I_\lambda,\ x_\lambda} |y(\lambda) - h(\lambda)x_\lambda|^2. \tag{5.7}$$

3. Estimate the constellation index, $\hat{I}_{\lambda_N}$, on the λ_Nth channel according to the

[9]Here, we omit the block superscript for convenience since one can execute the decoding independently for each block.

[10]Ordering the channel gains and decoding the signals regarding the highest channel gains at first improve the detection performance of the LC-ML detectors considerably since the channels are dependent on each other through an MDS code.

equation

$$(\hat{I}_{\lambda_1} + \hat{I}_{\lambda_2} + \ldots + \hat{I}_{\lambda_N}) \bmod Q = 0. \tag{5.8}$$

4. Determine the modulation symbol, $\hat{x}_{\lambda_N}$, on the λ_Nth channel by substituting M symbols belonging to the constellation $\mathcal{M}_{\hat{I}_{\lambda_N}}$.

The proposed LC-ML detector compares $QM(N-1)$ squared Euclidean distances for $N-1$ channels, which have the highest $N-1$ channel gains, and M squared Euclidean distances for the channel that has the smallest channel gain. The total number of squared Euclidean distance calculations per channel is $QM - QM/N + M/N$. Hence, the computational complexity of the proposed detector is of order $O(QM)$ per channel. It can be shown that the proposed ML detector exhibits a lower complexity than the LLR detector of MM-IM [1] when $Q = N$ and the constellation size of MM-IM is chosen in a way that the overall SEs of both schemes are equal.

5.2.2.2 MDS-APM

By using the optimum ML detector, the indices of amplitude levels, $\mathcal{I}_\zeta^p$, and disjoint phase sets, $\mathcal{I}_v^p$, as well as the phase vector, ϕ_v^p, related to the transmitted symbols for the pth block can be detected as

$$(\hat{\mathcal{I}}_\zeta^p, \hat{\mathcal{I}}_v^p, \hat{\phi}_v^p) = \arg \min_{\mathcal{I}_\zeta, \mathcal{I}_v, \phi_v} ||\mathbf{y}^p - \mathbf{X}^p \mathbf{h}^p||^2 \tag{5.9}$$

where $\mathbf{y}^p = [y((p-1)N+1), \ldots, y(pN)]^T$, $\mathbf{X}^p = \text{diag}(\mathbf{x}^p)$, and $\mathbf{h}^p = [h((p-1)N+1), \ldots, h(pN)]$. Moreover, $\phi_v \coloneqq \{\varphi_{v,1}, \varphi_{v,2}, \ldots, \varphi_{v,N}\}$ is the set of phases drawn from disjoint phase sets with the indices $\mathcal{I}_v$, i.e., $\varphi_{v,n} \in \mathcal{M}_{I_{v,n}}$. The cardinality of this set, $|\phi_v|$, is M^N since we have M phases in each disjoint phase set.

Since there are $(ZJ)^{N-1}M^N$ different realizations related to the amplitude levels,

disjoint phase sets, and corresponding phases, the number of metric calculations in (5.9) is $(ZJ)^{N-1}M^N$. Thus, the complexity of the optimum ML detector is of order $O((ZJM)^N)$. To overcome this exponential complexity, one can employ the LC-ML detector in the same way as described in the previous subsection. During the LC-ML detection, we order the channel gains of N channels in descending order except for the channel that has the minimum channel gain, we perform independent ML detection on each channel by choosing one of Z amplitude levels, J disjoint phase sets, and M phases that constitutes the closest point (among ZJM constellation symbols) to the received signal in two-dimensional Euclidean space. Then, for the remaining channel, we determine the indices of amplitudes and disjoint phase sets in a way that the indices related to the first $N-1$ channels and the remaining channel form an MDS code. Finally, we perform an ML detection on this channel by choosing one of the M phases drawn from the determined phase set. Consequently, the LC-ML detector performs $ZJM(N-1)+M$ metric calculations in total, thus the complexity is of order $O(ZJMN)$. It is important to note that the complexity of the LC-ML detector scales linearly with N, whereas the complexity increases exponentially with N for the optimum ML detector.

5.2.2.3 MDS-IQM

By exploiting the optimum ML detection of MDS-IQM, the estimated index sets, $\hat{\mathcal{I}}_\rho^p, \hat{\mathcal{I}}_\delta^p$, related to the disjoint PAM constellations and sets of PAM symbols, $\hat{\Gamma}_\rho^p, \hat{\mathcal{Q}}_\delta^p$, drawn from these constellations can be written as

$$(\hat{\mathcal{I}}_\rho^p, \hat{\mathcal{I}}_\delta^p, \hat{\Gamma}_\rho^p, \hat{\mathcal{Q}}_\delta^p) = \arg \min_{\mathcal{I}_\rho, \mathcal{I}_\delta, \Gamma_\rho^p, \mathcal{Q}_\delta^p} ||\mathbf{y}^p - \mathbf{X}^p \mathbf{h}^p||^2 \tag{5.10}$$

where $\Gamma_\rho^p \coloneqq \left\{x_{I_{\rho,1}}^p, x_{I_{\rho,2}}^p, \ldots, x_{I_{\rho,N}}^p\right\}$ and $\mathcal{Q}_\delta^p \coloneqq \left\{x_{I_{\delta,1}}^p, x_{I_{\delta,2}}^p, \ldots, x_{I_{\delta,N}}^p\right\}$ are the sets of PAM symbols whose nth elements are drawn from $\mathcal{P}_{I_{\rho,n}}$ and $\mathcal{P}_{I_{\delta,n}}$, respectively.

Table 5.4. Detection complexity comparisons for MDS-IQM, IM, MM-IM, and conventional modulation.

	MDS-IQM (N, R, T, M)		IM (N, K, M_1)	MM-IM (N, M_2)	Conv. Mod. M_3-PSK/QAM
	ML	LC-ML	LLR	LLR	ML
$N = 4, R = 2, T = 2, M = 4, K = 3,$ $M_1 = 102, M_2 = 23, M_3 = 46$	1.05×10^6	52	102	57.5	46
$N = 8, R = 2, T = 4, M = 4, K = 7,$ $M_1 = 142, M_2 = 27, M_3 = 99$	1.13×10^{15}	114	142	121.5	99
$N = 16, R = 4, T = 4, M = 4, K = 15,$ $M_1 = 256, M_2 = 32, M_3 = 216$	1.33×10^{36}	241	256	272	216

Since the MDS-IQM codebook consists of $(RT)^{N-1}M^{2N}$ symbol vectors, the number of metric calculations in (5.10) is $(RT)^{N-1}M^{2N}$. Thus, the complexity of (5.10) is the order of $O((RTM^2)^N)$. Like MDS-PSK/QAM and MDS-APM, the optimum ML detection becomes impractical for MDS-IQM when the codebook size is large. To overcome this high complexity, we exploit the same detection strategy as the LC-ML detection of MDS-PSK/QAM and MDS-APM. Hence, we decide the indices of disjoint PAM sets and related PAM symbols independently for all channels except for the channel corresponding to the minimum channel gain. Then, we determine the indices of disjoint PAM constellations related to the in-phase and quadrature components of the symbol carried by the remaining channel in a way that the estimated indices form a valid N-tuple of the MDS code. Finally, we perform an ML detection for the remaining channel and choose the closest symbol to the received symbol among M^2 symbols drawn from the determined PAM sets. Here, the LC-ML detector performs $(RTM^2)(N-1) + M^2$ metric calculations, thus the complexity is the order of $O(RTM^2N)$.

5.2.2.4 Detection Complexity

In this subsection, we compare the detection complexities related to the low-complexity detectors of the proposed schemes with those of the IM schemes. We also provide detection complexity of conventional modulation for benchmarking.

In Table 5.4, we compare the detection complexities of MDS-IQM, IM, MM-IM, and conventional modulation (Conv. Mod.) in terms of the number of metric calculations per channel.[11] Such a number can be given as $(RT)^{N-1}M^{2N}/N$, $RTM^2 - RTM^2/N + M^2/N$, M_1, $M_2N/2 + M_2/2$, and M_3 for the optimum and low-complexity ML detectors of MDS-IQM (N, R, T, M), the LLR detector of IM (N, K, M_1) [2], the LLR detector of MM-IM (N, M_2) [1], and the optimum ML detector of M_3-PSK/QAM, respectively. Here, MDS-IQM (N, R, T, M) signifies the MDS-IQM concept with N channels in each block as well as R and T disjoint M-PAM constellations. Moreover, IM (N, K, M_1) is the IM scheme that activates K out of N channels in each IM block and carries M_1-ary constellation symbols on the activated channels, whereas MM-IM (N, M_2) stands for the MM-IM concept that uses permutations of N disjoint M_2-ary constellations to form the symbols carried by N channels. To send the same number of information bits for each scheme, thus to make fair comparisons, we adjust M_1, M_2, and M_3, sizes of the constellation in the IM, MM-IM, and conventional modulation schemes, respectively. Although the most values of M_1, M_2, and M_3 may not be practical, we aim to demonstrate how the complexities of different techniques scale compared to one another when the SE is the same. As seen from the table, the LC-ML detector of MDS-IQM is capable of exhibiting a lower complexity than the LLR detectors of IM and MM-IM. Moreover, its complexity is comparable to the ML detector of conventional modulation. It is also worth mentioning that the LC-ML detectors of the proposed schemes do not require knowledge of the noise variance unlike the LLR detectors of IM and MM-IM.

The number of operations per decoded bit is commonly used as a metric for the computational complexity [124]. Let us denote by ζ_s the decoding complexity measured in the receiver operations per information bit where the subscript s is used to

[11] The number of metric calculations of the LC-ML detector will be the same for MDS-PSK/QAM, MDS-APM, and MDS-IQM when the codebooks of these concepts have the same number of index and conventional modulation symbols. Thus, we included only the complexity results related to MDS-IQM in Table 5.4 for brevity.

signify one of the modulation schemes, i.e., $s \in \{\text{IM}, \text{MM-IM}, \text{MDS-APM}\}$. Assuming each scheme has the same SE, η, per channel and number of channels, N, in each block, the decoding complexity for the concepts in the question is given by $\zeta_s = \frac{c_s}{\eta}$ where c_s stands for the complexity in terms of the number of metric calculations per channel. Following the discussion above, the complexities of the MDS-APM, MM-IM, and IM schemes can be written as $c_{\text{MDS-APM}} = ZJM - ZJM/N + M/N$, $c_{\text{MM-IM}} = M_2N/2 + M_2/2$, and $c_{\text{IM}} = M_1$. The SEs of MM-IM and IM are given by $\eta_{\text{MM-IM}} = \frac{\log_2 N! + N\log_2 M_2}{N}$ and $\eta_{\text{IM}} = \frac{\log_2 \binom{N}{K} + K\log_2 M_1}{N}$, respectively. Let us assume that the modulation order employed by each channel is equal for the MDS-APM and MM-IM schemes, i.e., $M_2 = M$. Moreover, to make the SEs of these schemes equal, we should have $\log_2 N! = (N-1)\log_2(ZJ)$, i.e., $ZJ = (N!)^{\frac{1}{N-1}}$. In this way, we ensure that the MDS-APM and MM-IM techniques have the same SE along with the same number of index and conventional modulation symbols, i.e., $\eta = (1/N)\log_2 N! + \log_2 M$. Let us also choose $K = N - 1$ for the IM scheme as this choice is shown in [77] to achieve a high SE for the IM method. It is straightforward to show that one should pick the modulation order on the activated channels as $M_1 = (M^N(N-1)!)^{\frac{1}{N-1}}$ to make the SE of the IM scheme equal to η. Thus, in this case, the decoding complexity per bit for the MDS-APM, MM-IM, and IM schemes can be written respectively as

$$\zeta_{\text{MDS-APM}} = \frac{(N!)^{\frac{1}{N-1}} M(1 - 1/N) + M/N}{\eta}, \tag{5.11}$$

$$\zeta_{\text{MM-IM}} = \frac{MN/2 + M/2}{\eta}, \tag{5.12}$$

$$\zeta_{IM} = \frac{(M^N(N-1)!)^{\frac{1}{N-1}}}{\eta}. \tag{5.13}$$

Remark. When $N \to \infty$, the SEs of MDS-APM and MM-IM can easily be given by $\eta \sim \log_2 ZJM$ and $\eta \sim \log_2 M_2 + \log_2 N - \log_2 e$, respectively, where the latter follows from the Stirling's approximation. To achieve the same SE as MM-IM, one should satisfy $ZJM = M_2N/e$. In this case, the decoding complexities can be written as

$\zeta_{\text{MDS-APM}} \sim \frac{M_2 N}{e \log_2(M_2 N/e)}$ and $\zeta_{\text{MM-IM}} \sim \frac{M_2 N}{2 \log_2(M_2 N/e)}$. Assuming the same modulation order for each channel, one can also show the same result for MDS-PSK/QAM and MDS-IQM. Thus, for the MDS and MM-IM concepts that achieve the same SE with the same number of channels N, the decoding complexity per bit is lower for MDS concepts when $N \to \infty$.

We provide decoding complexity values as a function of $\log_2 N$ in Fig. 5.7 for $M = 4$ and $M = 16$. We also plot the decoding complexity of the conventional modulation in this figure for benchmarking. The figure verifies the result stated in the remark. The proposed MDS-APM method outperforms the MM-IM method since it achieves the same SE as MM-IM by using fewer disjoint constellations. Moreover, its complexity approaches those of the IM and conventional modulation techniques as N increases. It is important to note that although the proposed concepts employing the LC-ML detectors are slightly outperformed by conventional modulation in terms of detection complexity, they exhibit superior error performance compared to conventional modulation as will be shown in the Numerical Results Section.

5.3 Implementation Issues

One-to-one mapping between natural numbers and combinations/permutations eliminates the need for look-up tables, thus reducing the transceiver complexity significantly for the IM concepts. Moreover, the LLR detectors related to the uncoded IM techniques are highly effective in overcoming the excessive complexity of the optimum ML detectors and achieving near-ML performance [1,2,36,38,75]. As a result, one can implement the uncoded IM schemes with a complexity close to that of conventional modulation techniques. However, dependence between the codeword elements makes it challenging to perform the optimum bits-to-index symbols mapping (Gray coding) and low-complexity LLR calculations of each bit related to the IM techniques. Unlike

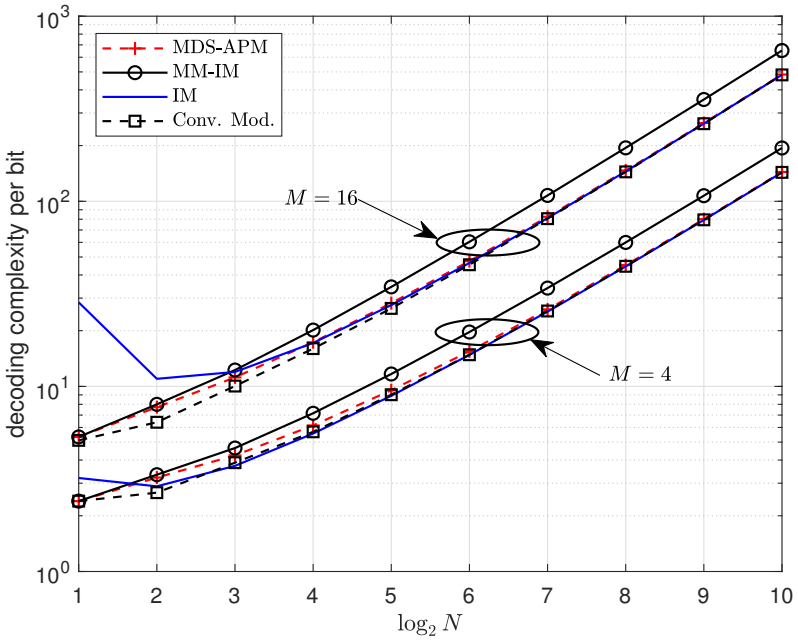


Figure 5.7. Complexity comparison of MDS-APM with MM-IM, IM, and convetional modulation in terms of decoding complexity per bit.

the IM techniques, it is reasonably easy to execute Gray coding and low-complexity LLR calculations for the MDS techniques. We devote this section to manifesting these practical advantages of the MDS concepts. In this regard, we present that element-wise low-complexity implementations are possible for Gray coding and LLR calculations of the MDS modulation techniques.

5.3.1 Gray Coding

We showed in Section 5.1.1 that the mapping of the information bits to the MDS tuples can be performed without implementing a look-up table. Hence, the transceiver complexity can be effectively reduced by implementing such a mapping. We noticed that the structure of the MDS code formed by the integers $1, 2, \ldots, Q$ enables Gray-like mapping when Q is an integer power of two, i.e., $Q = 2^l$ where $l \geq 1$ is an integer. In this case, we have $l(N-1)$ bits to map to the N-tuples of the MDS code, and one can map l information bits to the first $N-1$ elements of an N-tuple separately by

Table 5.5. An example of Gray mapping for the MDS code when $N = 3$, $Q = 4$, and $l = 2$.

Mapping for the First Eight N-Tuples			Mapping For The Remaining N-Tuples		
The Mapping in Sec. 5.1.1	Gray Mapping	N-Tuple MDS Code	The Mapping in Sec. 5.1.1	Gray Mapping	N-Tuple MDS Code
[0 0 0 0]	[0 0 0 0]	$\mathcal{I}_1 = (1, 1, 2)$	[1 1 0 1]	[1 0 0 1]	$\mathcal{I}_9 = (4, 2, 2)$
[0 0 0 1]	[0 0 0 1]	$\mathcal{I}_2 = (1, 2, 1)$	[1 1 0 0]	[1 0 0 0]	$\mathcal{I}_{10} = (4, 1, 3)$
[0 0 1 0]	[0 0 1 1]	$\mathcal{I}_3 = (1, 3, 4)$	[1 0 0 0]	[1 1 0 0]	$\mathcal{I}_{11} = (3, 1, 4)$
[0 0 1 1]	[0 0 1 0]	$\mathcal{I}_4 = (1, 4, 3)$	[0 1 0 0]	[0 1 0 0]	$\mathcal{I}_{12} = (2, 1, 1)$
[0 1 1 1]	[0 1 1 0]	$\mathcal{I}_5 = (2, 4, 2)$	[0 1 0 1]	[0 1 0 1]	$\mathcal{I}_{13} = (2, 2, 4)$
[1 0 1 1]	[1 1 1 0]	$\mathcal{I}_6 = (3, 4, 1)$	[0 1 1 0]	[0 1 1 1]	$\mathcal{I}_{14} = (2, 3, 3)$
[1 1 1 1]	[1 0 1 0]	$\mathcal{I}_7 = (4, 4, 4)$	[1 0 1 0]	[1 1 1 1]	$\mathcal{I}_{15} = (3, 3, 2)$
[1 1 1 0]	[1 0 1 1]	$\mathcal{I}_8 = (4, 3, 1)$	[1 0 0 1]	[1 1 0 1]	$\mathcal{I}_{16} = (3, 2, 3)$

using Gray coding. Thus, we map l information bits to these elements in a way that information bits related to consecutive integers differ by only one bit. In this way, we ensure a one-bit difference between the information bits related to the pairs of $(N-1)$-tuples that differ by one element, and the different elements are consecutive integers. Since the pairs of disjoint constellations related to consecutive integers are closer to each other, such a mapping provides a Gray-like coding structure for the MDS modulation methods. However, this structure may not be optimum since the last integer, Q, rolls over to the first integer, 1, with only a one-bit difference as in classic Gray coding, and the disjoint constellations related to the first and last integers are not the closest constellation pairs. On the other hand, Gray code for the tuples of the MDS code can be obtained by starting with an $(N-1)$-tuple of all ones and increasing or decreasing elements of such a tuple one by one until we have Q^{N-1} different tuples.[12] In this case, consecutive N-tuples will differ by only two elements and these elements will be successive integers.

In Table 5.5, we give an example of Gray mapping of bit strings to N-tuples of an MDS code when $N = 3$, $Q = 4$, and $l = 2$. First, we obtain the Gray code for 3-tuples of the MDS code by increasing or decreasing the first two integers one-by-one until

[12]Here, we exploit the cyclic property of Gray code where the first and the last integers are regarded as consecutive integers.

we have $Q^{N-1} = 16$ different tuples. Then, the mapper maps the first $l = 2$ bits of the bit strings to the first element of the 3-tuples. The remaining $l = 2$ bits of the bit strings are mapped to the second element of the 3-tuples. Notice that there is only a one-bit difference between the bit strings related to consecutive integers due to the Gray mapping. Note also that the minimum Hamming distance between the consecutive 3-tuples is two, and the bit strings related to the pairs of 3-tuples that have such a distance with consecutive integers differ only one bit. In the table, we also provide the bit strings related to the mapping in Section 5.1.1 to show the difference between the mappings.

The complexity of the proposed Gray-like mapping is order $O(N2^l)$ since the mapping of 2^l bits is performed on $N-1$ elements separately. For the index symbols of the generalized IM technique, the complexity of the Gray coding algorithm in [13] is of order $O(\beta^2)$ where $\beta = 2^{\lfloor \log_2 \binom{N}{K} \rfloor}$, N, and K are the number of total and activated codeword elements, respectively. By assuming the number of information bits related to the tuples of the MDS code is equal to that of the index symbols, i.e., $2(N-1)l = \lfloor \log_2 \binom{N}{K} \rfloor$, it can be shown that the complexity of the proposed Gray-like mapping becomes only order $O(\log_2 \beta)$. Thus, the Gray coding complexity is much lower for the MDS technique than the IM technique.

5.3.2 Log-Likelihood Ratio Calculation

We need to calculate LLR values for each bit of the MDS modulation techniques to obtain markedly improved error performance in the presence of channel encoding. That is because soft-decision decoding based on the LLR values has a clear BER advantage against hard-decision decoding for coded techniques. Thus, in this subsection, we calculate the LLR values of the MDS-IQM concept. Here, we only consider the MDS-IQM method for ease of representation. It is also because one can calculate the LLR values for other MDS methods by following the steps in this subsection.

Since MDS-IQM employs the MDS mechanism separately on in-phase and quadrature parts of the codeword elements, corresponding LLR values can be calculated for these parts independently. To better illustrate how we calculate these values, let us denote the received signal related to the nth codeword element by dropping the block indices as follows

$$y(n) = h(n)x_n + w(n) \tag{5.14}$$

where $h(n)$, x_n and $w(n)$ are the nth elements of $\mathbf{h}$, $\mathbf{x}$, and $\mathbf{w}$, respectively. Dividing both sides of the equation by $h(n)$, one can obtain

$$\tilde{y}(n) = x_n + \tilde{w}(n) \tag{5.15}$$

where $\tilde{y}(n) = y(n)/h(n)$ and $\tilde{w}(n) = w(n)/h(n)$. Now, let us denote $\tilde{y}(n)$ in terms of its in-phase and quadrature components and vectors related to these components as $\tilde{y}(n) = \tilde{y}^I(n) + j\tilde{y}^Q(n)$, $\tilde{\mathbf{y}}^I = [\tilde{y}^I(1), \tilde{y}^I(2), \ldots, \tilde{y}^I(N)]$, and $\tilde{\mathbf{y}}^Q = [\tilde{y}^Q(1), \tilde{y}^Q(2), \ldots, \tilde{y}^Q(N)]$, respectively. Moreover, vectors related to the in-phase and quadrature components of a signal vector $\mathbf{x}$ can be denoted in a similar way as $\mathbf{x}^I = [x_1^I, x_2^I, \ldots, x_N^I]$ and $\mathbf{x}^Q = [x_1^Q, x_2^Q, \ldots, x_N^Q]$, respectively, where $x_n = x_n^I + jx_n^Q$. Following this notation, LLR of the τth bit, b_τ^I, related to the in-phase parts of the transmitted signal can be calculated as

$$L(b_\tau^I) = \ln \frac{p(b_\tau^I = 0|\tilde{\mathbf{y}}^I)}{p(b_\tau^I = 1|\tilde{\mathbf{y}}^I)} \propto \ln \frac{\sum_{\mathbf{x}^I \in \mathcal{X}_{\tau,0}} p(\tilde{\mathbf{y}}^I|\mathbf{x}^I)}{\sum_{\mathbf{x}^I \in \mathcal{X}_{\tau,1}} p(\tilde{\mathbf{y}}^I|\mathbf{x}^I)} \tag{5.16}$$

where $\mathcal{X}_{\tau,0}$ and $\mathcal{X}_{\tau,1}$ are signal spaces that consist of signal vectors regarding binary bit sequences with $b_\tau^I = 0$ and $b_\tau^I = 1$, respectively. The proportionality at the right-hand side of (5.16) is a result of the assumption that the signal vectors are equally likely. Due to the independence among the elements of the received signal vector, the

likelihood probability $p(\tilde{\mathbf{y}}^I|\mathbf{x}^I)$ can be obtained by

$$p(\tilde{\mathbf{y}}^I|\mathbf{x}^I) = p(\tilde{y}^I(1)|x_1^I)p(\tilde{y}^I(2)|x_2^I)\dots p(\tilde{y}^I(N)|x_N^I) \tag{5.17}$$

where

$$p(\tilde{y}^I(n)|x_n^I) = \frac{|h(n)|}{\pi\sqrt{N_0}} \exp\left(- \frac{|h(n)|^2(\tilde{y}^I(n) - x_n^I)^2}{N_0}\right). \tag{5.18}$$

It is important to note that LLR values of the bits related to quadrature components of the transmitted signals can be calculated in the same way as (5.16). However, assuming $R = T = Q$ for MDS-IQM, (5.16) and (5.18) require $Q^{2(N-1)}M^{2N}$ squared Euclidean distance calculations per bit. Hence, the receiver complexity becomes dramatically high when Q, N, and M are high. To solve this issue and reduce the receiver complexity, one can simply calculate the LLR values in an element-wise manner for the MDS modulation technique. This is only possible because each group of $\log_2 Q$ bits in $(\log_2 Q^{N-1})$-length bit sequence related to the MDS tuples are mapped to the first $N-1$ elements of an N-tuple code, separately. For example, let us consider the mapping example in Table 5.5 where we have $N = 3$, $Q = 4$, and $l = 2$. If we have a look at the bit sequence [0 0 0 1], we see that the first two $(\log_2 4 = 2)$ bits, i.e., "00", are used to determine the first element of regarding N-tuple as 1, whereas the second two bits, "01", are used to determine the second element as 2. The last element is $I_3 = 1$ since $(1 + 2 + I_3) \mod 4 = 0$ where $I_3 = 1$. Hence, we divide $(N-1)\log_2 Q$ bits related to the MDS code into groups of $\log_2 Q$ bits and calculate the LLR value of the ϕth bit in the Λth $(\log_2 Q)$-length bit sequence related to the MDS code as

$$L(b_{\phi,\Lambda}^I) = \ln \frac{p(b_{\phi,\Lambda}^I = 0|\tilde{y}^I(\Lambda))}{p(b_{\phi,\Lambda}^I = 1|\tilde{y}^I(\Lambda))} \propto \ln \frac{\sum_{x_\Lambda^I \in \mathcal{X}_{\phi,0}} p(\tilde{y}^I(\Lambda)|x_\Lambda^I)}{\sum_{x_\Lambda^I \in \mathcal{X}_{\phi,1}} p(\tilde{y}^I(\Lambda)|x_\Lambda^I)} \tag{5.19}$$

where $\phi \in \{1, \dots, \log_2 Q\}$, $\Lambda \in \{1, \dots, N-1\}$, and $\mathcal{X}_{\phi,0}$ is the signal space that

consists of one-dimensional signals with $b^I_{\phi,\Lambda} = 0$. Moreover, LLR values of the mth ($m = 1, 2, \ldots, \log_2 M$) bit related to nth ($n = 1, 2, \ldots, N$) conventional PAM symbols drawn by the in-phase part of the nth element can be calculated as

$$L(b^I_{m,n}) = \ln \frac{p(b^I_{m,n}=0|\tilde{y}^I(n))}{p(b^I_{m,n}=1|\tilde{y}^I(n))} \propto \ln \frac{\sum\limits_{x^I_n \in \mathcal{X}_{m,0}} p(\tilde{y}^I(n)|x^I_n)}{\sum\limits_{x^I_n \in \mathcal{X}_{m,1}} p(\tilde{y}^I(n)|x^I_n)}. \tag{5.20}$$

For the information bits related to the quadrature components of the constellation points, LLR values can be calculated in the same way as in (5.19) and (5.20). A receiver using these equations performs QM squared Euclidean distance calculations per bit, thus such a receiver exhibit much lower complexity than that of a receiver using (5.16) and (5.18). However, as will be shown in Fig. 5.17, a receiver performing only (5.19) and (5.20) in the presence of channel encoding fails to provide near-optimal error performance when $N < 4$. To solve this problem, instead of employing the MDS coding mechanism among symbol vectors, equivalent $(N, N-1)$ single parity check (SPC) code can be employed among information bits, and the resulting bit sequence can be mapped to the symbol vectors. Then, extrinsic values related to the information bits can be used to improve the decoding performance. Let us again focus on the information bits related to the in-phase parts. In an $(N, N-1)$ SPC code, the ϕth parity check bit, $b^I_{\phi,N}$, satisfies $b^I_{\phi,N} = b^I_{\phi,1} \oplus b^I_{\phi,2} \oplus \ldots \oplus b^I_{\phi,N-1}$ where $\oplus$ is addition operation under modulo-2 arithmetic. For this code, it is also clear that $b^I_{\phi,\Delta} = \bigoplus\limits_{\kappa \in \mathcal{N}\setminus\Delta} b^I_{\phi,\kappa}$ where $\mathcal{N} \coloneqq \{1, 2, \ldots, N\}$ and $\bigoplus$ is the summation operation under modulo-2 arithmetic. Hence, the extrinsic value related to $b^I_{\phi,\Delta}$ can be calculated as [30]

$$L_e(b^I_{\phi,\Delta}) = L\left(\bigoplus_{\kappa \in \mathcal{N}\setminus\Delta} b^I_{\phi,\kappa} \right) \tag{5.21}$$

Table 5.6. Complexity order comparison of LLR calculations per bit for MDS-IQM, IM, and MM-IM.

System Parameters	MDS-IQM (N, Q, M)		IM (N, K, M_1)	MM-IM (N, M_2)
	Opt. Soft	LC Soft		
$N = 2, Q = 2, M = 2, K = 1, M_1 = 32, M_2 = 6$	64	4	32	6
$N = 4, Q = 4, M = 4, K = 3, M_1 = 406, M_2 = 64$	2.68×10^8	16	406	64
$N = 8, Q = 8, M = 8, K = 7, M_1 = 5513, M_2 = 664$	1.24×10^{27}	64	5513	664

$$= 2 \operatorname{arctanh}\left(\prod_{\kappa \in \mathcal{N} \setminus \Delta} \tanh\left(L(b^I_{\phi,\kappa})/2\right) \right)$$

$$\approx \left(\prod_{\kappa \in \mathcal{N} \setminus \Delta} \operatorname{sign}(L(b^I_{\phi,\kappa})) \right) \min_{\kappa \in \mathcal{N} \setminus \Delta} |L(b^I_{\phi,\kappa})|.$$

Then, we use this extrinsic value to update the LLR of $b^I_{\phi,\Delta}$ as follows [30]

$$\breve{L}(b^I_{\phi,\Delta}) = L(b^I_{\phi,\Lambda}) + L_e(b^I_{\phi,\Lambda}). \tag{5.22}$$

On the other hand, the order of complexity per bit related to the proposed low-complexity LLR calculation method is $O(QM)$, whereas it can be lower-bounded by $O(M_1)$ and $O(M_2)$ for the low-complexity LLR calculation methods of IM (N, K, M_1) in [125] and MM-IM (N, M_2) in [126]. Although the IM techniques seem more beneficial than the MDS method, the MDS method can achieve the same SE as the IM techniques while $QM < M_1, M_2$. To better understand these complexities, we provide numerical complexity results in Table 5.6. In addition, the table incorporates the complexity results related to the optimum LLR calculations of the MDS-IQM techniques. As seen from the table, the low-complexity method is highly efficient in overcoming the excessive complexity of the optimum detection for the MDS-IQM system. Moreover, it can achieve lower complexity for different codeword lengths and SEs than the low-complexity methods of the IM and MM-IM schemes.

5.4 Performance Analysis

5.4.1 Spectral Efficiency

Based on the SE expressions in (5.2), (5.3), and (5.4), when $N \to \infty$, the asymptotic SE of MDS-PSK/QAM, MDS-APM, and MDS-IQM can be simply given as

$$\eta_{MDS-PSK/QAM} \sim \log_2(QM), \tag{5.23}$$

$$\eta_{MDS-APM} \sim \log_2(ZJM), \tag{5.24}$$

$$\eta_{MDS-IQM} \sim \log_2(RTM). \tag{5.25}$$

Remark. The proposed techniques exhibit outstanding flexibility since Q, Z, J, R, T, N, and M can be adjusted independently to achieve a desired SE. On the other hand, it is easy to notice that the proposed techniques are equivalent to parallel transmission of the M-QAM, M-PSK, and M-PAM symbols when $Q = 1$ for MDS-QAM, $Z = J = 1$ for MDS-APM, and $R = T = 1$ for MDS-IQM. Moreover, when $Q = N$, we arrive at N^{N-1} index codewords for the proposed MDS-PSK/QAM concept and the SE scales as $\eta \sim \log_2(NM)$, which is the highest SE among the IM schemes employing N disjoint constellations over N channels. That is also the case for the MDS-APM and MDS-IQM concepts.

Proposition 6. As $N \to \infty$, the index codewords of MDS-PSK/QAM are capable of achieving the same SE as the overall SEs of the MM-IM and OFSPM methods, which employ N disjoint M-ary constellations, when $Q = NM/e$ and $Q = NM/(e \ln 2)$, respectively.

Proof. The proof can easily be obtained by looking at the SEs of MM-IM and OFSPM, and the SE provided by only the index symbols of the MDS-PSK/QAM scheme. They can be given, respectively, as $\eta \sim \log_2 N - \log_2 e + \log_2 M$, $\eta \sim \log_2 N - \log_2(e/\ln 2) + \log_2 M$ and $\eta \sim \log_2 Q$ when $N \to \infty$. □

Based on the fact in Proposition 6, the proposed MDS-PSK/QAM method can achieve the same SE as that of the MM-IM scheme without employing conventional modulation while utilizing fewer constellation points. Such a result makes the proposed MDS scheme not only spectrally more efficient but also more reliable in terms of error rate than MM-IM since the index symbols[13] are capable of achieving a higher minimum Hamming distance than the conventional modulation symbols. Note also that the SE analysis in Proposition 6 is limited with MDS-PSK/QAM since one can obtain similar results for the MDS-APM and MDS-IQM techniques.

5.4.2 Average Bit Error Rate

In this subsection, for the sake of completeness, we derive an upper-bound based on the well-known union bound for the average BER of the proposed schemes. As shown in the Numerical Results Section, the derived bound is tight, especially at high SNR.

Based on the ML detection rules in (5.6), (5.9) and (5.10), the PEP conditioned on the channel coefficients, $\mathbf{h}$, can be given as

$$P(\mathbf{X}_i \rightarrow \mathbf{X}_j|\mathbf{h}) = Q\left(\sqrt{\frac{\gamma||(\mathbf{X}_i - \mathbf{X}_j)\mathbf{h}||^2}{2}}\right) \tag{5.26}$$

where $\mathbf{X}_i = \text{diag}(\mathbf{x}_i)$, $\mathbf{X}_j = \text{diag}(\mathbf{x}_j)$, $i \neq j$, $i, j \in \left\{1, 2, \ldots, 2^f\right\}$, and $\gamma = 1/N_0$ is the average received SNR. Here, $\mathbf{x}_i$ and $\mathbf{x}_j$ are the ith and jth codewords of the corresponding MDS method.

By substituting the identity $Q(x) = \frac{1}{12}e^{-x^2/2} + \frac{1}{4}e^{-2x^2/3}$ into (5.26) and taking an average over $\mathbf{h}$, we obtain the following approximate unconditional PEP expression

[13]Here, we call the symbols related to the N-tuples of the MDS code as index symbols, since we use the MDS code to embed information into the indices of codeword elements, and they have the same minimum Hamming distance as those of the index symbols of the IM benchmarks.

as in the previous chapters:

$$P(\mathbf{X}_i \rightarrow \mathbf{X}_j) = \mathrm{E}_{\mathbf{h}}[P(\mathbf{X}_i \rightarrow \mathbf{X}_j|\mathbf{h})] \approx \frac{1/12}{\det(\mathbf{I}_N + \frac{\gamma}{4}\mathbf{C}\mathbf{A}_{ij})} + \frac{1/4}{\det(\mathbf{I}_N + \frac{\gamma}{3}\mathbf{C}\mathbf{A}_{ij})} \tag{5.27}$$

where $\mathbf{I}_N$ is the identity matrix of size $N \times N$, $\mathbf{C} = \mathrm{E}_{\mathbf{h}}[\mathbf{h}\mathbf{h}^H]$, and $\mathbf{A}_{ij} = (\mathbf{X}_i - \mathbf{X}_j)^H(\mathbf{X}_i - \mathbf{X}_j)$.

The average BER of the proposed MDS methods can be upper-bounded in the same way as in [1, 2] by using the well-known union bound as follows

$$P_b \leq \frac{1}{f2^f} \sum_{i=1}^{2^f} \sum_{j=1}^{2^f} P(\mathbf{X}_i \rightarrow \mathbf{X}_j) D(\mathbf{X}_i \rightarrow \mathbf{X}_j) \tag{5.28}$$

where $D(\mathbf{X}_i \rightarrow \mathbf{X}_j)$ is the Hamming distance between bit sequences related to codewords $\mathbf{x}_i$ and $\mathbf{x}_j$.

Remark. The codebooks of MDS-PSK/QAM, MDS-APM, and MDS-IQM consist of Q^{N-1}, $(ZJ)^{N-1}$ and $(RT)^{N-1}$ codewords related to the indices of constellation points, respectively, and the minimum Hamming distance related to these codewords is two. However, they also include M^N, M^N, and M^{2N} codewords, respectively, related to the conventional modulation symbols (QAM/PSK, PSK, and PAM symbols for MDS-PSK/QAM, MDS-APM, and MDS-IQM) that limit the minimum Hamming distance to one. Thus, the diversity order, i.e., the slope of the BER curves, related to the proposed systems is one. On the other hand, a codebook can incorporate only the codewords related to the indices and exhibit a diversity order equal to two. In this case, the SEs of the proposed concepts become $\eta = \frac{\lfloor \log_2 Q^{N-1} \rfloor}{N}$, $\eta = \frac{\lfloor \log_2 Z^{N-1} \rfloor + \lfloor \log_2 J^{N-1} \rfloor}{N}$ and $\eta = \frac{\lfloor \log_2 R^{N-1} \rfloor + \lfloor \log_2 T^{N-1} \rfloor}{N}$ for MDS-PSK/QAM, MDS-APM and MDS-IQM, respectively. As we will show in the Numerical Results Section, these special cases of the proposed schemes can achieve a superior error performance at especially high SNR.

5.4.3 Minimum Codeword Distances

For fading channels, the error performance at low SNR is mainly limited by the MEDs between codeword pairs and the number of codeword pairs that are apart from each other with this distance. However, the minimum L-product distance between the codeword pairs and the number of these pairs limit the error performance at high SNR. Since the minimum Hamming distance is equal to one, i.e., $L = 1$, for the proposed MDS methods due to the presence of conventional modulation symbols, the minimum L-product distance becomes the MED between the unit distance codeword pairs. In this subsection, we calculate the minimum codeword distances of the proposed schemes that are critical measures to assess the error performance.

We provided the formal definitions of the minimum L-product distance and MED in (2.53) and (2.54), respectively. In this subsection, we will follow the same notation in Section 2.5.3 and use d_1 and d_2 for the former and latter distances, respectively.

5.4.3.1 MDS-PSK

For MDS-PSK, we observe the unit distance codeword pairs when we draw two codewords from the same set of disjoint PSK constellations and choose the same M-PSK symbols from these constellations for all codeword elements except for a codeword element. In this case, the minimum L-product distance is the MED between the symbols of a disjoint M-PSK constellation, i.e.,

$$d_1 = 2\sin\left(\frac{\pi}{M}\right). \tag{5.29}$$

On the other hand, the MED between any elements of codeword pairs is $2\sin\left(\frac{\pi}{QM}\right)$ since the union of all disjoint constellations constitutes a (QM)-PSK constellation. However, we observe such a distance in at least two elements of the codeword pairs

due to the MDS structure. Therefore, the MED of MDS-PSK can be given as

$$d_2 = 2\sqrt{2}\sin\left(\frac{\pi}{QM}\right). \tag{5.30}$$

Note that it is straightforward to show $d_1 \geq d_2$. This is also why d_2 is the MED between the codeword pairs of an MDS-PSK scheme.

5.4.3.2 MDS-QAM

As we mentioned in Section 5.2.1.1, MDS-QAM carves Q disjoint M-QAM constellations from a (QM)-QAM constellation. The MED between the points of a (QM)-QAM constellation can be given by

$$d_{min} = \begin{cases} \frac{2\sqrt{6}}{\sqrt{5QM-4}} & \text{for rectangular } (QM)-\text{QAM}, \\ \frac{\sqrt{6}}{\sqrt{QM-1}} & \text{for square } (QM)-\text{QAM}. \end{cases} \tag{5.31}$$

Like MM-IM, we use the well-known set partitioning technique in [100] to from the disjoint constellations in MDS-QAM. Therefore, one can calculate the minimum L-product distance for MDS-QAM in the same way as that of the MM-IM technique employing disjoint QAM constellations. Such a distance is given by [1]

$$d_1 = \begin{cases} d_{min}\frac{\sqrt{5Q}}{2} & \text{for rectangular } (QM)-\text{QAM}, \\ d_{min}\sqrt{Q} & \text{for square } (QM)-\text{QAM}. \end{cases} \tag{5.32}$$

On the other hand, the MED of an MDS-QAM codebook is given by

$$d_2 = \sqrt{2}d_{min} \tag{5.33}$$

since we observe d_{min} in at least two elements of the MDS-QAM codeword pairs.

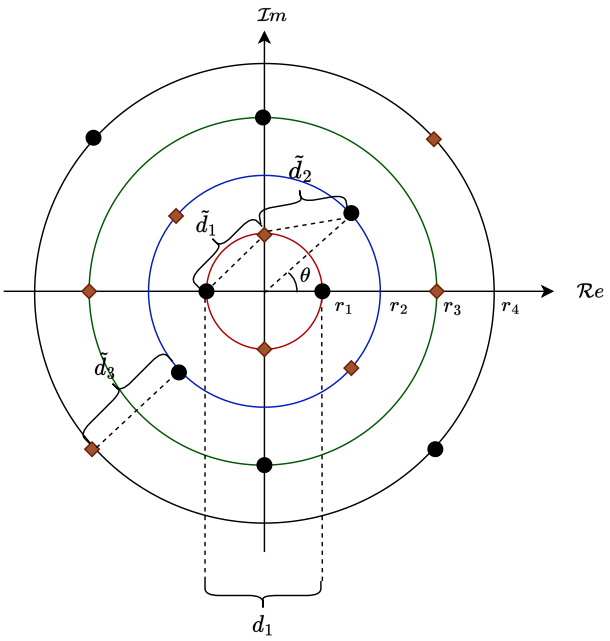


Figure 5.8. Constellation diagram of MDS-APM when $Z = 4$, $J = 2$, and $M = 2$.

5.4.3.3 MDS-APM

MDS-APM exhibits similar characteristics to the star-QAM constellation since its codebook consists of multiple concentric circles with phase symbols on them. To make it clearer, we revisit the constellation diagram in Fig. 5.5 and show the Euclidean distances among the regarding constellation points in Fig. 5.8.

In MDS-APM, the minimum Hamming distance codeword pairs are the unit distance pairs in the disjoint phase sets. The MED between those pairs is determined by the pairs in the innermost circle and calculated by

$$d_1 = \frac{2\sqrt{2}}{\sqrt{Z+1}} \sin\left(\frac{\pi}{M}\right) \tag{5.34}$$

since we have M phases in each disjoint phase set. As mentioned earlier, d_1 is the limiting factor for the error performance of MDS-APM at high SNR.

As can be seen from Fig. 5.8, the MED between the constellation points is limited by either $\tilde{d}_1$, the MED between the constellation symbols in the inner-most circle, or

$\tilde{d}_2$, the MED between the points in the consecutive circles, or $\tilde{d}_3$, the MED between the points of κth and $(\kappa+2)$th circles where $\kappa \in \{1, 2, \ldots, Z-2\}$. However, unlike star-QAM, we observe these distances at least two elements of the codeword pairs since the minimum Hamming distance between the index symbols related to the amplitude levels and disjoint phase sets is two. Thus, the MED of the codebook is limited by the MEDs $\sqrt{2}\tilde{d}_1$, $\sqrt{2}\tilde{d}_2$, and $\sqrt{2}\tilde{d}_3$. It can also be shown that d_1 does not yield the MED as $d_1 \geq \sqrt{2}\tilde{d}_1$.

In general case where we have Z amplitude levels, J disjoint phase sets, and M phases, we can write $r_z = \sqrt{\frac{2z}{Z+1}}$ and

$$\tilde{d}_1 = \frac{2\sqrt{2}}{\sqrt{Z+1}} \sin\left(\frac{\pi}{JM}\right). \tag{5.35}$$

In this case, $\tilde{d}_2$ becomes the MED between points in the innermost circles and can be calculated by

$$\tilde{d}_2 = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\left(\frac{\pi}{JM}\right)} = \sqrt{\frac{6 - 4\sqrt{2}\cos(\pi/JM)}{Z+1}}. \tag{5.36}$$

On the other hand, although rotating circles by π/JM maximizes the distance between the points of consecutive circles, it minimizes the distance between the points of κth and $(\kappa+2)$th circles where $\kappa \in \{1, 2, \ldots, Z-2\}$ since the constellation points on these circles will have the same phase due to the rotation. It can be shown that the MED between these circles is observed when $\kappa = Z-2$, thus such a distance can be given by

$$\tilde{d}_3 = \sqrt{\frac{2}{Z+1}}\left(\sqrt{Z} - \sqrt{Z-2}\right). \tag{5.37}$$

By using $\tilde{d}_1$, $\tilde{d}_2$, and $\tilde{d}_3$, the MED of the codebook can be lower-bounded by

$$d_2 \geq \min\left\{\sqrt{2}\tilde{d}_1, \sqrt{2}\tilde{d}_2, \sqrt{2}\tilde{d}_3\right\}. \tag{5.38}$$

The inequality arises from the fact that the MDS code may not allow us to observe the distances $\tilde{d}_1$, $\tilde{d}_2$, and $\tilde{d}_3$ in two elements of the codeword pairs.

5.4.3.4 MDS-IQM

In MDS-IQM, the codeword pairs that have the minimum Hamming distance are the M-ary symbols in each disjoint PAM constellation. The MED between these symbols can be given by [1]

$$d_1 = \frac{\sqrt{6}}{\sqrt{M^2 - \xi^{-2}}} \sim \frac{\sqrt{6}}{M} \tag{5.39}$$

where $\xi = \max\left\{R, T\right\}$.

The MED between the constellation points of MDS-IQM is equivalent to the MED between the symbols of a ξM-ary PAM constellations as the R-ary and T-ary PAM constellations form the codebook of MDS-IQM. Since such a distance is observed in at least two elements of the codeword pairs due to the MDS mechanism, the related MED can be calculated by multiplying the MED of a ξM-ary PAM constellation by $\sqrt{2}$, i.e.,

$$d_2 = \frac{2\sqrt{3}}{\sqrt{(\xi M)^2 - 1}}. \tag{5.40}$$

Remark. One can find an MDS scheme that achieves higher d_1 and d_2 than the techniques employing multiple disjoint constellations such as DM-IM, MM-IM, and SPM. That is because the MDS technique can achieve the same SE as those of these techniques with a lower number and smaller size of disjoint constellations. Moreover,

the MDS approach is much more flexible than those schemes as it is applicable for a broader range of the number of disjoint constellations and codeword length.

5.5 Numerical Results

In this section, we provide numerical BER results along with numerical achievable rate and distance results. Moreover, we compare these results with those of the benchmark techniques. We also confirm the theoretical results by computer simulation results in this section.

In tables and figures of this section, "MDS-APM (N, Z, J, M)" stands for the proposed MDS-APM concept with N channels in each block, Z amplitude levels, J phase sets, and M phases in each phase set, whereas "MDS-IQM (N, R, T, M)" is the proposed MDS-IQM scheme with N channels in each channel group with R and T disjoint M-PAM constellations related to the in-phase and quadrature components, respectively. "MDS-PSK/QAM (N, Q, M)" represents the MDS-PSK/QAM scheme that employs Q disjoint M-PSK/QAM constellations over N channels. We denote the special cases of the proposed schemes as "MDS-PSK/QAM $(N, Q, 1)$", "MDS-APM $(N, Z, J, 1)$", and "MDS-IQM $(N, R, T, 1)$". We do not embed information into the conventional modulation symbols for these special cases, thus each disjoint constellation has only one element. Moreover, "IM $(N, K, M$-PSK/QAM$)$" and "CI-IM $(N, K, M$-PSK/QAM$)$" stand for the conventional IM and CI-IM[14] techniques, respectively, which activate K codeword elements out of N elements and employ M-PSK/QAM symbols on the activated elements. "IQ-IM (N, K, M)" is the IQ-IM technique that has K nonzero in-phase and quadrature components out of N elements and chooses each nonzero element from an M-ary PAM constellation. "ICM (N, K, I, M)" is the ICM concept that activates K channels out of N, uses the com-

[14]We rotate the constellation points for the CI-IM techniques by the optimum angles provided in [75].

positions of an integer I with K parts to embed information into the energies of the activated channels, and carries M-PSK symbols on them, whereas, "ICM (N, K, I)" signifies the special case of ICM that embeds information into only indices of channels and compositions of an integer I with K parts. "OFSPM (N, M-PSK/QAM)" and "MM-IM (N, M-PSK/QAM)" stand for the OFSPM and MM-IM techniques with N channels in each channel block and N disjoint M-PSK/QAM constellations to specify the set partitions and permutations for former and latter, respectively. Moreover, "MM-IQ-IM (N, M)" is the MM-IQ-IM scheme of [1] with in-phase and quadrature components of N channels employing N disjoint M-PAM constellations. Finally, "M-PSK/QAM" signifies the conventional M-PSK/QAM.

The simulations related to the optimum and suboptimum detectors are executed based on the assumption that channel estimation is perfect. We further assume that the average channel energy is unity for the simulated systems in this section.

5.5.1 Uncoded Bit Error Rate

In this subsection, we provide uncoded BER results to show the effectiveness of the proposed MDS methods against the combinatorial and conventional modulation techniques. We divide this subsection into two parts to conduct BER comparisons. The first part demonstrates BER comparisons where the receivers of the proposed and benchmark methods employ the optimum ML detector. Besides, in the second part, we present and compare low-complexity detector performance for these methods.

5.5.1.1 The Optimum ML Performance

In Fig. 5.9, we compare the BER performance of the proposed MDS-APM and MDS-IQM schemes with those of the ICM, IM, and M-PSK benchmarks for an SE equal to 1 bpcu. As seen from the figure, unlike the BPSK and IM schemes, the proposed techniques yield a diversity gain equal to two due to the sole presence of

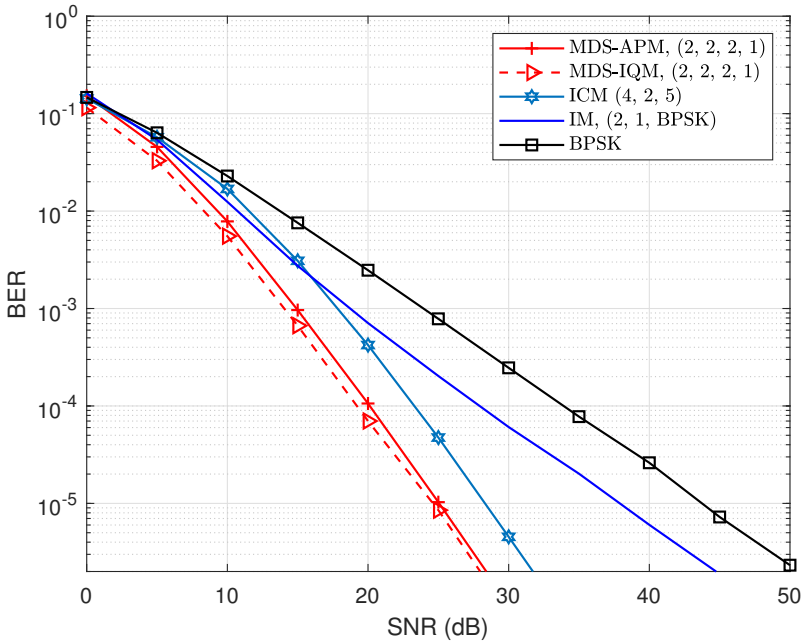


Figure 5.9. BER comparison of MDS-APM (2, 2, 2, 1) and MDS-IQM (2, 2, 2, 1) with ICM (4, 2, 5), IM (2, 1, BPSK), and BPSK.

index symbols. They considerably outperform the BPSK and IM concepts for all values of SNR illustrated in the figure. They also outperform the ICM scheme whose BER curve achieves the same diversity gain as the proposed schemes. These results confirm that MDS-APM and MDS-IQM are promising concepts for cell-edge users in cellular networks since the deployment of low order modulation is of paramount importance to achieve satisfactory error performance for cell-edge users [127].

Fig. 5.10 compares the BER performance of the proposed systems, MDS-PSK $(4, 4, 2)$ and MDS-PSK $(4, 8, 1)$, with OFSPM $(4, \text{BPSK})$, MM-IM $(4, \text{BPSK})$, IM $(4, 3, \text{QPSK})$, and QPSK. The SEs of the simulated systems range between 2 bpcu and 2.5 bpcu as shown in the legend of Fig. 5.10. MDS-PSK $(4, 8, 1)$ scheme is capable of achieving the same SE as OFSPM $(4, \text{BPSK})$ by employing only index symbols, and it considerably outperforms the other concepts by introducing an additional diversity gain. Moreover, MDS-PSK $(4, 4, 2)$ outperforms OFSPM $(4, \text{BPSK})$ and MM-IM $(4, \text{BPSK})$ slightly, and IM $(4, 3, \text{QPSK})$ and QPSK considerably at high

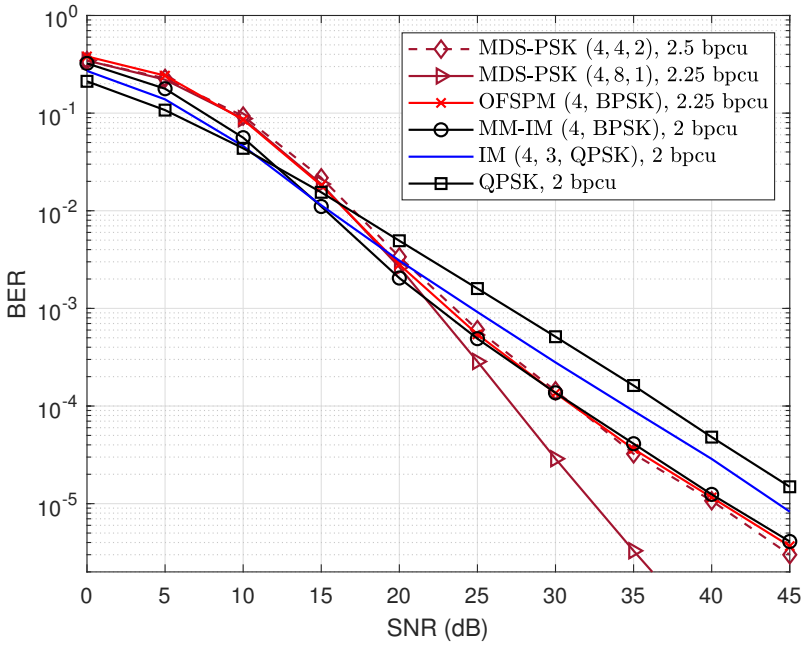


Figure 5.10. BER comparison of MDS-PSK $(4, 4, 2)$ and MDS-PSK $(4, 8, 1)$ with OFSPM $(4, \text{BSPK})$, MM-IM $(4, \text{BSPK})$, IM $(4, 3, \text{QPSK})$, and QPSK.

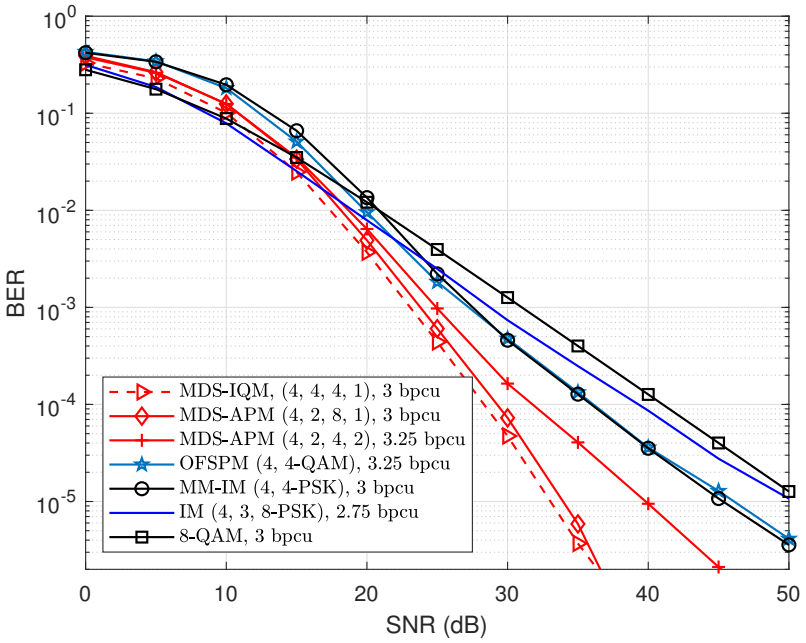


Figure 5.11. BER comparison of MDS-IQM (4, 4, 4, 1), MDS-APM (4, 2, 8, 1), and MDS-APM (4, 2, 4, 2) with OFSPM (4, 4-QAM), MM-IM (4, 4-PSK), IM (4, 3, 8-PSK), and 8-QAM.

SNR although it achieves the highest SE.

Fig. 5.11 compares the BER performance of the proposed MDS-IQM and MDS-APM techniques with those of OFSPM, MM-IM, IM, and 8-QAM. The SEs of the simulated systems range between 2.75 bpcu and 3.25 bpcu. Here, MDS-IQM (4, 4, 4, 1) and MDS-APM (4, 2, 8, 1) achieve a higher diversity order than the benchmarks and outperform them for a wide range of SNR. Although MDS-APM (4, 2, 4, 2) is not capable of achieving a higher diversity gain than those of the benchmarks, it considerably outperforms them since it achieves a similar SE to those of the benchmarks by producing a higher number of index symbols and employing a lower modulation order than them.

5.5.1.2 The Low-complexity Detector Performance

In this subsection, we compare the BER performance of the proposed schemes with those of the existing IM benchmarks by considering the low-complexity detectors of these schemes. Thus, in our simulations, we employ the LC-ML detector for the proposed concepts as well as the LLR detector in [2] for the IM concepts, in [1] for the MM-IM and MM-IQ-IM concepts, and the LC-ML detector in Section 4.3.3.2 for the ICM concept. We further provide BER results related to the optimum ML detector of the conventional modulation scheme for benchmarking. Also, we compare the performance of the LC-ML detector with that of the ML detector for the proposed concepts.

In Fig. 5.12, we compare the BER performance of the proposed schemes, MDS-QAM $(4, 8, 2)$, MDS-PSK $(4, 8, 2)$, and MDS-QAM $(4, 16, 1)$, with that of OFSPM $(4,$ 4-QAM), OFSPM $(4,$ QPSK), MM-IM $(4,$ QPSK), IM $(4, 3,$ 8-PSK), and 8-PSK methods. Here, except for MDS-QAM $(16, 4, 1)$, the MDS and OFSPM schemes exhibit the same SE of 3.25 bpcu whereas the remaining concepts, except for IM, have the same SE of 3 bpcu. Also, the SE of IM $(4, 3,$ 8-PSK) is 2.75 bpcu. As seen from the figure, MDS-QAM $(4, 8, 2)$ and MDS-PSK $(4, 8, 2)$ achieve an outstanding

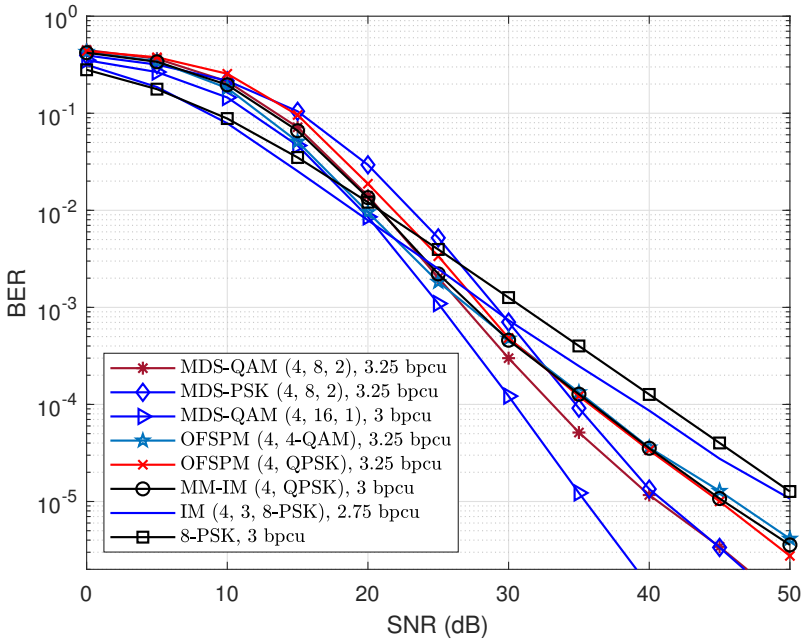


Figure 5.12. BER comparison of MDS-QAM (N, Q, M) and MDS-PSK (N, Q, M) with OFSPM $(4, 4\text{-QAM})$, OFSPM $(4, \text{QPSK})$, MM-IM $(4, \text{QPSK})$, IM $(4, 3, 8\text{-PSK})$ and 8-PSK for $N = 4, Q \in \{8, 16\}, M \in \{1, 2\}$.

BER performance at high SNR and outperform the OFSPM and MM-IM schemes as well as the IM and 8-PSK techniques by providing almost 5 dB and 10 dB SNR gains, respectively, at a BER value of 10^{-5}. These results arise from the fact that the MDS concepts achieve a similar SE as those of the other concepts by employing a lower order modulation. Such an order reduction results in higher MED for the modulation symbols as well as better error performance at high SNR. Moreover, MDS-QAM $(4, 16, 1)$ achieves the best BER performance with an SNR gain of more than 10 dB at a BER value of 10^{-5} compared to the OFSPM and MM-IM methods. Such a scheme achieves an additional diversity order due to the index symbols, whereas the diversity orders of the benchmark techniques are limited by the minimum Hamming distance between the conventional modulation symbols.

In Fig. 5.13, we compare the low-complexity detection performance of the proposed systems with those of the MM-IQ-IM, MM-IM, ICM, and IM systems when $\eta = 4$ bpcu. As seen from the figure, although the IM benchmarks slightly outperform

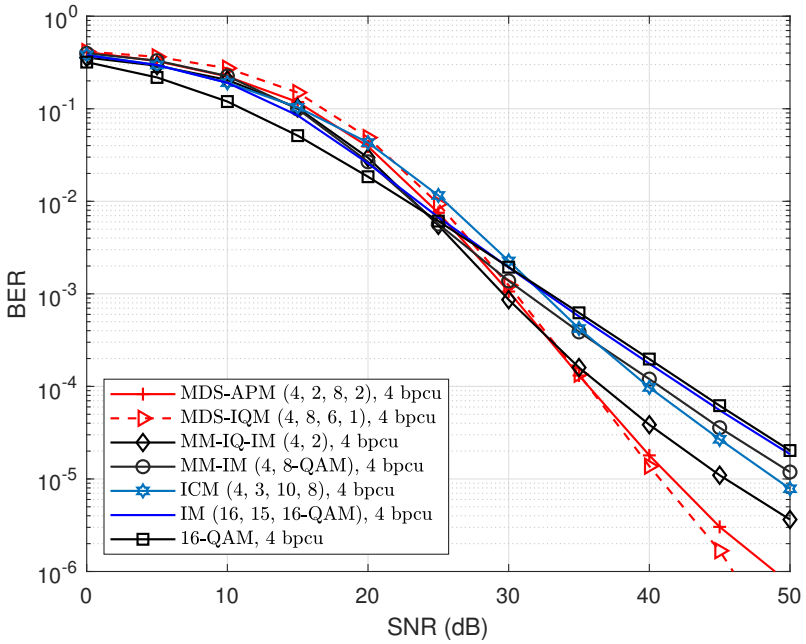


Figure 5.13. BER comparison of MDS-APM (4, 2, 8, 2) and MDS-IQM (4, 8, 6, 1) with MM-IQ-IM (4, 2), MM-IM (4, 8-QAM), ICM (4, 3, 10, 8), IM (16, 15, 16-QAM), and 16-QAM.

the proposed concepts at low SNR, our proposals begin to surpass the benchmarks when we increase SNR. Moreover, the SNR loss of the proposed schemes compared to the 16-QAM method is not slight at low SNR like the IM benchmarks. However, MDS-APM (4, 2, 8, 2) and MDS-IQM (4, 8, 6, 1) considerably outperform 16-QAM at mainly high SNR. The behavior of the BER curves at low and high SNR can be explained by the MED between the codeword pairs and the MED between the minimum Hamming distance codeword pairs, respectively. The presence of index symbols for the proposed schemes decreases the overall MED between the codeword pairs, leading to an SNR loss at low SNR. However, it increases the MED between the minimum Hamming distance codeword pairs as the codebook structure enables us to perform a lower order modulation on the channels. The proportion of index symbols to the conventional modulation symbols is higher in the proposed schemes than in the IM benchmarks. That explains the BER improvement compared to these benchmarks

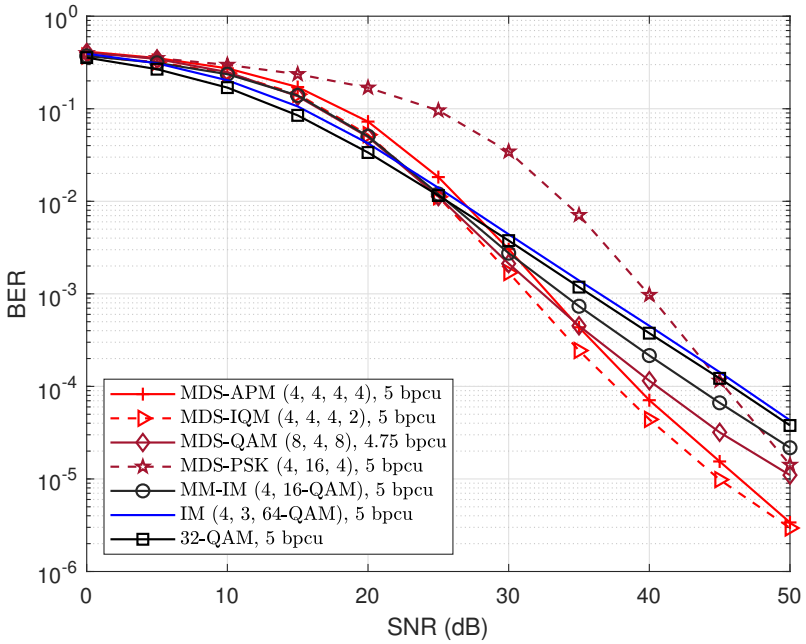
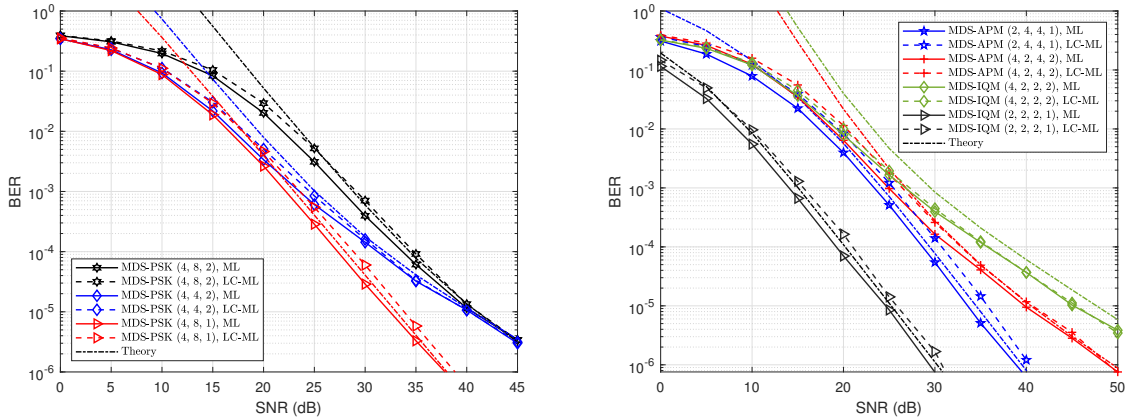


Figure 5.14. BER comparison of MDS-APM (4, 4, 4, 4) and MDS-IQM (4, 4, 4, 2) with MDS-QAM (8, 4, 8), MDS-PSK (4, 16, 4), MM-IM (4, 16-QAM), IM (4, 3, 64-QAM), and 32-QAM.

at high SNR. As a result, the proposed codebooks yield a trade-off between error performance at low and high SNR.

In Fig. 5.14, we compare the low-complexity detection performance of the proposed schemes with those of the MM-IM and IM concepts. All concepts provide an SE of 5 bpcu except for MDS-QAM (8, 4, 8), which achieves an SE of 4.75 bpcu. Like Fig. 5.13, the performance of the proposed concepts is superior to those of the IM and classical benchmarks at high SNR. Different than Fig. 5.13, we provide simulation results related to the MDS-QAM and MDS-PSK systems. The use of amplitude levels in addition to the phase sets enables MDS-APM (4, 4, 4, 4) to achieve a considerably better error performance than MDS-PSK (4, 16, 4) which incorporates only phase sets for information encoding. Moreover, both MDS-APM (4, 4, 4, 4) and MDS-IQM (4, 4, 4, 2) outperform MDS-QAM (8, 4, 8) despite they achieve a higher SE than MDS-QAM (8, 4, 8).



(a) MDS-PSK (b) MDS-IQM and MDS-APM

Figure 5.15. BER comparison of the LC-ML detector with the optimum ML detector for the MDS-PSK, MDS-IQM, and MDS-APM systems.

In Fig. 5.15, we compare the BER performance of the LC-ML detector with that of the optimum ML detector for the proposed schemes with different values of SE. The figure demonstrates that the performance of the LC-ML algorithm is considerably close to that of the ML detector, especially at high and low SNR. Although the LC-ML detection results in a slight SNR loss at medium SNR, it decreases the complexity substantially compared to the optimum ML detection. Also, the curves labeled "Theory" illustrate theoretical upper-bound results for the average BER. As we observe from the figure, the upper-bound curves are consistent with computer simulations, especially at high SNR.

5.5.2 Coded Bit Error Rate

In Fig. 5.16, we compare the BER performance of the MDS-IQM schemes employing the mapping in Section 5.1.1 and the Gray-like mapping described in Section 5.3.1. Here, the BER results correspond to the LC-ML detector of the MDS-IQM schemes. As seen from the figure, the systems with Gray-like mapping (solid curves) achieve lower BER than the ones with the mapping in Section 5.1.1 (dashed curves). More-

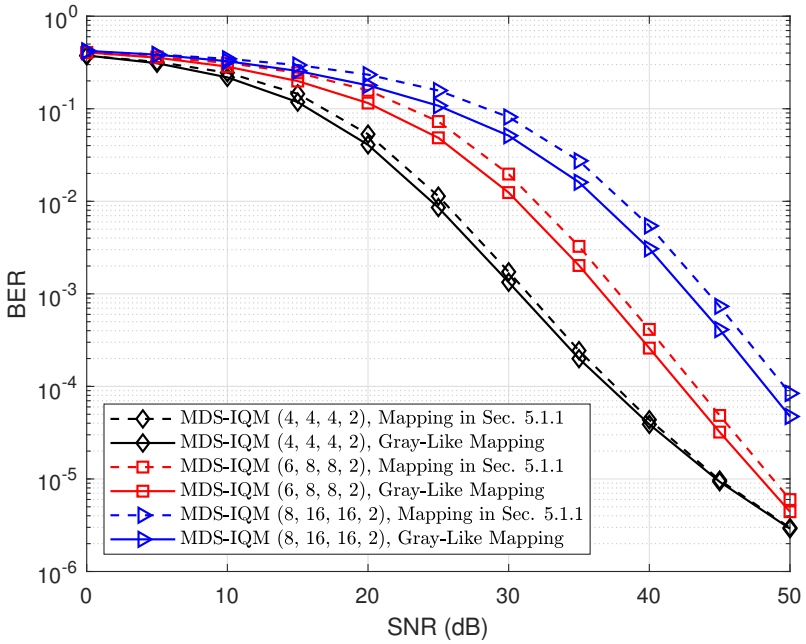


Figure 5.16. BER comparison of the MDS-IQM schemes employing the mapping in Section 5.1.1 and the Gray-like mapping.

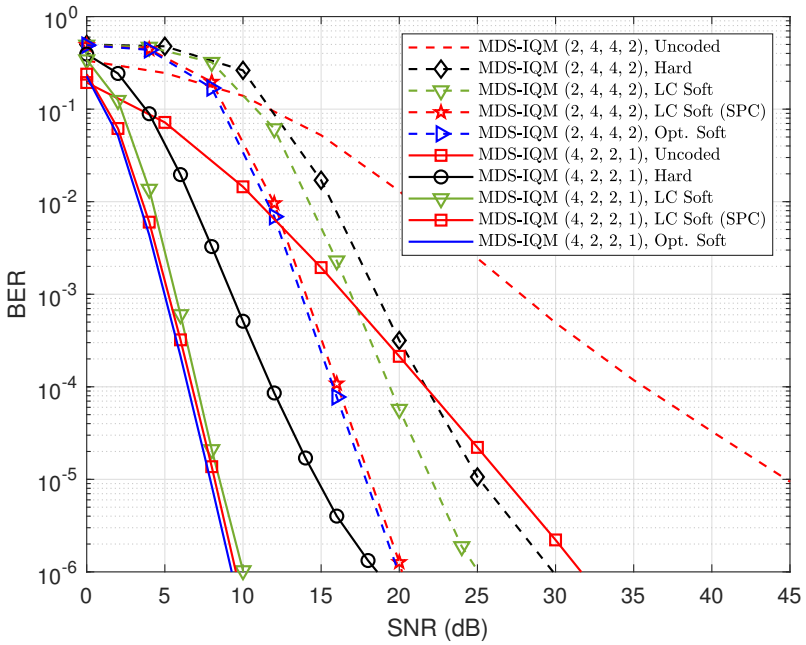


Figure 5.17. BER comparison of the uncoded and coded MDS-IQM schemes.

over, the gap between the BER curves increases when we increase the number of N-tuples by increasing the number of disjoint PAM sets, i.e., R and T.

In Fig. 5.17, we compare the BER performance of uncoded and coded MDS-IQM techniques.[15] For the coded schemes, we employ a rate 1/2 convolutional code with generator matrix $G(D) = [1 + D + D^2 + D^3 + D^6, 1 + D^2 + D^3 + D^5 + D^6]$. Here, the BER curves of the uncoded schemes are obtained by employing the optimum ML detector, whereas those of the coded schemes with labels "Hard", "LC Soft", and "Opt. Soft" are obtained by performing hard-decision decoding based on the optimum ML detector, soft-decision decoding based on the low-complexity LLR values in (5.19)-(5.20), and based on the optimum LLR values in (5.16), respectively. Moreover, we obtained the coded schemes with the label "LC Soft (SPC)" by employing the equivalent SPC code and performing low-complexity LLR values in (5.19)-(5.22). As seen from the figure, the channel encoding improves the BER performance of the MDS-IQM techniques remarkably. The system performing only (5.19) and (5.20) suffers from significant performance loss compared to the optimum system when especially $N = 2$. On the other hand, the performance of the soft-decision decoder using the LLR values in (5.19)-(5.22) is very close to that of the one using optimum LLR values in (5.16). Such a result is critical for the MDS modulation techniques as the complexity of calculating (5.19)-(5.22) is considerably lower than that of the optimum decoder. It also makes the MDS modulation techniques more suitable to channel encoding than the IM techniques as it is not straightforward to calculate reliable LLR values for the IM techniques.

In Fig. 5.18, we compare the coded BER performance of the MDS methods with that of the combinatorial methods. For the MDS systems, we employ the equivalent SPC code along with soft-decision decoding based on the low-complexity LLR values given in (5.19)-(5.22). For the remaining systems, we exploit the optimum soft-decision decoding based on Eqs. (9) and (11) in [125]. Here, all systems employ a rate

[15]It is important to note that we do not employ interleaving between the codeword elements of the MDS techniques related to Fig. 5.17 since we assume the channel fading coefficients are changing independently over each symbol interval. However, for block fading channels, one should employ interleaving to yield optimum coded performance.

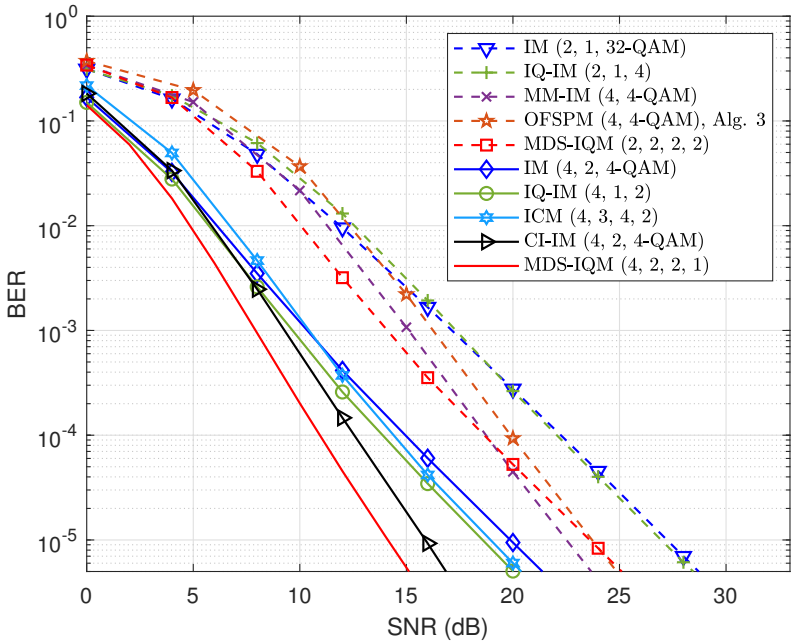
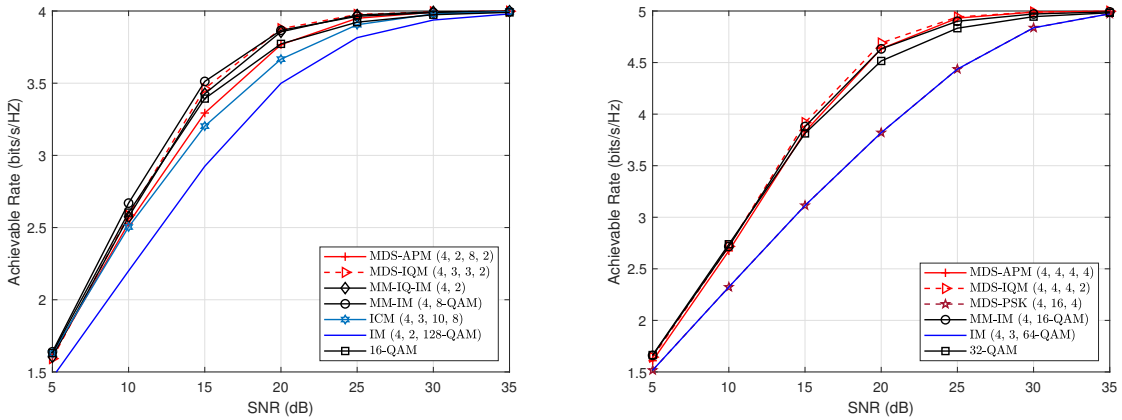


Figure 5.18. Coded BER comparison of the MDS-IQM and combinatorial techniques.

1/2 convolutional code with generator matrix $G(D) = [1+D^2, 1+D+D^2]$. Moreover, we adjust the system parameters to achieve the same level of SE for the benchmark schemes, and we used solid and dashed curves to indicate the results related to the systems achieving the same level of SE. As seen from Fig. 5.18, although the MDS concepts execute a suboptimum detection with much less complexity per bit, they outperform the combinatorial concepts and exhibit a comparable performance to that of the MM-IM and OFSPM techniques in the presence of channel encoding.

5.5.3 Achievable Rate

As discussed in Section 2.5.1.2, to gain a practical insight into the achievable rate as well as the coded performance, one can use the same approach in [76] and compute the achievable rate numerically by assuming equally likely codewords and evaluating (2.49). In this subsection, we use this approach to compare the achievable rate of the proposed and benchmark concepts.



(a) $\eta = 4$ bpcu (b) $\eta = 5$ bpcu

Figure 5.19. Achievable rate comparison of the proposed methods with the combinatorial and conventional concepts.

In Figs. 5.19a and 5.19b, we compare the achievable rate performance of the proposed MDS coded modulation techniques with that of the combinatorial and conventional methods when the SE is 4 bpcu and 5 bpcu, respectively. As seen from the figures, the proposed MDS-IQM schemes outperform all the benchmarks, except for the MM-IM scheme when the SE is 4 bpcu, in terms of achievable rate for a wide range of SNR. Moreover, although the MDS-APM scheme is outperformed by the MM-IM and 16-QAM schemes for the SE of 4 bpcu, it performs closely to the MM-IM scheme and exhibits superior performance compared to the MDS-PSK, IM, and 32-QAM methods when the SE is 5 bpcu. It is also important to note that one can improve the performance of the MDS-APM technique by carefully choosing the ring ratios and rotation angles related to the constellation of this technique.

5.5.4 Minimum Codeword Distance

This subsection compares the minimum codeword distances for the MDS and IM methods. We exploit tables to demonstrate the minimum L-product distance, d_1, and MED, d_2, of different modulation concepts. We use bold characters in our tables

Table 5.7. Minimum codeword distance comparison of the MDS modulation techniques with the QAM, IM, and IQ-IM techniques when $N = 2$.

SE per codeword element (η)	System	Diversity Order (L)	d_1	d_2
1 bpcu	BPSK	1	2.0000	2.0000
	IM (2, 1, BPSK)	1	2.8284	2.0000
	MDS-QAM (2, 4, 1)	2	2.0000	2.0000
2 bpcu	4-QAM	1	1.4142	1.4142
	IM (2, 1, 8-QAM)	1	1.1547	1.1547
	IQ-IM (2, 1, 2)	1	2.0000	1.4142
	MDS-QAM (2, 4, 2)	1	1.8257	1.1547
	MDS-IQM (2, 4, 4, 1)	2	0.4000	0.8944
3 bpcu	8-QAM	1	0.8165	0.8165
	IM (2, 1, 32-QAM)	1	0.6325	0.6325
	IQ-IM (2, 1, 4)	1	0.8944	0.6325
	MDS-IQM (2, 2, 2, 2)	1	1.2649	0.8944
4 bpcu	16-QAM	1	0.6325	0.6325
	IM (2, 1, 128-QAM)	1	0.3123	0.3123
	IQ-IM (2, 1, 8)	1	0.4364	0.3086
	MDS-IQM (2, 4, 4, 2)	1	1.2344	0.4364

to emphasize the methods with the highest d_1 among benchmarks achieving the same diversity order. We further make numerical BER comparisons between the MDS and IM techniques to support the distance comparisons.

In Table 5.7, we compare the minimum codeword distances of MDS-IQM and MDS-QAM with conventional modulation, IM, and IQ-IM when $N = 2$. It is clear from the table that the MDS concepts provide higher d_1 and d_2 than the IM and IQ-IM benchmarks when the SE per codeword element is higher than 2 bpcu. However, in the case where we have $\eta = 1$ bpcu and $\eta = 2$ bpcu, one can find MDS schemes that achieve a higher diversity order than the IM and IQ-IM schemes. Moreover, the MDS techniques outperform conventional modulation in terms of d_1 for almost all values of SE. As a result, the MDS concepts can attain a superior error performance against the benchmarks. Such superiority is best illustrated with a BER figure.

In Fig. 5.20, we compare the BER performance of the MDS concepts with con-

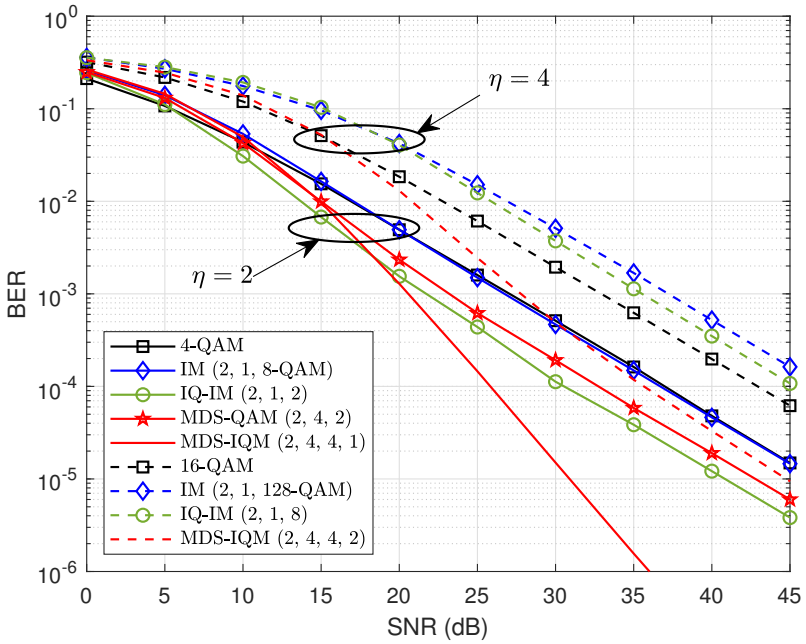


Figure 5.20. BER comparison of the MDS-IQM and MDS-QAM systems with the QAM, IM, and IQ-IM techniques when $N = 2$ and $\eta \in \{2, 4\}$.

ventional modulation, IM, and IQ-IM when $N = 2$ and $\eta \in \{2, 4\}$. To obtain the BER curves, we employ the optimum ML detectors for all concepts as in Eq. (10) of [2]. As seen from the figure, IQ-IM (2, 1, 2) outperforms MDS-QAM (2, 4, 2) at both low and high SNR, since the IQ-IM scheme has better distance properties, or in other words, higher d_1 and d_2, than the MDS-QAM scheme as shown in Table 5.7. Although MDS-IQM (2, 4, 4, 1) is outperformed by 4-QAM, IM (2, 1, 8-QAM), and IQ-IM (2, 1, 2) at low SNR due to its relatively low d_2, it considerably outperforms these benchmarks at high SNR as it does achieve an additional diversity gain. On the other, when $\eta = 4$, the MDS-IQM technique outperforms the IM and IQ-IM techniques at both low and high SNR because MDS-IQM (2, 4, 4, 2) has higher d_1 and d_2 than IM (2, 1, 128-QAM) and IQ-IM (2, 1, 8).

In Tables 5.8 and 5.9, we compare the distance results of the MDS schemes with those of the QAM, IM, IQ-IM, ICM,[16] and CI-IM techniques for $N = 4$ and $N = 8$,

[16]Note that the ICM constellation has a similar structure to that of star-QAM where the energies or regarding amplitude levels of the ICM symbols can be viewed as rings/circles like in star-QAM.

Table 5.8. Minimum codeword distance comparison of the MDS techniques with the QAM, IM, IQ-IM, ICM, and CI-IM techniques when $N = 4$.

Groups	System	SE per Codeword Element (η)	Diversity Order (L)	d_1	d_2
Group 1	4-QAM	2	1	1.4142	1.4142
	IM (4, 3, 4-QAM)	2	1	1.6330	1.6330
	IQ-IM (4, 3, 2)	2.5	1	1.6330	1.1547
	ICM (4, 3, 6, 2)	2	1	1.6330	0.7655
	MDS-QAM (4, 4, 2)	2.5	1	1.8257	1.1547
	CI-IM (4, 2, 16-QAM)	2.5	2	0.1169	0.4880
	MDS-IQM (4, 3, 3, 1)	2	2	0.7500	1.2247
Group 2	8-QAM	3	1	0.8165	0.8165
	IM (4, 3, 8-QAM)	2.75	1	0.9428	0.9428
	IQ-IM (4, 2, 4)	3	1	0.8944	0.6325
	ICM (4, 3, 8, 4)	3	1	1.0000	0.5464
	MDS-QAM (4, 8, 2)	3.25	1	1.7889	0.8944
	CI-IM (4, 2, 64-QAM)	3.5	2	0.0149	0.1954
	MDS-QAM (4, 32, 1)	3.75	2	0.2000	0.6325
Group 3	32-QAM	5	1	0.4472	0.4472
	IM (4, 3, 64-QAM)	5	1	0.3563	0.3563
	IQ-IM (4, 2, 16)	5	1	0.2169	0.1534
	ICM (4, 3, 13, 16)	5	1	0.2164	0.2164
	MDS-IQM (4, 4, 4, 2)	5	1	1.2344	0.4364
	CI-IM (4, 2, 512-QAM)	5	2	-	-
	MDS-IQM (4, 12, 12, 1)	5	2	0.0420	0.2897

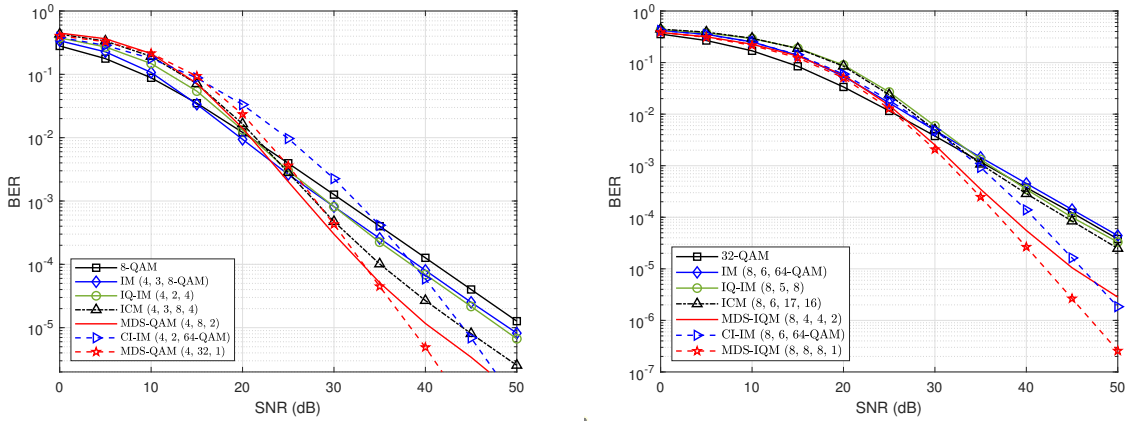
respectively. In these tables, we group the systems in a way that those with similar SE would be in the same group. As seen from the tables, one can find an MDS system that achieves a higher d_1 than the benchmarks for any group except for Group 1 in Table 5.9 where IQ-IM (8, 4, 2) attains higher d_1 and d_2 than MDS-QAM (8, 4, 2). Moreover, the gap between d_1 related to the MDS and benchmark techniques increases in favor of the MDS techniques when the SE increases. In this case, the MDS methods are expected to achieve increasingly better asymptotic error performance compared to the benchmark methods. On the other hand, we observe from the tables that the MDS approach exhibits a d_2 which is comparable to that of the QAM techniques and

Thus, for the ICM techniques in this subsection, we rotate the consecutive circles by $\frac{\pi}{M}$ to increase the MED between the constellation points.

Table 5.9. Minimum codeword distance comparison of the MDS techniques with the QAM, IM, IQ-IM, ICM, and CI-IM techniques when $N = 8$.

Groups	System	SE per Codeword Element (η)	Diversity Order (L)	d_1	d_2
Group 1	4-QAM	2	1	1.4142	1.4142
	IM (8, 6, 4-QAM)	2	1	1.6330	1.6330
	IQ-IM (8, 4, 2)	2.5	1	2.0000	1.4142
	ICM (8, 4, 10, 4)	2.5	1	1.2649	0.6375
	MDS-QAM (8, 4, 2)	2.5	1	1.8257	1.1547
	CI-IM (8, 6, 4-QAM)	2	2	0.6667	1.1547
	MDS-IQM (8, 3, 3, 1)	2.75	2	0.7500	1.2247
Group 2	8-QAM	3	1	0.8165	0.8165
	IM (8, 6, 16-QAM)	3.5	1	0.7303	0.7303
	IQ-IM (8, 4, 4)	3.5	1	0.8944	0.6325
	ICM (8, 6, 10, 8)	3.5	1	0.6846	0.6846
	MDS-IQM (8, 2, 2, 2)	3.75	1	1.2649	0.8944
	CI-IM (8, 6, 16-QAM)	3.5	2	0.0780	0.3984
	MDS-IQM (8, 4, 4, 1)	3.5	2	0.4000	0.8944
Group 3	32-QAM	5	1	0.4472	0.4472
	IM (8, 6, 64-QAM)	5	1	0.3563	0.3563
	IQ-IM (8, 5, 8)	5	1	0.3904	0.2760
	ICM (8, 6, 10, 8)	5	1	0.2677	0.2677
	MDS-IQM (8, 4, 4, 2)	5.5	1	1.2344	0.4364
	CI-IM (8, 6, 64-QAM)	5	2	0.0099	0.1595
	MDS-IQM (8, 8, 8, 1)	5.25	2	0.0952	0.4364

higher than that of the IM techniques except for Group 1. These results demonstrate the clear advantage of such an approach against the IM method in error performance for relatively high SEs. Although the IM method seems more beneficial for low SEs, the MDS approach still achieves the best asymptotic error performance at low SEs by exhibiting additional diversity gain compared to the IM methods and better distance properties than the CI-IM methods. Note that we do not provide d_1 and d_2 for CI-IM (4, 2, 512-QAM) in Table 5.8 as the searching space for the optimum rotation angle related to this system is exhaustive. Note also that we provide upper-bound distance results on d_2 for the ICM techniques in Groups 2 and 3 since it is not straightforward to calculate the MED for these techniques when the codebook size is large.



(a) $N = 4$ (b) $N = 8$

Figure 5.21. BER comparison of the MDS modulation techniques with the QAM, IM, IQ-IM, ICM, and CI-IM techniques.

In Figs. 5.21a and 5.21b, we compare the BER performance of the MDS-QAM and MDS-IQM systems with that of the QAM, IM, IQ-IM, ICM, and CI-IM systems for $N = 4$ and $N = 8$, respectively. We employ the low-complexity detectors of the IM and MDS techniques to obtain the corresponding BER curves. In this regard, we adopt LLR detectors in [2], [36], and [75] for the IM, IQ-IM, and CI-IM schemes, respectively, whereas we operate the LC-ML detectors of the ICM and MDS systems. As seen from Fig. 5.21a, MDS-QAM (4, 8, 2) exhibits the best error performance at medium-to-high SNR among the systems having unit diversity order. That verifies the results of Table 5.8 as MDS-QAM (4, 8, 2) has the highest d_1 in Group 2 of the table. One can also expect MDS-IQM (4, 8, 2) to achieve better BER performance at low SNR than most IM techniques as it has the second-highest d_2 in the group. However, this is not the case due to the suboptimal performance of the LC-ML detector and the number of codeword pairs having the MED, i.e., d_2. On the other hand, Fig. 5.21b demonstrates the superiority of the MDS techniques where the MDS techniques outperform the IM techniques at almost all SNR values.

In Fig. 5.22, we conduct more general distance comparisons among the MDS and IM methods. We also provide distance results related to the QAM method for

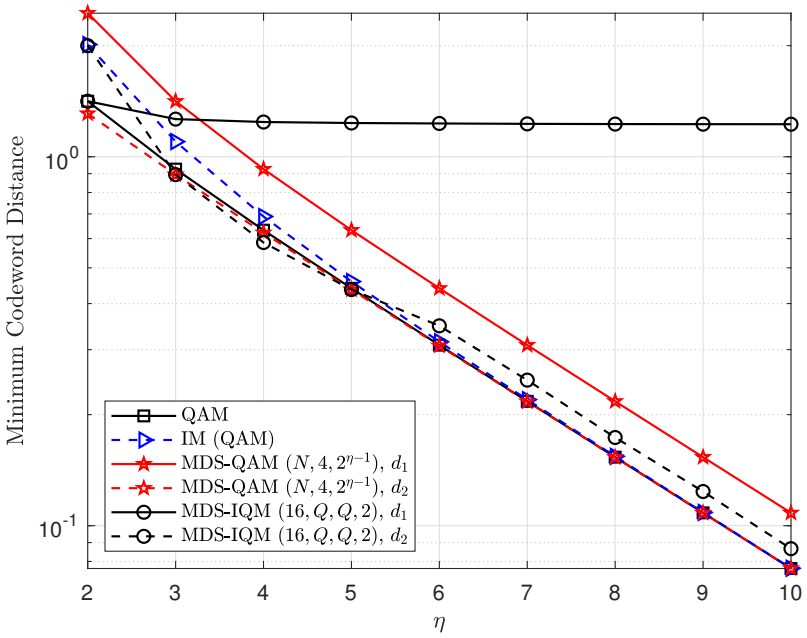


Figure 5.22. Minimum codeword distance vs. η for the MDS modulation, QAM, and IM techniques.

benchmarking. Here, the distance results are provided as a function of η, i.e., SE. For the IM technique, we assume QAM symbols employed by the activated codeword elements and use the upper-bound on the minimum L-product distance calculated in [76]. Here, we do not provide any numerical results for the MED of the IM method as it is equal to the L-product distance for most values of SE. For MDS modulation, we consider two extreme cases where M is large and small, i.e., $M = 2^{\eta-1}$ and $M = 2$, respectively. MDS-QAM $(N, 4, 2^{\eta-1})$ can achieve a SE greater than or equal to those of other methods, and the distance results of MDS-QAM $(N, 4, 2^{\eta-1})$ are valid for any value of N. For MDS-IQM $(16, Q, Q, 2)$, we alter Q to achieve a target SE, i.e., $Q = \lceil 2^{\frac{N(\eta-2)}{N-1}} \rceil$. As seen from the figure, the MDS techniques can deliver considerably higher minimum L-product distance, d_1, than IM and QAM. Moreover, MDS-IQM $(16, Q, Q, 2)$ outperforms IM and QAM in terms of d_2 when $\eta \geq 6$ in addition to considerably exceeding other systems in terms of d_1.

Remark. The results presented in this section illustrated that the MDS mechanism is

quite useful when especially SE is high, and there is channel encoding in the system. A recent wireless local area network standard called IEEE 802.11ax, or in other words, Wifi 6, targets to achieve high SEs by deploying high order modulation such as 1024-QAM and employing forward error correction codes (convolutional or low-density parity-check codes) [128]. In this regard, one can use the MDS modulation to improve the performance of IEEE 802.11ax.

5.6 Summary

In this chapter, we proposed two novel modulation concepts that use constellation points to form the codewords of a simple MDS code. We showed that the proposed concepts are substantially different from the PM concepts, and they can produce more codewords and achieve a higher SE. To depict an example of practical application and conduct fair comparisons with the recent combinatorial techniques, we presented practical implementations of our modulation designs in parallel channels. For these implementations, efficient low-complexity decoding and very simple bits-to-symbols/symbols-to-bits mappings were shown possible. We further compared the detection complexity of the proposed techniques with those of the combinatorial and classical methods and demonstrated the effectiveness of our concepts against them. In addition, we touched upon implementation issues and presented simple ways to perform Gray coding and low-complexity LLR calculation for the MDS-coded methods. Then, we investigated the performance of the MDS-coded concepts in terms of SE, average BER, and minimum codeword distance. We provided extensive numerical results to prove the effectiveness of the MDS methods against the combinatorial and conventional methods. Our findings revealed that the proposed concepts are practical, flexible, and low-complexity modulation concepts that are capable of considerably outperforming the existing benchmarks.

As we listed above, the MDS methods have several beneficial properties. However, they also come with a couple of drawbacks. As we showed by the BER figures, the MDS methods may struggle to provide competitive error performance at relatively low SNR. Although this could be prevented by exploiting the channel encoding techniques, this may limit the application of the MDS techniques in the uncoded systems. Moreover, the low complexity detection of the MDS concepts is dependent on ordering the channel gains. Therefore, the low complexity detector may induce a poor error performance when the channel estimation is imperfect.

Chapter 6

Conclusion

6.1 Thesis Summary

Mobile data traffic per subscriber and the number of connected devices are expected to increase dramatically in the near future. As foreseen by [8], global mobile traffic and mobile traffic per subscription will increase approximately 80-fold and 50-fold, respectively, from 2020 to 2030. They will reach 5016 Exabytes and 257.1 Gigabytes per month. To fuel this ever-increasing demand in data communication and support the massive connectivity, the next-generation networks will need to deploy high-speed, reliable, and low-latency modulation techniques [129]. In this regard, IM has attracted remarkable attention as a promising modulation technique due to its attractive advantages involving better error performance and enhanced energy efficiency over conventional modulation techniques as well as low-complexity implementation. On the other hand, IM owes these properties to its combinatorial structure where combinations and permutations play a vital role to construct constellation points. In this thesis, motivated by the benefits of such a combinatorial structure, we uncovered two modulation concepts based on using different tools of combinatorics called set partitions and compositions. Then, we turned our attention to a well-known mech-

anism called MDS code and showed that modulation based on a simple MDS code could be more beneficial than the combinatorial modulation concepts.

In **Chapter 3**, we proposed a novel modulation concept called SPM. We used a set partitioning mechanism to form codewords in an SPM codebook. Depending on set partitioning and permutations of subsets in the set partitions, we designed four different variants of SPM. For these modulation concepts, we illustrated practical implementations that deploy disjoint constellations to specify the subsets of the set partitions. To overcome the prohibitive complexity of the optimum ML detection and improve the distance properties of the SPM codebooks, we proposed low-complexity detection and codebook selection algorithms. We exhibited the efficacy of these algorithms by complexity calculations and numerical results. We further investigated the data rate and error performance of the SPM techniques and provided related numerical and simulation results. These results proved that these techniques outperform the classical modulation and IM techniques in achievable rate and error performance.

In **Chapter 4**, we defined the basic idea of composition encoding and proposed novel modulation concepts called WCM, CM, and ICM that exploit weak and strict compositions of an integer to construct their codewords. In particular, WCM and CM operate over weak and strict compositions of an integer, respectively. The ICM method utilizes combinations of the activated codeword elements in addition to the strict compositions. We illustrated a practical model which exploits the proposed methods in parallel channels to embed information into the energies of well-known PSK symbols. We further showed that ICM is a more general concept that achieves higher SE than IM and encompasses IM and classic PSK modulation in particular cases. To overcome the complexity of the optimum ML detector, we designed a low-complexity detection method that achieves near-ML performance. In addition, we presented a simple bits-to-compositions mapping algorithm for the proposed techniques that eliminates the need for look-up tables. We investigated the BER, SE, and

mutual information of the proposed concepts. We provided numerical results verifying our theoretical derivations and proving the superiority of the proposed methods against the classical modulation and IM techniques.

In **Chapter 5**, we took a different approach and revisited a well-known MDS code for codebook design. Instead of using combinations, permutations, set partitions, and compositions of codeword elements, we used tuples of a simple MDS code to structure a modulation technique, *MDS modulation*. Unlike the vast majority of studies in the literature showing the effectiveness of IM against classical modulation techniques, we exhibited a different perspective that a well-known encoding mechanism is compared with the IM and PM concepts. Moreover, we implemented MDS modulation in parallel channels over the disjoint PSK, QAM, and PAM constellations. The MDS approach enabled us to show low-complexity bits-to-index symbols mapping, detection, Gray coding, and LLR calculations are possible for our implementations. We examined the performance in terms of the minimum codeword distance and BER. Our analytical and numerical results demonstrated that the MDS modulation techniques achieve considerably better codeword distance, error, and achievable rate performance than the IM and classical modulation concepts for a wide range of SE. More importantly, the proposed schemes prosper in keeping the detection complexity lower than the benchmark schemes while achieving better performance.

6.2 Future Directions

This thesis proposed reliable and spectrally efficient modulation techniques based on combinatorial and conventional encoding mechanisms. Although we showed that the proposed concepts provide substantial improvements compared to the existing methods, one can still enhance the performance. On the other hand, although the proposed modulation concepts have more desirable properties than the IM and clas-

sical modulation techniques, they still come with a few challenges. In this context, this section provides the points where one can improve the proposed approaches with further research efforts. We also deliver our perspectives in addressing these points.

6.2.1 Performance Enhancement

The proposed SPM and MDS schemes form their multidimensional codewords by using disjoint constellations. Moreover, the disjoint constellations are obtained by set partitioning of well-known PSK, QAM, and PAM constellations. Although exploiting this approach with the set partitioning and MDS mechanisms was proved to outperform existing schemes and provide promising error performance at low and high SNR, one can still improve the error performance by using the optimum disjoint constellations. In addition, one can increase the minimum codeword distances between the ICM codewords and improve the error performance by using the optimum energy levels and angular difference between these levels.

Finding the optimum two-dimensional constellations for a multidimensional codebook is an open and challenging problem that requires further research. Moreover, such constellations will be particular to the SNR value since different distance measures alter error performance at low and high SNR. A simple yet suboptimal solution to this problem is to use the optimum two-dimensional constellations, *hexagonal constellations* in [32], along with the optimum set partitioning of these constellations to obtain improved error performance for the proposed modulation techniques. On the other hand, one can define the upper-bounds on the BER as an objective function and try to minimize this function for a specific SNR value. One can also set the minimum codeword distances among the points of well-known QAM, PSK, and APSK constellations as a threshold and try to minimize average symbol energy as in [130]. Although the resulting optimization problem is a nonconvex quadratically constrained quadratic programming problem, it is possible to obtain good small to

medium-sized constellations by linearizing inequality constraints as shown in [130].

In the next-generation mobile networks, the battery life of the mobile devices will remain as a challenging issue since these networks will need to support several new applications and enabling technologies [131]. In this regard, improving the error performance of the proposed techniques may benefit the next-generation mobile networks as this may potentially decrease the transmit power, make the information transfer more energy-efficient, and increase the battery life of the mobile devices.

6.2.2 Noncoherent Transmission

In practical communication networks, the receiver may struggle to estimate CSI. That is especially the case when the channel changes fast. The channel estimation is more challenging when we have multiple parallel channels associated with the signal transmission since they require higher pilot overhead and consume more power for channel estimation than a SISO channel. Some communication networks may not have enough computing power and system complexity to address this challenging problem. In this case, the transmission can be conducted without CSI at the transmitter and receiver, i.e., noncoherent transmission. However, error performance is sacrificed in noncoherent transmission as CSI plays a critical role in detecting the transmitted signals correctly. That makes noncoherent communication more suitable for networks where the transceiver structure is not sophisticated and relatively high error rates are admissible.

Several studies consider noncoherent transmission in the IM systems, i.e., noncoherent IM [132–135]. These studies investigate the error performance of the noncoherent IM systems, design optimum and low-complexity sub-optimum receivers for these systems, and attempt to improve their performance by deriving optimum amplitude levels and introducing an additional diversity to the system. Similarly, one can consider noncoherent detection for our proposed techniques, conduct analyses,

and attempt to improve the performance as in the IM studies. It would be beneficial to compare the resulting noncoherent systems with the noncoherent IM systems to see the potential when we have no access to CSI.

Massive MIMO will play a vital role to satisfy the requirements of beyond 5G networks in terms of spectral/energy efficiency, data rate, and reliability [136]. However, channel estimation overhead is high in the massive MIMO technology due to the presence of a high dimensional channel. Noncoherent transmission can effectively combat this problem and reduce the complexity of massive MIMO systems. In this regard, further research efforts on optimizing the proposed modulation methods for noncoherent transmission can help these systems to overcome the channel estimation problem and satisfy the requirements of beyond 5G networks.

6.2.3 Adaptive Modulation and Bit Loading

When CSI is available at the receiver, a feedback channel can convey it to the transmitter. Then, the transmitter can adapt the communication parameters such as constellation size, power, symbol rate, and coding rate according to the CSI. That allows us to utilize the potential of the channel and leads to a significant improvement in system performance. Therefore, for a fixed BER, adaptive modulation techniques can yield high SEs by transmitting at high rates when the channel conditions are good and reducing the throughput as the conditions deteriorate [137]. One can also optimize BER or transmit power for a given SE by exploiting adaptive modulation techniques. Consequently, one can apply these techniques to the proposed modulation concepts to optimize BER, SE, and transmit power of these concepts according to the changing channel conditions.

6G networks will bring several new applications such as extended reality, wireless brain-computer interactions, connected robotics, and autonomous systems. The majority of these applications will require higher bit rates and reliability than what the

existing networks can support [7]. Adaptive counterparts of the proposed modulation techniques can be beneficial in fulfilling these requirements.

6.2.4 Variable-Length Encoding and Distribution Matching

In the IM methods, the number of codewords related to combinations/permutations is usually not a power of two. To exploit all of the codewords and increase the information rate, these methods can employ well-known variable-length encoding techniques such as Huffman and arithmetic encoding. In this regard, the authors of [80] and [90] employ Huffman encoding in a reverse manner where they map the output of a binary source to the IM symbols based on Huffman encoding. The resulting Huffman encoding-based IM implementations achieve higher information rates than the classical IM schemes.

Similarly, the number of codewords related to the set partitioning, composition, and MDS mechanisms may not be a power of two. This situation makes variable-length encoding techniques a natural source of rate improvement for the SPM, CM, and MDS methods. Therefore, as in [80] and [90], one can use Huffman encoding in a reverse manner and map the information bits to the symbols of SPM, CM, and MDS techniques based on Huffman encoding to increase the rate and reliability of transmission in these techniques. These efforts can potentially result in efficient modulation concepts that support rate/reliability-hungry applications and services of 6G networks.

On the other hand, for many discrete memory-less channels, where a discrete input is drawn from a finite alphabet source, the channel capacity is achieved by a non-uniform probability mass function (PMF). Therefore, using uniformly distributed input symbols results in the degradation of the mutual information. In this regard, it is desirable to map independent Bernoulli distributed input bits to input symbols with a desired distribution [138]. A distribution matcher undertakes this task, and

one can match the capacity-achieving PMFs using a distribution matcher to improve the information rate. Although there exist studies attempting to derive capacity-achieving PMFs for IM [80, 90], the potential of employing a distribution matcher in an IM system has not been fully explored to the best of the author's knowledge. Such application of distribution matching can be further extended to SPM, CM, and MDS techniques, and the effectiveness compared to the IM methods can be investigated.

6.3 Conclusions

Although IM attracted remarkable interest recently, the idea of embedding information into combinations/permutations of codeword elements is not new. Its roots date back to the 60s when David Slepian described PM in [9] as a modulation concept that uses element permutations to form codewords. Unlike Slepian's original PM idea, the recent PM applications exploit well-known constellations and achieve superior error and rate performance compared to the classical modulation concepts. Moreover, although the searching space for the optimal detectors of these applications may be excessive, the receiver complexity is not an issue for them as the combinatorial structures allow them to execute low-complexity detection algorithms. Motivated by these practical advantages, we explored the combinatorial structure that IM, DM-IM, MM-IM, and many other PM-based techniques have. Then, we uncovered different tools of enumerative combinatorics and proposed new modulation concepts using set partitioning and composition. Our analytical and numerical results show that the proposed methods can achieve better error and achievable rate performance than the PM variants while preserving the transceiver complexity at a similar level. We further noticed that the combinatorial modulation concepts involving our novel concepts do not exploit the full potential of the index domain since they do not produce the maximum number of codewords in such a domain. Our research efforts revealed that

modulation based on a simple MDS code utilizes the full potential of the index domain and yields considerably better performance than the combinatorial methods. More importantly, the performance superiority does not come at the cost of increased transceiver complexity.

As the IM studies and this thesis indicated, the combinatorial encoding structure has the potential to improve performance. One can employ such an encoding mechanism through combinations, permutations, set partitions, and compositions on top of the classical modulation techniques and ameliorate their error and rate performance. Moreover, this comes at the cost of only a slight increase in the transceiver complexity. Therefore, it appears not burdensome to build the combinatorial modulation methods into the existing and future wireless networks and utilize the enhanced performance these methods offer. However, the potential of the combinatorial concepts is restricted. That is because the MDS mechanism delivers a better error, rate, and distance performance than these concepts by producing more codewords with desirable properties and exhibiting an even less complex structure. In addition, the lack of straightforward algorithms for LLR calculations and Gray coding may limit the application of combinatorial schemes in future wireless networks that require ultra-reliability. The MDS structure permits simple yet effective solutions for these problems. This structure further comes with an inherent flexibility that enables us to adjust the system performance through modulation parameters such as codeword length, modulation order, and the number of disjoint constellations. These properties make the MDS modulation methods more suitable for future wireless networks.

Bibliography

[1] M. Wen, E. Basar, Q. Li, B. Zheng, and M. Zhang, "Multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Commun*, vol. 65, pp. 3892–3906, Sep. 2017.

[2] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, pp. 5536–5549, Nov. 2013.

[3] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Commun. Lett.*, vol. 12, pp. 225–227, Apr. 2008.

[4] X. Guan, Y. Cai, and W. Yang, "On the mutual information and precoding for spatial modulation with finite alphabet," *IEEE Wireless Commun. Lett.*, vol. 2, pp. 383–386, Aug. 2013.

[5] M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor, "On the achievable rate of OFDM with index modulation," *IEEE Trans. Signal Process.*, vol. 64, pp. 1919–1932, Apr. 2016.

[6] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Jul. 1948.

[7] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, pp. 134–142, May 2020.

[8] "IMT traffic estimates for the years 2020 to 2030." `https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2370-2015-PDF-E.pdf`. Accessed: 2021-12-23.

[9] D. Slepian, "Permutation modulation," *Proc. IEEE*, vol. 53, pp. 228–236, Mar. 1965.

[10] N. Ishikawa, S. Sugiura, and L. Hanzo, "50 years of permutation, spatial and index modulation: From classic RF to visible light communications and data storage," *IEEE Commun. Surveys Tuts.*, vol. 20, pp. 1905–1938, 3rd Quart. 2018.

[11] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Tech.*, vol. 57, pp. 2228–2241, Jul. 2008.

[12] T. Mao, Q. Wang, Z. Wang, and S. Chen, "Novel index modulation techniques: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, pp. 315–348, 1st Quart. 2019.

[13] L. Xiao, D. Chen, I. A. Hemadeh, P. Xiao, and T. Jiang, "Graph theory assisted bit-to-index-combination Gray coding for generalized index modulation," *IEEE Tran. Wireless Commun.*, vol. 19, pp. 8232–8245, Dec. 2020.

[14] F. Yarkin and J. P. Coon, "Set partition modulation," *IEEE Trans. Wireless Commun.*, vol. 19, pp. 7557–7570, Nov. 2020.

[15] F. Yarkin and J. P. Coon, "Set partition modulation," in *IEEE 30th Int. Sym. Pers., Indoor Mobile Radio Commun.*, pp. 1–7, Sep. 2019.

[16] F. Yarkin and J. P. Coon, "Index and composition modulation," *IEEE Commun. Lett.*, vol. 25, pp. 911–915, Mar. 2021.

[17] F. Yarkin and J. P. Coon, "Composition modulation," in *2020 IEEE 20th ICCT*, pp. 131–135, Oct. 2020.

[18] F. Yarkin and J. P. Coon, "Modulation based on a simple MDS code: Achieving better error performance than index modulation and related schemes," *IEEE Trans. Commun.*, vol. 70, pp. 118–131, Jan. 2022.

[19] F. Yarkin and J. P. Coon, "Q-ary multi-mode OFDM with index modulation," *IEEE Wireless Commun. Letts.*, vol. 9, pp. 1110–1114, Jul. 2020.

[20] F. Yarkin and J. P. Coon, "Simple Gray coding and LLR calculation for MDS modulation systems," in *2022 IEEE WCNC*, pp. 268–273, 2022.

[21] F. Yarkin and J. P. Coon, "Comparing minimum codeword distances and error performance for index modulation and maximum distance separable coded modulation," in *2022 EuCNC/6G Summit*, pp. 106–111, 2022.

[22] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill International Edition, McGraw-Hill, 2008.

[23] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[24] G. D. Forney and L. . Wei, "Multidimensional constellations. i. introduction, figures of merit, and generalized cross constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, pp. 877–892, Aug. 1989.

[25] H. Jafarkhani, *Space-Time Coding: Theory and Practice*. Cambridge University Press, 2005.

[26] J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 44, pp. 1453–1467, Jul. 1998.

[27] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[28] B. Vucetic and J. Yuan, *Space-Time Coding*. New York, NY, USA: John Wiley & Sons, Inc., 2003.

[29] E. Larsson, "MIMO detection methods: How they work [lecture notes]," *IEEE Signal Process. Mag.*, vol. 26, pp. 91–95, May 2009.

[30] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, pp. 429–445, Mar. 1996.

[31] V. Tarokh and H. Jafarkhani, "On the computation and reduction of the peak-to-average power ratio in multicarrier communications," *IEEE Trans. Commun.*, vol. 48, pp. 37–44, Jan. 2000.

[32] G. Forney, R. Gallager, G. Lang, F. Longstaff, and S. Qureshi, "Efficient modulation for band-limited channels," *IEEE J. Sel. Areas Commun*, vol. 2, pp. 632–647, Sep. 1984.

[33] J. Boutros, E. Viterbo, C. Rastello, and J.-C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian channels," *IEEE Trans Inf. Theory*, vol. 42, pp. 502–518, Mar. 1996.

[34] A. Nordio and E. Viterbo, "Permutation modulation for fading channels," in *Int. Conf. Telecommunications*, vol. 2, pp. 1177–1183, 2003.

[35] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Commun. Mag.*, vol. 54, pp. 168–175, Jul. 2016.

[36] R. Fan, Y. J. Yu, and Y. L. Guan, "Generalization of orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 5350–5359, Oct. 2015.

[37] B. Zheng, F. Chen, M. Wen, F. Ji, H. Yu, and Y. Liu, "Low-complexity ML detector and performance analysis for OFDM with in-phase/quadrature index modulation," *IEEE Commun. Lett.*, vol. 19, pp. 1893–1896, Nov. 2015.

[38] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-mode index modulation aided OFDM," *IEEE Access*, vol. 5, pp. 50–60, Feb. 2017.

[39] T. Mao, Q. Wang, and Z. Wang, "Generalized dual-mode index modulation aided OFDM," *IEEE Commun. Lett.*, vol. 21, pp. 761–764, Apr. 2017.

[40] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 3692–3703, Jul. 2009.

[41] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, pp. 545–547, Aug. 2008.

[42] M. D. Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: a survey," *IEEE Commun. Mag.*, vol. 49, pp. 182–191, Dec. 2011.

[43] M. D. Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, pp. 56–103, Jan. 2014.

[44] R. Mesleh, S. S. Ikki, and H. M. Aggoune, "Quadrature spatial modulation," *IEEE Trans. Veh. Tech.*, vol. 64, pp. 2738–2742, Jun. 2015.

[45] S. Sugiura, S. Chen, and L. Hanzo, "Coherent and differential space-time shift keying: A dispersion matrix approach," *IEEE Trans. Commun.*, vol. 58, pp. 3219–3230, Nov. 2010.

[46] S. Sugiura, S. Chen, and L. Hanzo, "Generalized space-time shift keying designed for flexible diversity-, multiplexing- and complexity-tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1144–1153, Apr. 2011.

[47] G. Zafari, M. Koca, and H. Sari, "Spatial modulation with dual-polarized antennas," in *2015 IEEE IEEE Int. Conf. Commun. (ICC)*, pp. 2375–2380, 2015.

[48] J. Zhang, Y. Wang, J. Zhang, and L. Ding, "Polarization shift keying (PolarSK): System scheme and performance analysis," *IEEE Trans. Veh. Tech.*, vol. 66, pp. 10139–10155, Nov. 2017.

[49] Y. Naresh and A. Chockalingam, "On media-based modulation using RF mirrors," *IEEE Trans. Veh. Tech.*, vol. 66, pp. 4967–4983, Jun. 2017.

[50] A. K. Khandani, "Media-based modulation: A new approach to wireless transmission," in *Proc. IEEE Int. Symp. Inf. Theory,*, pp. 3050–3054, Jul. 2013.

[51] E. Basar, "Media-based modulation for future wireless systems: A tutorial," *IEEE Wireless Commun.*, vol. 26, pp. 160–166, Oct. 2019.

[52] G. Kaddoum, M. F. A. Ahmed, and Y. Nijsure, "Code index modulation: A high data rate and energy efficient communication system," *IEEE Commun. Lett.*, vol. 19, pp. 175–178, Feb. 2015.

[53] G. Kaddoum, Y. Nijsure, and H. Tran, "Generalized code index modulation technique for high-data-rate communication systems," *IEEE Trans. Veh. Tech.*, vol. 65, pp. 7000–7009, Sep. 2016.

[54] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *Proc. IEEE Int. Sym. Pers., Indoor Mobile Radio Commun.*, pp. 177–181, Sep. 2009.

[55] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *Proc. IEEE GLOBECOM Workshops*, pp. 728–732, Dec. 2011.

[56] B. Shamasundar, S. Bhat, S. Jacob, and A. Chockalingam, "Multidimensional index modulation in wireless communications," *IEEE Access*, vol. 6, pp. 589–604, 2018.

[57] S. Lu, I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Compressed sensing-aided multi-dimensional index modulation," *IEEE Trans. Commun.*, vol. 67, pp. 4074–4087, Jun. 2019.

[58] Y. Cui, X. Fang, and L. Yan, "Hybrid spatial modulation beamforming for mmWave railway communication systems," *IEEE Trans. Veh. Tech.*, vol. 65, pp. 9597–9606, Dec. 2016.

[59] P. Yang, Y. Xiao, Y. L. Guan, M. Di Renzo, S. Li, and L. Hanzo, "Multidomain index modulation for vehicular and railway communications: A survey of novel techniques," *IEEE Veh. Tech. Mag.*, vol. 13, pp. 124–134, Sep. 2018.

[60] M. C. Gursoy, E. Basar, A. E. Pusane, and T. Tugcu, "Index modulation for molecular communication via diffusion systems," *IEEE Trans. Commun.*, vol. 67, pp. 3337–3350, May 2019.

[61] Y. Huang, M. Wen, L.-L. Yang, C.-B. Chae, and F. Ji, "Spatial modulation for molecular communication," *IEEE Trans. Nanobiosci.*, vol. 18, pp. 381–395, Jul. 2019.

[62] M. Herceg, L. Filipović, T. Matić, and G. Kaddoum, "Inductance index modulation for human body communication systems," *IEEE Wireless Commun. Lett.*, vol. 8, pp. 937–940, Jun. 2019.

[63] R. Mesleh, H. Elgala, and H. Haas, "Optical spatial modulation," *IEEE/OSA J. Opt. Commun. Netw*, vol. 3, pp. 234–244, Mar. 2011.

[64] A. Yesilkaya, E. Basar, F. Miramirkhani, E. Panayirci, M. Uysal, and H. Haas, "Optical MIMO-OFDM with generalized LED index modulation," *IEEE Trans. Commun.*, vol. 65, pp. 3429–3441, Aug. 2017.

[65] F. J. Escribano, A. Wagemakers, G. Kaddoum, and J. V. C. Evangelista, "Design and performance analysis of an index time-frequency modulation scheme for optical communications," *IEEE J. Sel. Topics Signal Process.*, vol. 13, pp. 1403–1416, Oct. 2019.

[66] L. Wang, S. Bashar, Y. Wei, and R. Li, "Secrecy enhancement analysis against unknown eavesdropping in spatial modulation," *IEEE Commun. Lett.*, vol. 19, pp. 1351–1354, Aug. 2015.

[67] X.-Q. Jiang, M. Wen, H. Hai, J. Li, and S. Kim, "Secrecy-enhancing scheme for spatial modulation," *IEEE Commun. Lett.*, vol. 22, pp. 550–553, Mar. 2018.

[68] R. Mesleh, S. Ikki, and M. Alwakeel, "Performance analysis of space shift keying with amplify and forward relaying," *IEEE Commun. Lett.*, vol. 15, pp. 1350–1352, Dec. 2011.

[69] M. Wen, X. Cheng, H. V. Poor, and B. Jiao, "Use of SSK modulation in two-way amplify-and-forward relaying," *IEEE Trans. Veh. Tech.*, vol. 63, pp. 1498–1504, Mar. 2014.

[70] S. Dang, J. P. Coon, and G. Chen, "Adaptive OFDM with index modulation for two-hop relay-assisted networks," *IEEE Trans. Wireless Commun.*, vol. 17, pp. 1923–1936, Mar. 2018.

[71] S. Sugiura, T. Ishihara, and M. Nakao, "State-of-the-art design of index modulation in the space, time, and frequency domains: Benefits and fundamental limitations," *IEEE Access*, vol. 5, pp. 21774–21790, 2017.

[72] S. Dogan Tusha, A. Tusha, E. Basar, and H. Arslan, "Multidimensional index modulation for 5G and beyond wireless networks," *Proc. IEEE*, vol. 109, no. 2, pp. 170–199, 2021.

[73] J. Choi, "Coded OFDM-IM with transmit diversity," *IEEE Trans. Commun.*, vol. 65, pp. 3164–3171, Jul. 2017.

[74] J. Choi and Y. Ko, "TCM for OFDM-IM," *IEEE Wireless Commun. Lett.*, vol. 7, pp. 50–53, Feb. 2018.

[75] E. Basar, "OFDM with index modulation using coordinate interleaving," *IEEE Wireless Commun. Lett.*, vol. 4, pp. 381–384, Aug. 2015.

[76] N. Ishikawa, S. Sugiura, and L. Hanzo, "Subcarrier-index modulation aided OFDM - will it work?," *IEEE Access*, vol. 4, pp. 2580–2593, Jun. 2016.

[77] A. I. Siddiq, "Effect of subcarrier activation ratio on the performance of OFDM-IM over Rayleigh fading channel," *IEEE Commun. Lett.*, vol. 21, pp. 1293–1296, Jun. 2017.

[78] N. H. Nguyen, B. Berscheid, and H. H. Nguyen, "Fast-OFDM with index modulation for NB-IoT," *IEEE Commun. Lett.*, vol. 23, pp. 1157–1160, Jul. 2019.

[79] M. Irfan and S. Aïssa, "Generalization of index-modulation: Breaking the conventional limits on spectral and energy efficiencies," *IEEE Trans. Wireless Commun.*, vol. 20, pp. 3911–3924, Jun. 2021.

[80] J. Coon, M. Badiu, Y. Liu, F. Yarkin, and S. Dang, "Binary-tree encoding for uniform binary sources in index modulation systems," *IEEE J. Sel. Top. Signal Process.*, vol. 13, pp. 1270–1285, Oct. 2019.

[81] M. Chafii, J. P. Coon, and D. A. Hedges, "DCT-OFDM with index modulation," *IEEE Commun. Lett.*, vol. 21, pp. 1489–1492, Jul. 2017.

[82] X. Zhang, H. Bie, Q. Ye, C. Lei, and X. Tang, "Dual-mode index modulation aided OFDM with constellation power allocation and low-complexity detector design," *IEEE Access*, vol. 5, pp. 23871–23880, 2017.

[83] Q. Li, M. Wen, H. V. Poor, and F. Chen, "Information guided precoding for OFDM," *IEEE Access*, vol. 5, pp. 19644–19656, 2017.

[84] T. Mao, Q. Wang, J. Quan, and Z. Wang, "Zero-padded tri-mode index modulation aided OFDM," in *Proc. GLOBECOM*, pp. 1–5, Dec. 2017.

[85] M. Wen, Q. Li, E. Basar, and W. Zhang, "Generalized multiple-mode OFDM with index modulation," *IEEE Trans. Wireless Commun.*, vol. 17, pp. 6531–6543, Oct. 2018.

[86] J. Li, M. Wen, X. Jiang, and W. Duan, "Space-time multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Access*, vol. 5, pp. 23212–23222, 2017.

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

[88] D. A. Basnayaka, M. D. Renzo, and H. Haas, "Massive but few active MIMO," *IEEE Trans. Veh. Tech.*, vol. 65, pp. 6861–6877, Sep. 2016.

[89] Z. An, J. Wang, J. Wang, S. Huang, and J. Song, "Mutual information analysis on spatial modulation multiple antenna system," *IEEE Trans. Commun.*, vol. 63, pp. 826–843, Mar. 2015.

[90] W. Wang and W. Zhang, "Huffman coding-based adaptive spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 5090–5101, Aug. 2017.

[91] A. Younis and R. Mesleh, "Information-theoretic treatment of space modulation MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 67, pp. 6960–6969, Aug. 2018.

[92] S. Guo, H. Zhang, J. Zhang, and D. Yuan, "On the mutual information and constellation design criterion of spatial modulation MIMO systems," in *IEEE International Conf. Commun. Sys.*, pp. 487–491, Nov. 2014.

[93] R. Bohnke, D. Wubben, V. Kuhn, and K.-D. Kammeyer, "Reduced complexity MMSE detection for BLAST architectures," in *Proc. IEEE Globecom'03*, vol. 4, pp. 2258–2262 vol.4, 2003.

[94] N. Ishikawa, "IMtoolkit: An open-source index modulation toolkit for reproducible research based on massively parallel algorithms," *IEEE Access*, vol. 7, pp. 93830–93846, 2019.

[95] C.-M. Yu, S.-H. Hsieh, H.-W. Liang, C.-S. Lu, W.-H. Chung, S.-Y. Kuo, and S.-C. Pei, "Compressed sensing detector design for space shift keying in MIMO systems," *IEEE Commun. Lett.*, vol. 16, pp. 1556–1559, Oct. 2012.

[96] L. Xiao, P. Yang, Y. Xiao, S. Fan, M. Di Renzo, W. Xiang, and S. Li, "Efficient compressive sensing detectors for generalized spatial modulation systems," *IEEE Trans. Veh. Tech.*, vol. 66, pp. 1284–1298, Feb. 2017.

[97] A. Younis, S. Sinanovic, M. Di Renzo, R. Mesleh, and H. Haas, "Generalised sphere decoding for spatial modulation," *IEEE Trans. Commun.*, vol. 61, pp. 2805–2815, Jul. 2013.

[98] B. Zheng, M. Wen, F. Chen, N. Huang, F. Ji, and H. Yu, "The K-best sphere decoding for soft detection of generalized spatial modulation," *IEEE Trans. Commun.*, vol. 65, pp. 4803–4816, Nov. 2017.

[99] M. Wen, B. Zheng, K. J. Kim, M. Di Renzo, T. A. Tsiftsis, K.-C. Chen, and N. Al-Dhahir, "A survey on spatial modulation in emerging wireless systems: Research progresses and applications," *IEEE J. Sel. Areas Commun.*, vol. 37, pp. 1949–1972, Sep. 2019.

[100] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, pp. 55–67, Jan. 1982.

[101] R. Carraghan and P. M. Pardalos, "An exact algorithm for the maximum clique problem," *Oper. Res. Lett.*, vol. 9, no. 6, pp. 375 – 382, 1990.

[102] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo, *The Maximum Clique Problem*, pp. 1–74. Boston, MA: Springer US, 1999.

[103] A. Amin and S. Hakimi, "Upper bounds on the order of a clique of a graph," *SIAM J. of Appl. Math*, vol. 22, no. 4, pp. 569–573, 1972.

[104] D. E. Knuth, *The Art of Computer Programming, Volume 1 (3rd Ed.): Fundamental Algorithms.* Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc., 1997.

[105] E. W. Weisstein, "Clique." `http://mathworld.wolfram.com/Clique.html`.

[106] N. M. Temme, "Asymptotic estimates of Stirling numbers," *Stud. Appl. Math.*, vol. 89, no. 3, pp. 233–243, 1993.

[107] B. Rennie and A. Dobson, "On Stirling numbers of the second kind," *Journal of Combinatorial Theory*, vol. 7, no. 2, pp. 116 – 121, 1969.

[108] E. R. Canfield and C. Pomerance, "On the problem of uniqueness for the maximum Stirling number(s) of the second kind," *INTEGERS*, vol. 2, 2002.

[109] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the LambertW function," *Adv. Comput. Math.*, vol. 5, pp. 329–359, Dec. 1996.

[110] I. Mező, "Asymptotics of the modes of the ordered Stirling numbers," *arXiv e-prints*, p. arXiv:1504.06970, Apr. 2015.

[111] N. G. de Bruijn, *Asymptotic Methods in Analysis*. New York: Dover, second ed., 1981.

[112] J. Barthelemy, "An asymptotic equivalent for the number of total preorders on a finite set," *Discrete Math.*, vol. 29, no. 3, pp. 311 – 313, 1980.

[113] Soon Xin Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Trans. Veh. Tech.*, vol. 55, pp. 528–536, Mar. 2006.

[114] A. Gatherer, "What will 6G be?." `https://www.comsoc.org/publications/ctn/what-will-6g-be`, Jun. 2018.

[115] R. P. Stanley, *Enumerative Combinatorics*, vol. 1 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2 ed., 2011.

[116] E. Arslan, A. T. Dogukan, and E. Basar, "Sparse-encoded codebook index modulation," *IEEE Trans. Veh. Tech.*, vol. 69, pp. 9126–9130, Aug. 2020.

[117] A. V. Sills, "Compositions, partitions, and Fibonacci Numbers," *arXiv e-prints*, p. arXiv:1312.0693, Dec. 2013.

[118] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 2012.

[119] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between Gaussian mixture models," in *2007 IEEE ICASSP*, vol. 4, pp. IV–317–IV–320, 2007.

[120] R. Lucky and J. Hancock, "On the optimum performance of N-ary systems having two degrees of freedom," *IRE Trans. Commun. Sys.*, vol. 10, no. 2, pp. 185–192, 1962.

[121] C. Thomas, M. Weidner, and S. Durrani, "Digital amplitude-phase keying with M-ary alphabets," *IEEE Trans. Commun.*, vol. 22, pp. 168–180, Feb. 1974.

[122] R. Singleton, "Maximum distance q-nary codes," *IEEE Trans. Inf. Theory*, vol. 10, pp. 116–118, Apr. 1964.

[123] P. Yang, Y. Xiao, B. Zhang, S. Li, M. El-Hajjar, and L. Hanzo, "Star-QAM signaling constellations for spatial modulation," *IEEE Trans. Veh. Tech.*, vol. 63, pp. 3741–3749, Oct. 2014.

[124] A. Khandekar and R. J. McEliece, "On the complexity of reliable communication on the erasure channel," in *Proc. 2001 IEEE Int. Symp. Inf. Theory (IEEE Cat. No.01CH37252)*, pp. 1–, Jun. 2001.

[125] Z. Hu, F. Chen, M. Wen, F. Ji, and H. Yu, "Low-complexity LLR calculation for OFDM with index modulation," *IEEE Wireless Commun. Lett.*, vol. 7, pp. 618–621, Aug. 2018.

[126] Z. Hu, F. Chen, Y. Liu, S. Liu, H. Yu, and F. Ji, "Low-complexity detection for multiple-mode OFDM with index modulation," *Phys. Commun.*, vol. 34, pp. 38–47, Feb. 2019.

[127] S. Hong, M. Sagong, C. Lim, S. Cho, K. Cheun, and K. Yang, "Frequency and quadrature-amplitude modulation for downlink cellular OFDMA networks," *IEEE J. Sel. Area Comm.*, vol. 32, pp. 1256–1267, Jun. 2014.

[128] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, "A tutorial on IEEE 802.11ax high efficiency WLANs," *IEEE Commun. Surveys Tuts.*, vol. 21, pp. 197–216, 1st Quart. 2019.

[129] M. Alsabah, M. A. Naser, B. M. Mahmmod, S. H. Abdulhussain, M. R. Eissa, A. Al-Baidhani, N. K. Noordin, S. M. Sait, K. A. Al-Utaibi, and F. Hashim, "6G wireless communications networks: A comprehensive survey," *IEEE Access*, vol. 9, pp. 148191–148243, 2021.

[130] M. Beko and R. Dinis, "Designing good multi-dimensional constellations," *IEEE Wireless Commun. Lett.*, vol. 1, pp. 221–224, Jun. 2012.

[131] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?," *Nature Electronics*, vol. 3, pp. 20–29, Jan. 2020.

[132] J. Choi, "Noncoherent OFDM-IM and its performance analysis," *IEEE Trans. Wireless Commun.*, vol. 17, pp. 352–360, Jan. 2018.

[133] S. Gopi, S. Kalyani, and L. Hanzo, "Coherent and non-coherent multilayer index modulation," *IEEE Access*, vol. 7, pp. 79677–79693, 2019.

[134] M. Hanif and H. H. Nguyen, "Non-coherent index modulation in Rayleigh fading channels," *IEEE Commun.s Lett.*, vol. 23, pp. 1153–1156, Jul. 2019.

[135] A. Fazeli, H. H. Nguyen, H. D. Tuan, and H. V. Poor, "Non-coherent multi-level index modulation," *IEEE Trans. Commun.*, vol. 70, pp. 2240–2255, Apr. 2022.

[136] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, 2021.

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

[138] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, pp. 430–434, Jan. 2016.