
Leonardo Jiménez Rodríguez

Department of Electrical & Computer Engineering
McGill University
Montréal, Canada

July 2014

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

© 2014 Leonardo Jiménez Rodríguez
Para mi familia
Abstract
Relaying techniques, in which a source node communicates to a destination node with the help of a relay, have been proposed as a cost-effective solution to address the ever increasing demand for high data rates and reliable services over the air. As such, it is crucial to design relay systems that are able to not only provide high spectral efficiency, but also fully exploit the diversity of the relay channel. With this objective in mind, this thesis investigates achievable rates, power allocation schemes, and code designs for half-duplex (HD) and full-duplex (FD) amplify-and-forward (AF) single-relay systems.

The first part of the thesis is concerned with systems in which the relay operates in HD mode. First, we study the capacity and respective power allocation strategy for a non-orthogonal AF relay system over static channel gains. By assuming full channel state information (CSI) at all nodes, we derive the optimal input covariance matrix at the source and power allocation scheme at the relay that maximize the achievable rate under both per-node and sum power constraints. Second, we investigate the ergodic capacity for a wide range of AF protocols over Rayleigh fading channels. Assuming full CSI only at the destination node, a general method is proposed to obtain tight and simple approximations to the achievable rates in high and low transmission power regimes. The proposed approach is applicable to several one-way (OW) and two-way (TW) schemes using the conventional channel inversion or fixed-gain amplification coefficients. The derived approximations are then used not only to compare the protocols of interest, but also to characterize the asymptotically optimal allocations that maximize the achievable rate and sum rate for OW and TW schemes, respectively. Third, optimal amplification coefficients at the relay are developed based on the rate and sum rate criteria for OW and TW relaying. Specifically, optimal power adaptation schemes are established under a long-term average power constraint and assuming full CSI at the relay. Insights into proposed schemes are then discussed and numerical results are presented to quantify the gain of the adaptation techniques over the conventional approaches. Finally, turning our attention to the performance with finite constellations, we propose the idea of precoding over multiple cooperative frames for a non-orthogonal AF system over Rayleigh fading channels. In particular, by adopting the framework of bit-interleaved coded modulation (BICM), we show that multi-frame precoding is able to improve the diversity and approach the capacity of such an AF channel.

The second part of this thesis deals with the FD mode of operation, where the residual
self-interference due to FD transmission is explicitly taken into account. In particular, we assume that the self-interference variance is proportional to the $\lambda$-th power of the transmitted power. First, we investigate the optimal power allocation scheme and corresponding capacity of a static dual-hop AF system with full CSI at all nodes. By analyzing such system in different high power regions, the role of $\lambda$ on the capacity is examined and comparisons to its HD counterpart are carried. Then, we focus on the error performance of several FD protocols over Rayleigh fading channels. Specifically, considering the same interference model, we analyze the error and diversity behavior of uncoded and coded systems, where we again adopt the framework of BICM. Similar to the capacity analysis, insights into the effect of $\lambda$ and comparisons to HD schemes are discussed.
Sommaire

La technique du relais, dans laquelle le nœud transmetteur communique avec le nœud destinataire à l’aide de relais, a été proposée comme une solution à coût réduit pour satisfaire l’augmentation de la demande du très haut débit et garantir les services par la communication sans fil. C’est pour cela qu’il est crucial de concevoir des systèmes de relais qui sont capables non seulement d’offrir une grande efficacité spectrale, mais aussi d’exploiter la diversité du canal relais. En considérant ces objectifs, cette thèse traite les débits atteignables, des schémas d’allocation de puissance, et la conception de codes pour le système à relai avec amplification-et-retransmission (amplify-and-forward ou AF) sous les modes half-duplex (HD) et full-duplex (FD).

La première partie de la thèse est consacrée aux systèmes où les relais opèrent en mode HD. En premier lieu, nous étudions la capacité et les stratégies d’allocation de puissance pour un système AF non-orthogonal dans un canal à gains statiques. En assumant une connaissance parfaite du canal dans tous les nœuds, nous dérivons la matrice de covariance optimale à la source et un schéma d’allocation de puissance au relai qui maximisent le débit sous les contraintes de puissance individuelle et totale. En second lieu, nous examinons la capacité ergodique pour une multitude de protocoles AF dans un canal à évanouissement Rayleigh. En supposant une connaissance complète du canal à la destination, une méthode générale est proposée pour obtenir une approximation simple et étroite du débit dans les régions de large et basse puissance de transmission. L’approche proposée est applicable à plusieurs schémas unidirectionnels (one-way ou OW) et bidirectionnels (two-way ou TW) utilisant les coefficients d’amplification du canal inversé ou à gain fixe. Les approximations dérivées sont utilisées non seulement pour comparer les protocoles en question, mais aussi pour caractériser les allocations asymptotiquement optimales qui maximisent le débit et la somme de débits pour les schémas OW et TW. En troisième lieu, les coefficients d’amplifications optimales au relai sont développés en se basant sur les critères de débit et de la somme de débit pour les relais OW et TW. Particulièrement, des schémas d’adaptations optimaux de puissance sont établis sous la contrainte de puissance à long terme et en assumant une connaissance complète du canal au relai. Les schémas proposés sont ensuite discutés et des résultats numériques sont présentés pour quantifier le gain de la technique d’adaptation par rapport aux approches conventionnelles. Finalement, en tournant notre attention à la performance avec des constellations finies, nous proposons...
l'idée de précodage pour un système AF non-orthogonal sur les canaux à évanouissement Rayleigh. En particulier, en adoptant le schéma de modulation codée avec entrelacement par bit (bit-interleaved coded modulation ou BICM), nous démontrons que le précodage est capable d'améliorer la diversité et d'approcher la capacité de ce canal AF.

La seconde partie de cette thèse traite le mode de transmission opérant en FD, où l'autobrouillage résiduel causé par la transmission en FD est explicitement pris en considération. En particulier, nous assurons que la variance de l'autobrouillage est proportionnelle à la puissance de transmission portée à la puissance $\lambda$. Premièrement, nous étudions un schéma d'allocation de puissance optimal et sa capacité pour un système AF statique avec connaissance totale du canal. En analysant ce système dans des régions à puissance élevée différentes, le rôle de $\lambda$ sur la capacité est examiné et comparé à son contrepartie HD. Ensuite, nous nous focalisons sur la performance en termes d'erreur de plusieurs protocoles FD dans les canaux à évanouissement Rayleigh. Particulièrement, en considérant le même modèle d'autobrouillage, nous analysons le taux d'erreur et la diversité pour les systèmes codés et non codés, où nous adaptons aussi le codage BICM. Tout comme l'analyse de la capacité, l'effet de $\lambda$ et une comparaison avec le schéma HD sont discutés.
Acknowledgments

First, I would like to thank my supervisor, Professor Tho Le-Ngoc, for giving me the opportunity to pursue my graduate studies at McGill University, and for his support and guidance throughout this journey. Not only have I benefited from his deep knowledge in and outside of the laboratory, but I have also learned a great deal from his way of doing research that will surely help in my career. This thesis would have not been possible without his continuous trust, patience, and encouragement. I would also like to thank Professor Nghi Tran of the University of Akron for the productive research collaborations and fruitful discussions over the years.

I wish to express my appreciation to the members of my PhD supervisory committee, Professor Benoît Champagne and Professor Bruce Shepherd, for taking time from their busy schedules to read my reports and attend several meetings. Their valuable feedback at different stages of my PhD has greatly improved the quality of this thesis.

In addition, I gratefully acknowledge the financial support provided by McGill University through the McGill Engineering Doctoral Award (MEDA), and by the Natural Sciences and Engineering Research Council of Canada (NSERC) and InterDigital Canada Ltée through a Collaborative Research and Development Grant. Without such generosity, I would not have been able to concentrate fully on my doctoral studies.

I am thankful to my friends and colleagues (past and present) of the Broadband Communications Research Laboratory: Chris, Sanjeewa, Suman, Amir, Mahsa, Loïc, Sean, Danny, Prabhat, Soham, Ahmed, Jamshid, Beri, Sinh, Kevin, Chaitanya, Saeideh and Seonghwan. You guys have made my stay at McGill a truly enjoyable one. In particular, I thank Ahmed Masmoudi and Loïc Canonne-Velasquez for the French translation of the abstract.

Finally, my deepest love and gratitude are devoted to all my family members: my grandparents, aunts, uncles, and cousins whose unconditional support has helped me to overcome the difficulties despite the distance. Above all, I am forever indebted to my mom, my dad, my sister Ziania, and my girlfriend Jing Jin for their love, encouragement and sacrifice. My life would not be the same without any of you.
# Contents

1 Introduction .......................................................... 1
   1.1 Half-duplex Relaying .............................................. 2
   1.2 Full-duplex Relaying ............................................. 4
   1.3 Relay Functions ................................................... 5
   1.4 Contributions and Thesis Outline ............................. 6

2 Background and Relay Protocols ................................. 10
   2.1 Wireless Channel ................................................ 10
   2.2 Error Performance and Diversity ............................... 12
   2.3 Achievable Rates and Capacity ................................. 14
      2.3.1 Single-input Single-output Channel ..................... 14
      2.3.2 MIMO Channel ............................................. 17
      2.3.3 Two-way Channel .......................................... 18
   2.4 BICM, BICM-ID and EXIT Charts ............................... 19
   2.5 Input-Output Relations ......................................... 20
      2.5.1 HD Protocols ............................................. 20
      2.5.2 FD Protocols ............................................. 26
   2.6 Chapter Summary ................................................. 30

3 Capacity and Power Allocation of the Static NAF Relay Channel 32
   3.1 Problem Formulation ............................................. 34
   3.2 Capacity under Individual Power Constraints ............... 35
   3.3 Capacity under Joint Power Constraints ..................... 42
   3.4 High and Low Power Regions .................................. 45
      3.4.1 Per-node Power Constraints .............................. 45
3.4.2 Global Power Constraint .................................. 47
3.5 Illustrative Examples ........................................ 48
  3.5.1 Symmetric Network Model ............................... 49
  3.5.2 Linear Network Model ................................... 52
3.6 Chapter Summary ............................................. 54

4 Achievable Rate and Power Allocation of AF over Fading Channels 56
  4.1 Problem Formulation ........................................ 59
  4.2 Achievable Rates and Closed-form Approximations .......... 61
    4.2.1 OW DHAF Systems ..................................... 63
    4.2.2 OW Cooperative Systems ............................... 65
    4.2.3 TW Systems ............................................ 68
    4.2.4 Remarks ............................................... 69
  4.3 Optimal Power Allocation ................................... 71
    4.3.1 DHAF Systems .......................................... 71
    4.3.2 Cooperative Systems .................................. 72
    4.3.3 TW Systems .......................................... 73
  4.4 Closed-form Solution for DHAF System with CI Coefficient 74
  4.5 Rate Comparisons .......................................... 76
    4.5.1 Individual Rate Comparison ............................ 76
    4.5.2 Sum Rate Comparison .................................. 77
  4.6 Illustrative Examples ...................................... 79
    4.6.1 Tightness of the Proposed Approximations ............. 79
    4.6.2 Optimal Power Allocation ............................. 81
    4.6.3 Rate Comparisons .................................... 86
  4.7 Chapter Summary ............................................ 86

5 Relay Adaptation Policies for AF Relay Systems over Fading Channels 88
  5.1 Problem Formulation ....................................... 90
  5.2 Optimal Power Adaptation Schemes ........................ 92
    5.2.1 OAF System ............................................ 92
    5.2.2 NAF System ........................................... 94
    5.2.3 TW System ............................................. 97
5.3 Discussions on Optimal Solutions ........................................... 99
  5.3.1 Zero and Non-zero Power Regions ................................... 99
  5.3.2 High and Low Power Regions ....................................... 102
5.4 Illustrative Examples .................................................. 104
5.5 Chapter Summary ...................................................... 108

6 Multi-frame Precoding for NAF Relaying over Fading Channels 109
  6.1 System Model .......................................................... 111
  6.2 Performance Analysis ................................................ 114
  6.3 Diversity Benefits in Error-floor Regions ......................... 117
    6.3.1 Diversity and Coding Gain Functions ......................... 117
    6.3.2 Optimal Class of Precoders .................................. 119
    6.3.3 Design of Superposition Angles For Independent Mappings .... 121
  6.4 Near-Capacity Design in Turbo Pinch-off Areas ............... 123
    6.4.1 Multi-D Mapping in Precoded Multiple Cooperative Frames ... 124
    6.4.2 EXIT Chart Analysis ........................................... 125
  6.5 Illustrative Examples ................................................ 128
    6.5.1 Tightness of the Derived Bound ................................ 128
    6.5.2 Superiority of Multiple-frame Precoding in Error-floor Regions ... 129
    6.5.3 Capacity-approaching Performances with Multiple-frame Precoding 132
  6.6 Chapter Summary ...................................................... 134

7 Capacity of FD AF Relaying under Residual Self-interference 135
  7.1 Problem Formulation ................................................ 137
    7.1.1 Residual Self-interference Model .......................... 137
    7.1.2 Achievable Rate and Capacity Formulation .................. 138
  7.2 Capacity and Optimal Power Allocation .......................... 138
  7.3 Asymptotic Analysis ................................................ 141
  7.4 Illustrative Examples and Comparisons to HD Schemes ....... 143
  7.5 Chapter Summary ...................................................... 146

8 Error Performance of FD AF under Residual Self-interference 147
  8.1 System Model ........................................................ 149
  8.2 PEP Analysis and Tight BER Bounds ............................... 151
8.2.1 PEP for the LR System .................................................. 151
8.2.2 PEP for the DH System .................................................. 154
8.2.3 Tight BER Bounds of BICM-ID Systems ......................... 155
8.3 Diversity and Coding Gain Analysis ................................. 155
  8.3.1 LR Analysis .......................................................... 156
  8.3.2 DH Analysis .......................................................... 160
8.4 Illustrative Examples ..................................................... 162
  8.4.1 LR Examples .......................................................... 162
  8.4.2 DH Examples .......................................................... 165
8.5 Chapter Summary .......................................................... 167

9  Conclusion ........................................................................ 168
  9.1 Future Work ..................................................................... 170
    9.1.1 Extension Problems .................................................... 170
    9.1.2 More Advanced Problems ............................................. 171

A  Proofs for Chapter 3 ......................................................... 174
  A.1 Proof of Lemma 3.1 ....................................................... 174
  A.2 Solutions to Subcases in Lemma 3.3 ................................. 175
  A.3 Proof of Lemma 3.4 ....................................................... 178
  A.4 Proof that $C_{BF} > C_{OAF}$ ............................................. 179

B  Proofs for Chapter 4 ......................................................... 181
  B.1 Proof of Proposition 4.1 ................................................... 181
  B.2 Positiveness of (4.17) ..................................................... 181

C  Proofs for Chapter 5 ......................................................... 183
  C.1 Properties of $r_i(\cdot)$ with respect to $\mu_3$ ....................... 183
  C.2 Properties of $r_i(\cdot)$ with respect to the Channel Gains .... 184

D  Proofs for Chapter 6 ......................................................... 185
  D.1 Proof of Proposition 6.1 .................................................. 185

E  Proofs for Chapter 8 ......................................................... 187
  E.1 Proof of Proposition 8.1 .................................................. 187
Contents

E.2 Proof of Proposition 8.2 .............................................. 192

References ................................................................. 193
# List of Figures

1.1 HD relay protocols. ................................................................. 3  
1.2 FD relay protocols. ................................................................. 5  
2.1 Channel gains. .......................................................................... 21  
3.1 Feasible region for individual power constraints. ......................... 37  
3.2 Feasible region when \( P(q_1) < z_r \) (left) and \( P(q_1) > z_r \) (right) for \( 0 \leq q_1 \leq q_s \). 40  
3.3 Feasible region for joint power constraint. ................................... 42  
3.4 Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against \( P_s \) \( (P_r = 0 \text{ dB}) \). ...................................................... 50  
3.5 Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against \( P_r \) \( (P_s = 0, 4.77 \text{ dB}) \). ...................................................... 51  
3.6 Normalized rates of the DT, NAF-BF, OAF, OAF-EQ and NAF-EQ protocols against \( P_t \). ................................................................. 51  
3.7 Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against \( P_s = P_r \) \( (d = 0.5) \). ...................................................... 52  
3.8 Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against \( P_r \) \( (P_s = 0 \text{ dB}, d = 0.1, 0.5, 0.9) \). ...................................................... 53  
3.9 Normalized rates of the DT, NAF-BF, OAF, OAF-EQ and NAF-EQ protocols against \( P_t \) \( (d = 0.6) \). ...................................................... 54  
4.1 Achievable rates and high power approximations of the systems using uniform power allocation \( (d = 0.5) \). ...................................................... 80  
4.2 Achievable rates and low power approximations of the systems using uniform power allocation \( (d = 0.5) \). ...................................................... 80
List of Figures

4.3 Achievable rates of the DH systems using different power allocation schemes in high power regions \((d = 0.3)\). ............................................................... 82

4.4 Achievable rates of the DH systems using different power allocation schemes in low power regions \((d = 0.3)\). ............................................................... 82

4.5 Achievable rates of the NAF and OAF systems using different power allocation schemes in high power regions \((d = 0.3)\). ............................................................... 83

4.6 Achievable rates of the NAF and OAF systems using different power allocation schemes in low power regions \((d = 0.3)\). ............................................................... 84

4.7 Achievable sum rates of the TW systems using different power allocation schemes in high power regions \((d = 0.3)\). ............................................................... 84

4.8 Achievable sum rates of the TW systems using different power allocation schemes in low power regions \((d = 0.3)\). ............................................................... 85

4.9 Achievable rates and sum rates of the DH and TW systems normalized by \(C_{DT} = C_{DT}^{\text{sum}} \) \((d = 0.5)\). ............................................................... 87

5.1 Zero and non-zero power regions for the OAF protocol \((\mu_1 = 1, q_1 = 1)\). . . 101

5.2 Zero and non-zero power regions for the NAF protocol \((\mu_2 = 1, q_1 = q_2 = 1)\). 102

5.3 Zero and non-zero power regions for the TW protocol \((\mu_3 = 1, q_1 = q_2 = 1)\). 103

5.4 Optimal adaptation scheme \(z_{2i}^{\text{NAF}} \) in (5.18) for \(\gamma_{0i} = \mu_2 = q_1 = q_2 = 1\). . . 104

5.5 Optimal adaptation scheme \(z_{2i}^{\text{TW}} \) in (5.25) for \(\mu_3 = q_1 = q_2 = 1\). . . . . 105

5.6 Normalized achievable rates for the OAF system using different power adaptation schemes. ............................................................... 106

5.7 Normalized achievable rates for the NAF system using different power adaptation schemes. ............................................................... 107

5.8 Normalized sum rates of the TW system using different power adaptation schemes. ............................................................... 107

6.1 Block diagram of a NAF-BICM-ID system using a precoder \(G\) over multiple cooperative frames. ............................................................... 112

6.2 Achievable rates over NAF channels with different inputs. . . . . . . . . . 125

6.3 Demapper EXIT curves for three different systems. ............................................................... 127

6.4 EXIT charts of 4-D demapper with \(G_{\text{st}}\) and rate-2/3 punctured convolutional code. ............................................................... 128
6.5 BER performance after 1, 2 and 3 iterations of the NAF-BICM-ID system with $N = 3$ using the $G_{\text{cyclo6}}$ precoder. ................................. 129
6.6 BER performances after 5 iterations of the NAF-BICM-ID system with $N = 2$ and different $4 \times 4$ precoders. ................................. 130
6.7 BER performances after 5 iterations of the NAF-BICM-ID system with $N = 2$ using the $I_4$, $G_{st}$, $G_{\pi/6 \text{ Bk}}$ and $G(\varphi)$ precoders. ................................. 131
6.8 BER performances after 1 and 5 iterations of the NAF-BICM-ID system with $N = 2$ using the $G_d$ and $G(\varphi)$ precoders. ................................. 132
6.9 BER performances with 50 iterations of the NAF-BICM-ID system using the $G_{st}$ precoder, multi-D mapping, and the rate-2/3 and rate-3/4 outer codes. 133

7.1 Achievable rates of DH systems against $P_s (\alpha_1 = \alpha_2 = 1, \lambda = [0, 0.5, 1], P_r = 5 \text{ dB})$. ............................................................... 144
7.2 Achievable rates of DH systems against $P_r (\alpha_1 = \alpha_2 = 1, \lambda = [0, 0.3, 0.5, 0.7, 1], P_s = 5 \text{ dB})$. ............................................................... 144
7.3 Normalized rates of DH systems against $P_s = P_r (\alpha_1 = \alpha_2 = 2, \lambda = [0, 0.3, 0.5, 0.7, 1])$. ............................................................... 145
7.4 Normalized rates of DH systems against $P_t (\alpha_1 = \alpha_2 = 2, \lambda = [0, 0.3, 0.5, 0.7, 1])$. 146

8.1 Uncoded LR and DH systems. ............................................................... 149
8.2 BER performances of the uncoded LR system using different $3 \times 3$ precoders $(\lambda = 0, 0.5, 1)$. ............................................................... 163
8.3 BER performances after 3 iterations of the LR BICM-ID system using different $3 \times 3$ precoders $(\lambda = 0, 0.5, 1)$. ............................................................... 164
8.4 BER performances of the uncoded LR system for $L = 3$ and $\lambda = 0.5, 0.7, 1$. 164
8.5 BER performances after 10 iterations of the LR BICM-ID system for $L = 3$ and $\lambda = 0, 0.5, 1$. ............................................................... 165
8.6 BER performances of the uncoded DH system for different values of $\lambda$. 166
8.7 BER performances of the DH BICM system for different values of $\lambda$. 167
List of Tables

2.1 HD relay protocols. .................................................. 21
2.2 LR protocol. ......................................................... 27

3.1 Capacity achieving protocols in high and low power regions. .............. 48

4.1 High power approximations for different AF protocols. ...................... 70
4.2 Low power approximations for different AF protocols. ...................... 71

6.1 NAF transmission in two consecutive cooperative frames. ................. 112

7.1 Capacity and power allocation behavior in high power regions. ........... 142
## List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>3GPP</td>
<td>Third Generation Partnership Project</td>
</tr>
<tr>
<td>AF</td>
<td>Amplify-and-forward</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive white Gaussian noise</td>
</tr>
<tr>
<td>BER</td>
<td>Bit error rate</td>
</tr>
<tr>
<td>BF</td>
<td>Beamforming</td>
</tr>
<tr>
<td>BICM</td>
<td>Bit-interleaved coded modulation</td>
</tr>
<tr>
<td>BICM-ID</td>
<td>Bit-interleaved coded modulation with iterative decoding</td>
</tr>
<tr>
<td>CDI</td>
<td>Channel distribution information</td>
</tr>
<tr>
<td>CF</td>
<td>Compress-and-forward</td>
</tr>
<tr>
<td>CI</td>
<td>Channel inversion</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel state information</td>
</tr>
<tr>
<td>D</td>
<td>Destination</td>
</tr>
<tr>
<td>dB</td>
<td>Decibels</td>
</tr>
<tr>
<td>DF</td>
<td>Decode-and-forward</td>
</tr>
<tr>
<td>DH</td>
<td>Dual-hop</td>
</tr>
<tr>
<td>DHAHF</td>
<td>Dual-hop amplify-and-forward</td>
</tr>
<tr>
<td>DT</td>
<td>Direct transmission</td>
</tr>
<tr>
<td>EXIT</td>
<td>Extrinsic information transfer</td>
</tr>
<tr>
<td>FD</td>
<td>Full-duplex</td>
</tr>
<tr>
<td>FG</td>
<td>Fixed-gain</td>
</tr>
<tr>
<td>FR</td>
<td>Full-rank</td>
</tr>
<tr>
<td>HD</td>
<td>Half-duplex</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush-Kuhn-Tucker</td>
</tr>
<tr>
<td>LLR</td>
<td>Log-likelihood ratio</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>LR</td>
<td>Linear relaying</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>MAP</td>
<td>Maximum a posteriori probability</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-input multiple-output</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum likelihood</td>
</tr>
<tr>
<td>MRC</td>
<td>Maximal-ratio combining</td>
</tr>
<tr>
<td>multi-D</td>
<td>Multidimensional</td>
</tr>
<tr>
<td>NAF</td>
<td>Non-orthogonal amplify-and-forward</td>
</tr>
<tr>
<td>OAF</td>
<td>Orthogonal amplify-and-forward</td>
</tr>
<tr>
<td>OW</td>
<td>One-way</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability density function</td>
</tr>
<tr>
<td>PEP</td>
<td>Pairwise error probability</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature amplitude modulation</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature phase-shift keying</td>
</tr>
<tr>
<td>R</td>
<td>Relay</td>
</tr>
<tr>
<td>RA</td>
<td>Relay adaptation</td>
</tr>
<tr>
<td>S</td>
<td>Source</td>
</tr>
<tr>
<td>$S_i$</td>
<td>Source node $i$</td>
</tr>
<tr>
<td>SER</td>
<td>Symbol error rate</td>
</tr>
<tr>
<td>SISO</td>
<td>Soft-input soft-output</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>TW</td>
<td>Two-way</td>
</tr>
<tr>
<td>TWAF</td>
<td>Two-way amplify-and-forward</td>
</tr>
<tr>
<td>VG</td>
<td>Variable-gain</td>
</tr>
</tbody>
</table>
List of Symbols

\( b \) Amplification coefficient
\( B \) Amplification matrix
\( c \) Coded sequence
\( C, C_{\text{sum}} \) Capacity and sum capacity
\( c_d \) Total information weight at Hamming distance \( d \)
\( d_H \) Free Hamming distance of the channel code
\( F^{-1} \) Coding gain function
\( G \) Power gain
\( G, g_{i,k} \) Precoder and \((i, k)\)-th component
\( h \) Vector of channel gains among all nodes
\( H_{\text{NAF}} \) NAF channel matrix
\( H_{\text{LR}} \) LR channel matrix
\( h_{\text{OAF}} \) OAF channel vector
\( h_0 \) \(S_1\)-\(S_2\) or \(S\)-\(D\) channel gain
\( h_1 \) \(S_1\)-\(R\) or \(S\)-\(R\) channel gain
\( h_2 \) \(S_2\)-\(R\) or \(R\)-\(D\) channel gain
\( I, I_i \) Mutual information at destination and at node \( i \)
\( I|h, I_i|h \) Conditional mutual information at destination and at node \( i \)
\( I_{\text{sum}} \) Sum rate
\( I_{\text{sum}}|h \) Conditional sum rate
\( K, K_v \) Covariance of equivalent noise vector and self-interference
\( k_c, n_c \) Parameters for a rate-\(k_c/n_c\) channel code
\( m_c \) Number of bits in constellation \( \Omega \)
\( n \) Equivalent noise vector
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_i$</td>
<td>Equivalent noise sample at node $i$, or at destination at time $i$</td>
</tr>
<tr>
<td>$n_d, n_r$</td>
<td>Noise vectors at destination and relay</td>
</tr>
<tr>
<td>$n_{d,i}$</td>
<td>Noise sample at node $i$, or at destination at time $i$</td>
</tr>
<tr>
<td>$n_{r,i}$</td>
<td>Noise sample at relay at time $i$</td>
</tr>
<tr>
<td>$N_{d}, N_r, N_{di}$</td>
<td>Noise variance at destination, relay and node $i$</td>
</tr>
<tr>
<td>$P_b$</td>
<td>Bit error rate</td>
</tr>
<tr>
<td>$P_s, P_r, P_{si}, P_t$</td>
<td>Source, relay, node $i$, and total power</td>
</tr>
<tr>
<td>$q$</td>
<td>Power allocation vector containing $q_1$, $q_2$ and $z_2$</td>
</tr>
<tr>
<td>$Q$</td>
<td>Covariance matrix of $x$</td>
</tr>
<tr>
<td>$q_i, q_{ik}$</td>
<td>$i$-th diagonal component and $(i, k)$-component of $Q$</td>
</tr>
<tr>
<td>$q_s, q_t$</td>
<td>Power constraint at source, and global power constraint</td>
</tr>
<tr>
<td>$s, s_i$</td>
<td>Signal vector in $\Psi$ and component in $\Omega$</td>
</tr>
<tr>
<td>$T_c$</td>
<td>Coherence time</td>
</tr>
<tr>
<td>$t$</td>
<td>Vector transmitted by the relay</td>
</tr>
<tr>
<td>$t_i$</td>
<td>Signal transmitted by the relay at time $i$</td>
</tr>
<tr>
<td>$u$</td>
<td>Information sequence</td>
</tr>
<tr>
<td>$U$</td>
<td>$X - \hat{X}$</td>
</tr>
<tr>
<td>$v$</td>
<td>Interleaved sequence</td>
</tr>
<tr>
<td>$v$</td>
<td>Self-interference vector</td>
</tr>
<tr>
<td>$v_i$</td>
<td>Self-interference sample at time $i$</td>
</tr>
<tr>
<td>$V$</td>
<td>Self-interference variance</td>
</tr>
<tr>
<td>$x$</td>
<td>Vector transmitted by source</td>
</tr>
<tr>
<td>$X$</td>
<td>Matrix representation of $x$</td>
</tr>
<tr>
<td>$x_i$</td>
<td>Signal transmitted by node $i$, or by source at time $i$</td>
</tr>
<tr>
<td>$y, y_r$</td>
<td>Vector received at destination and relay</td>
</tr>
<tr>
<td>$y_{r,i}$</td>
<td>Signal received at relay at time $i$</td>
</tr>
<tr>
<td>$y_i$</td>
<td>Signal received at node $i$, or at destination at time $i$</td>
</tr>
<tr>
<td>$Z$</td>
<td>Covariance matrix of $t$</td>
</tr>
<tr>
<td>$z_i$</td>
<td>Power allocation variable at relay at time $i$</td>
</tr>
<tr>
<td>$z_r$</td>
<td>Power constraint at relay</td>
</tr>
<tr>
<td>$\alpha_l$</td>
<td>Magnitude squared of $h_l$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Parameter for residual self-interference model</td>
</tr>
<tr>
<td>$\gamma_l$</td>
<td>Equivalent SNR</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>$\gamma_\theta(\cdot)$</td>
<td>Auxiliary function for BER bound</td>
</tr>
<tr>
<td>$\Delta_\theta(\cdot)$</td>
<td>Auxiliary function for PEP</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Error vector</td>
</tr>
<tr>
<td>$\theta_l$</td>
<td>Phase of $h_l$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Parameter for residual self-interference model</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Path-loss exponent</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Mapping rule</td>
</tr>
<tr>
<td>$\Xi$</td>
<td>Parameter for amplification coefficient</td>
</tr>
<tr>
<td>$\Pi$</td>
<td>Interleaver</td>
</tr>
<tr>
<td>$\phi_l$</td>
<td>Variance of $h_l$</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>Multi-dimensional constellation</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>Single-dimensional constellation</td>
</tr>
</tbody>
</table>
Mathematical Notation

\( A, a \) Matrix (bold upper-case) and vector (bold lower-case)
\( \text{bdiag}(A_1, \ldots, A_N) \) Block-diagonal matrix with diagonal components \( A_i \)
\( \mathcal{CN}(0, A) \) Zero-mean circularly Gaussian distributed vector with covariance \( A \)
\( \det(\cdot) \) Determinant
\( d^2(\cdot, \cdot) \) Squared Euclidean distance
\( \text{diag}(a) \) Diagonal matrix with vector \( a \) in the diagonal
\( \mathbb{E}[\cdot] \) Expectation operator
\( \exp, e \) Euler number
\( E_1(\cdot) \) Exponential integral
\( I_N \) Identity matrix of size \( N \times N \)
\( J(\cdot) \) \( \exp(\cdot)E_1(\cdot) \)
\( \mathcal{L}(\cdot, \cdot) \) Lagrangian
\( \ln(\cdot) \) Natural logarithm
\( \log(\cdot) \) Logarithm base two
\( O(\cdot) \) Big-O notation
\( \text{o.w.} \) Otherwise
\( \text{psd} \) Positive semidefinite
\( \mathbb{P}(\cdot) \) Probability of event (\( \cdot \))
\( Q(\cdot) \) Gaussian probability integral
\( \text{Re}(\cdot) \) Real part
\( \mathbb{R}^n \) \( n \)-dimensional real space
\( \text{s.t.} \) Such that
\( \text{tr}(\cdot) \) Trace
\( \dagger \) Hermitian
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>Transpose</td>
</tr>
<tr>
<td>γ</td>
<td>Euler constant</td>
</tr>
<tr>
<td>∇</td>
<td>Gradient</td>
</tr>
<tr>
<td>∇²</td>
<td>Hessian</td>
</tr>
<tr>
<td>0</td>
<td>Zero matrix of appropriate dimensions</td>
</tr>
<tr>
<td>(·)*</td>
<td>Optimal value</td>
</tr>
<tr>
<td>ℓ(·)</td>
<td>Estimated value</td>
</tr>
<tr>
<td>A ⪰ 0</td>
<td>psd matrix</td>
</tr>
<tr>
<td>∅</td>
<td>Empty set</td>
</tr>
<tr>
<td>⊂, ∩, ∈</td>
<td>Subset, intersection and in set</td>
</tr>
<tr>
<td>[·]+</td>
<td>max{0, ·}</td>
</tr>
<tr>
<td>ℓ(·)</td>
<td>Euclidean norm</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Current wireless communications systems face increasing challenges due to the ever growing demand for high data rates and reliable services over the air. As new applications for small wireless devices (e.g., mobile devices) arise, such rates are expected to be attained using as low power consumption as possible. In addition, bandwidth is a scarce resource and thus wireless systems are also expected to transmit high data rates in a spectrally-efficient manner, i.e., using the least possible spectrum. All of these requirements need to be considered in the design of emerging and future generations of wireless networks.

Relaying techniques, in which helper nodes aid in transmission from one node to another, have been recently proposed as a cost-effective solution to meet some of the demands in next generations of wireless systems [1–4]. In particular, in the context of cellular networks, the deployment of relay nodes has been shown to extend and/or improve the coverage, fill coverage holes, enhance the reliability, and improve the spectral efficiency per unit area. This can be achieved without incurring the associated high costs of adding extra base stations, e.g., site acquisition and backhaul costs. As such, relaying is one of the key features currently being considered in several wireless standards such as the Third Generation Partnership Project (3GPP) Long Term Evolution (LTE), among others [2,3]. Consequently, it is important for future wireless standards to have relay schemes that not only increase the reliability of the wireless network, but also present a high spectral efficiency [5,6].
1 Introduction

1.1 Half-duplex Relaying

The relay channel, in which a source node communicates to a destination node with the help of a relay, was first introduced by van der Meulen in [7]. Earlier information theoretical works assumed that the relay is capable of operating in full-duplex (FD) mode, i.e., the relay is able to transmit and receive at the same time and over the same frequency band [7–11]. This assumption was generally believed to be impractical due to the great difference in transmit and receive signal powers levels, which results in self-interference. Thus, motivated by wireless scenarios, the focus on the relay channel was shifted to half-duplex (HD) operation.

Seminal works on HD relaying concentrated on the dual-hop (DH) strategy, e.g., [12–16]. In DH protocols, information is transmitted from source to destination via the relay in two phases as shown in Fig. 1.1. In the first phase, the source node $S_1$ transmits to the relay $R$, whereas in the second phase, the relay communicates to the destination node $S_2$ and the source remains silent. Such a DH strategy can be easily implemented in practice to improve the coverage of the network and has already been considered for next-generation mobile communication standards [2, 3]. Furthermore, DH schemes are the only alternative when the direct link from source to destination is under severe shadowing. However, when the direct link is available for transmission, DH techniques present two main limitations. First, due to the HD constraint, DH systems are not spectrally-efficient as the source is only allowed to transmit every other phase. Second, although path-loss savings can be obtained, DH protocols provide no diversity gain as there exists only one path from source to destination.

By considering the direct source-destination link, cooperative relaying protocols have been introduced to counteract the drawbacks of DH relaying when the direct link is available [5,17–21]. Pioneering works on cooperative relaying focused on orthogonal schemes in which the source and relay alternate for transmission as in DH systems. However, contrary to DH schemes, the destination also listens to the source node in the first phase as shown in Fig. 1.1. Two independent paths between the source and the destination (the direct and the relay link) are thus created. With proper combining at the destination, rate and diversity advantages over DH protocols can be obtained. This diversity gain, commonly referred to as cooperative diversity, is particularly important for small devices which cannot accommodate multiple antennas due to practical constraints. In this case, spatial diversity
can still be realized through cooperative relaying in a distributed fashion.

Although orthogonal schemes provide rate and diversity advantages compared to DH transmission, the spectral efficiency of these cooperative protocols still suffers from the HD constraint of the relay. This is because similar to DH relaying, the source must remain silent in the second transmission phase. To mitigate the impact of HD relaying, non-orthogonal schemes in which the source is allowed to transmit continuously have been proposed in the literature [20, 21]. In non-orthogonal protocols, as illustrated in Fig. 1.1, the source and relay transmit concurrently in the second phase, maximizing the degrees of broadcasting and receive collision [20]. Moreover, non-orthogonal protocols are general in that they include orthogonal schemes and even direct transmission (i.e., when the relay is not used) as special cases. Unfortunately, the analysis and optimization of non-orthogonal networks are more challenging than their orthogonal counterparts and therefore have received less attention in the literature [20–22].

In the protocols discussed above, the node $S_1$ wants to communicate to $S_2$ via the relay, i.e., information flows from $S_1$ to $S_2$ ($S_1 \rightarrow S_2$). These protocols can thus be broadly classified as one-way (OW) relaying schemes. On the other hand, several applications require the nodes $S_1$ and $S_2$ to exchange information in a bidirectional or two-way (TW) fashion. A simple method to achieve this is by applying any of the above OW protocols, i.e., $S_1 \rightarrow S_2$ over the first two phases and $S_2 \rightarrow S_1$ over the remaining two. This exchange would require four transmission phases and only two symbols would be exchanged if we
were to use orthogonal or DH schemes. To mitigate the impact of HD relaying for bidirectional communication, two-phase and the three-phase TW protocols have recently been proposed in the literature [23,24]. Specifically, in the two-phase scheme, both source nodes simultaneously communicate to the relay in the first phase, whereas the relay broadcasts in the second phase (see Fig. 1.1). In the three-phase scheme, the source nodes take turns to communicate to the relay in the first two phases and the relay broadcasts in the third. Similar to OW non-orthogonal protocols, the potential benefits of TW protocols have just started to be investigated.

1.2 Full-duplex Relaying

It can be seen from the above discussion that the HD constraint of the relay has a great impact on the spectral efficiency of relaying protocols. As explained before, this HD constraint is motivated by the fact that the transmit signal power is usually orders of magnitude larger than the received signal power, resulting in heavy self-interference. Although originally believed to be impractical, FD wireless operation has been recently shown to be feasible through novel combinations of self-interference mitigation schemes (see for example [25–37] and references whithin). In particular, to avoid saturating the receiver front end, several techniques prior to analog-to-digital conversion have been proposed. For instance, basic analog cancellation methods include antenna separation [27–30], orientation [28,31] and directionality [26,32]. More involved methods include asymmetric placement of transmit antennas [33], symmetric placement of antennas with phase shifters [36,37], the use of a balanced/unbalanced transformer [34], the use of a circulator [35], and analog time domain subtraction [26,29–32,35], among others. These analog techniques can be combined with digital methods after quantization such as time domain subtraction for further mitigation [29,30,32,34,35]. Despite these advances in cancellation techniques, the self-interference remains a challenge as it cannot be completely mitigated in practice. As such, different from earlier information theoretical works, the residual self-interference must be explicitly taken into account when assessing, designing and analyzing FD protocols.

Similar to HD schemes, FD protocols can be classified depending on whether the direct source-destination link is used for transmission. In fact, the idea of cooperative relaying can be traced back to the works of van der Meulen and Cover in [7,8], respectively, which make use of the direct link for FD communication. Specifically, in cooperative FD schemes, the
source transmits to the relay and the destination, while the relay simultaneously receives
the signal from the source and transmits to the destination, as shown in Fig. 1.2. Analogous
to the HD DH scheme, the direct link is not used in FD DH protocols. Similar to their
HD counterparts, FD DH relaying has two main limitations when the direct link is not
under heavy shadowing. First, although the source is allowed to transmit continuously,
the rate of the FD DH scheme might be degraded due to the self-interference created at
the destination node from the direct link. Furthermore, this protocol does not provide any
diversity benefits. Thus, as in the HD scenario, cooperative FD techniques that make use
of the direct link for transmission might be able to provide rate and diversity advantages.

![FD relay protocols](image)

**Fig. 1.2** FD relay protocols.

It is important to note that for both FD protocols in Fig. 1.2, the potential gains of
FD relaying might not be realizable due to the level of residual self-interference at the
relay node. Hence, the ideal FD schemes previously proposed in the literature need to be
re-analyzed under such scenario.

### 1.3 Relay Functions

To this day, the capacity of the general relay channel along with its respective optimal
relay function are unknown. Thus, several relay functions have been proposed in the
literature. Due to practical constraints, relay functions in the literature are usually causal,
i.e., the signal transmitted by the relay at a given time can only be a function of previously
received signals. Common relay functions include *decode-and-forward* (DF), *compress-and-
forward* (CF), and *amplify-and-forward* (AF). In DF, the relay first decodes the information
received from the source node, encodes it and then forwards it to the destination node
[8]. On the other hand, in CF, the relay forwards a quantized version of the received signal to the destination [8]. Finally, the relay simply amplifies the signal received in the previous phase and forwards it to the destination in AF [19]. Among these three strategies, the AF technique is of practical interest as it requires lower implementation and computational complexity, it carries less delay at the relay terminal, and it is transparent to the modulation/coding used by the source nodes [38]. As such, the focus of this thesis is on AF relay protocols.

AF relaying can be further classified according to the availability of channel side information (CSI) at the relay node. In particular, to maintain a long-term average power constraint at the relay, most previous studies on AF relaying considered two power amplification techniques. The first method assumes that the relay has only channel distribution information (CDI) of the source-relay link and amplifies the received signal using a fixed-gain (FG) coefficient [20]. Although the instantaneous power of the FG relay is allowed to vary, the amplification gain is chosen to satisfy the average power constraint. The second technique assumes that the relay has an instantaneous knowledge of the source-relay channel and uses this channel gain to normalize the received signal to the desired power level [19]. The latter variable-gain (VG) method is therefore referred to as the channel inversion (CI) coefficient.

1.4 Contributions and Thesis Outline

As noted before, future wireless networks require relaying protocols that are able to fully exploit the diversity of the channel as well as provide high spectral efficiency. Motivated by this fact, the objective of this thesis is to design spectrally-efficient and reliable relaying techniques including both HD and FD relay modes of operation. In particular, this thesis investigates achievable rates, power allocation schemes, and code designs to improve the reliability and spectral efficiency of HD and FD AF relay systems. When the source nodes use Gaussian codebooks, our objective is to derive achievable rate expressions for AF relaying and to propose optimal power allocation schemes to maximize such expressions. When the sources use finite constellations, our objective is to design practical coding schemes that take full advantage of the diversity and approach the capacity of the AF relay channel. In the following, we describe the contributions of this thesis in more detail.

First, Chapter 2 provides some background material and introduces important concepts
1 Introduction

that will be used throughout the thesis. In particular, we first provide an overview of wireless channels. We then discuss error rate performance over fading channels along with the concept of diversity. We also introduce key concepts such as achievable rate and capacity for different systems. A state-of-the-art coded modulation structure, known as bit-interleaved coded modulation (BICM), is then briefly presented. Finally, the input-output relations of the relay protocols considered in this thesis are introduced. These include the HD non-orthogonal, orthogonal, dual-hop, and two-way AF protocols; and the FD dual-hop and linear relaying AF protocols.

Having introduced the background material, in Chapter 3 we analyze the capacity of the static HD non-orthogonal AF (NAF) channel under both per-node and joint power constraints. In particular, by deriving and comparing all local solutions, we characterize the optimal input covariance matrix at the source and the optimal power allocation scheme at the relay that maximize the achievable rate. First, for the individual power constraint scenario, it is shown that the capacity of the AF system is achieved by either a direct transmission (DT) scheme, a NAF beamforming (BF) protocol with a unit-rank covariance matrix, or a NAF system using a specific full-rank covariance matrix. Then, for the global power constraint scenario, it is shown that only a DT or a NAF-BF protocol can achieve the capacity. In both cases, orthogonal transmission is shown to be strictly suboptimal. The capacity of the NAF system is also analyzed for some concrete examples, such as under asymptotically high and low transmission powers and for several network models.

Considering Rayleigh fading channels, we propose in Chapter 4 a general method to analyze the achievable rate and to characterize the optimal power allocation scheme for a wide range of HD AF protocols. At high transmission power regimes, our main idea is to exploit the capacity of a two-branch maximal-ratio combining (MRC) system to obtain tight yet simple approximations to the achievable rates. In low transmission power regions, we use a simple approximation to the logarithm to compute the asymptotic achievable rates. For all considered AF protocols, the closed-form approximations are tight and easy to analyze, since they involve only the exponential integral. More importantly and different from previous works, the proposed approach is applicable to OW and TW protocols using either the CI or the FG coefficient. Then, using the derived approximations, we compare the FG and CI coefficients in high and low power regimes. In addition, we analytically quantify the asymptotic power allocation schemes among the nodes to achieve the maximum rate and sum rate for OW and TW schemes, respectively. Comparisons among the AF protocols
1 Introduction

of interest are also carried out.

As the CI and FG coefficients are simply power normalization factors, Chapter 5 develops optimal amplification coefficients at the relay for HD cooperative and TW AF systems. The achievable rate and sum rate criteria for OW and TW relaying are considered, respectively. In particular, assuming full knowledge of all channel gains at the relay and Gaussian inputs at the source, optimal power adaptation schemes under a long-term average power constraint are established in closed-form. Important insights on the proposed adaptation schemes with respect to the channel gains are also presented and discussed. Numerical results reveal that the derived relay adaptation techniques outperform the conventional FG and CI coefficients, thanks to the benefit of dynamic power allocation.

Turning our attention to the performance with finite constellations, in Chapter 6 we propose the idea of precoding over multiple cooperative frames to improve the diversity and approach the capacity of a HD NAF system over Rayleigh fading channels. At first, we derive a tight union bound on the bit error rate (BER) of the system by adopting the coded framework of BICM. Focusing on the error-floor region, we then show that a significantly higher diversity order can be achieved by precoding over multiple frames. To maximize the asymptotic coding gain, an optimal class of precoders is derived along with a pragmatic approach to obtain good rotation angles for such class. Subsequently, concentrating on the turbo pinch-off region, we demonstrate that a concatenation of multi-dimensional (multi-D) mapping and multiple-frame precoding can be used to approach the capacity of the HD NAF channel. In particular, for various spectral efficiencies, we obtain a BER of $10^{-5}$ or lower at a signal-to-noise ratio (SNR) that is within a few dB from the ergodic capacity with Gaussian inputs.

Concentrating on the FD mode of operation, Chapter 7 investigates the optimal power allocation scheme and corresponding capacity limit of a FD DH system over static channel gains. Assuming that the residual self-interference variance is proportional to the $\lambda$-th power of the transmitted power, the rate maximization problems under both per-node and sum power constraints are first shown to be quasiconcave. Given the non-linearity of the derivative, bisection is then proposed to obtain the optimal power strategies. The capacity and optimal schemes are then analyzed in different high power regions. Specifically, we apply the dominant balance method to show that full power at the relay is suboptimal when its power constraint approaches a large value. Following a similar approach, we then show that the multiplexing gain of the FD scheme with the optimal allocation is $1/(1 + \lambda)$. 
Comparisons between the HD and FD systems are finally carried out, where analytical and simulation results reveal that the FD system is superior in high source power regions with either fixed or large power constraints at the relay.

Still considering FD transmission and under the same interference model as above, Chapter 8 investigates the error and diversity performance of FD AF systems over Rayleigh fading channels. The study focuses on the cooperative linear relaying (LR) protocol with direct source-destination link and the DH scheme without direct link, both under uncoded and coded frameworks. At first, closed-form pairwise error probability (PEP) expressions are derived for the uncoded systems, which are then used to obtain tight bounds to the BER of the coded systems. To shed an insight on the diversity behavior, asymptotic expressions at high transmission powers are also presented. Different from previous works that treat the direct link as interference, we show that the FD LR systems with suitable precoder can attain the same diversity function as their HD counterparts. A non-zero diversity order is thus attained and no error floor is observed despite the existence of self-interference in FD. Furthermore, it is shown that the diversity order of FD DH systems is a decreasing function of $\lambda$ and is equal to zero only when $\lambda = 1$. Although HD relaying is shown to be asymptotically optimal for both relaying protocols, illustrative results show that FD relaying is advantageous at practical BER levels when $\lambda$ is sufficiently small.

Finally, Chapter 9 concludes the thesis by restating our contributions and providing suggestions for further studies.
Chapter 2

Background and Relay Protocols

This chapter presents the background material required for the developments of subsequent chapters. These include a brief overview of wireless channels, performance metrics such as error rate and achievable rate, and BICM techniques. Key concepts such as diversity and capacity are also introduced. Finally, the input-output relations for the relay schemes considered in this thesis are presented.

2.1 Wireless Channel

In wireless communications, the channel is characterized by variations of the channel strength over both time and frequency. These variations can be divided into two types [39]:

- **Large-scale fading**: This type of fading is caused by the path-loss of the signal as a function of distance from transmitter to receiver, and shadowing by large obstacles (such as buildings or hills) between transmitter and receiver. These variations occur over large distances, i.e., several meters.

- **Small-scale fading**: This kind of fading is caused by the constructive and destructive addition of multiple signal paths between transmitter and receiver. These variations occur over short distances, i.e., on the order of the carrier wavelength.

In this thesis, small-fading is of particular interest as it is the main challenge in designing wireless communication systems.
Consider a system composed of a single transmitter-receiver pair. When the multi-path components due to small-scale fading arrive at the receiver in a time interval which is much shorter than the symbol period, the discrete-time baseband equivalent input-output relation can be written as

\[ y = \sqrt{P_s} h x + n, \quad (2.1) \]

where \( y \) is the received signal, \( x \) is the transmitted symbol, \( P_s \) is the average transmitted power, \( h \) is the complex channel gain between transmitter and receiver, and \( n \) is the noise sample. The model in (2.1) is suitable for narrow-band transmission and is commonly referred to as frequency-flat (or non-selective) fading channel.

Due to the central limit theorem, a standard assumption in wireless systems is that the noise is a zero-mean additive white Gaussian noise (AWGN) process. The noise samples \( n \) are hence independent over time, and have independent Gaussian distributed real and imaginary components with variance \( N_0/2 \). This can be modeled by a circular symmetric complex Gaussian random variable, denoted as \( \mathcal{CN}(0, N_0) \). Assuming a large number of statistically independent reflected and scattered paths with random amplitudes, the channel gain can also be modeled by a circular Gaussian random variable as \( h \sim \mathcal{CN}(0, \phi) \). In this case, the magnitude and phase of the channel are Rayleigh and uniform distributed, respectively. This simple, yet important probabilistic fading model is called Rayleigh fading and applies to situations with no line-of-sight path from transmitter to receiver and where several small reflectors scatter the signal before it arrives at the receiver.

Another important property of wireless channels is how fast they change over time. The time interval over which these gains change significantly is known as the coherence time \( T_c \). In this regard, wireless channels can be categorized as fast or slow fading. In fast fading, the coherence time is much shorter than the delay requirement for the considered application and the transmitted codeword spans over multiple channel fades. On the other hand, in slow fading, the coherence time is larger than the delay requirement and the transmitted codeword spans over a single fade. Note that in a non-faded or static environment, \( h \) is a time-invariant constant and therefore \( T_c \to \infty \).
2.2 Error Performance and Diversity

The bit error rate (BER) and symbol error rate (SER) are two common measures of error performance in communication systems. For instance, consider an uncoded system (i.e., without channel coding) using a general unit-energy $M$-QAM constellation $\Omega$. In the static non-faded scenario where $h$ in (2.1) is constant, the SER can be upper bounded using the union bound as

$$P_M \leq \frac{1}{M} \sum_{s \in \Omega} \sum_{\hat{s} \in \Omega \atop \hat{s} \neq s} \mathbb{P}(s \rightarrow \hat{s}), \quad (2.2)$$

where $Q(\cdot)$ is the tail probability of a standard normal distribution

$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} \exp \left(-\frac{u^2}{2}\right) \, du. \quad (2.3)$$

In (2.2), $\mathbb{P}(s \rightarrow \hat{s})$ is known as the pairwise error probability (PEP), which is the probability of deciding in favor of $\hat{s}$ given that $s$ was transmitted ($s, \hat{s} \in \Omega, s \neq \hat{s}$) and can be expressed as

$$\mathbb{P}(s \rightarrow \hat{s}) = Q\left(\sqrt{\frac{d^2(s, \hat{s})}{2N_0}}\right),$$

where $d^2(s, \hat{s}) = P_s|h|^2 \cdot |s - \hat{s}|^2$ is the squared Euclidean distance between $s$ and $\hat{s}$ in the absence of noise. The SER in (2.2) is dominated by the worst-case PEP and can be further upper bounded as

$$P_M \leq (M - 1) \cdot Q\left(\sqrt{\frac{d_{\text{min}}^2}{2N_0}}\right),$$

with $d_{\text{min}}^2 = \min_{s, \hat{s} \in \Omega \atop s \neq \hat{s}} P_s|h|^2 \cdot |s - \hat{s}|^2$. Note that for a binary modulation with $M = 2$, the SER is equal to the BER $P_b$ and the above bound is exact $P_b = P_M = Q(\sqrt{d_{\text{min}}^2/2N_0})$. For $M > 2$, the BER can be approximated from the SER by assuming Gray mapping as $P_b \approx P_M/\log(M)$ [40]. From the Chernoff bound to (2.3), it is known $Q(\sqrt{2x}) < [1/2] \exp(-x)$. The BER and SER then decay exponentially with $P_s$ in this static environment.

In a faded scenario, the wireless channel in (2.1) might suffer from a sudden decline in receive power. This is because there is a non-zero probability that the magnitude of the
channel gain $|h|$ drops dramatically. The channel is then said to be in deep fading and the average probability of error is severely degraded. In particular, the SER over Rayleigh fading ($h \sim \mathcal{CN}(0, \phi)$) can be approximated for high transmission powers as

$$P_M \approx a_M N_0 P_s \phi, \quad (2.4)$$

where $a_M$ is a constant that depends on the modulation scheme [40]. Different from the static case, observe from (2.4) that the SER decays only linearly with $P_s$ due to the deleterious effects of fading.

An effective way to combat the adverse effects of fading is by applying diversity techniques. The idea behind diversity is to transmit the same information over multiple independently faded paths. The probability that all paths experience a deep fade simultaneously is then greatly reduced. Thus, with appropriate combining at the receiver, the error performance over fading channels can be improved significantly. These independent paths can be obtained through several dimensions. For instance, time diversity can be exploited by applying error-control coding and interleaving over multiple coherence intervals so that different parts of the codeword experience different fades. In a similar fashion, frequency diversity techniques can be used by transmitting over multiple subcarriers in a frequency-selective environment. As an alternative, space diversity can be obtained when the transmitter or receiver are equipped with multiple separated antennas. As explained before, space diversity can also be exploited in cooperative relaying systems. Another method to extract diversity is through the signal space domain [41]. In signal space diversity, a group of consecutive symbols is first mapped to a signal point in a multi-dimensional (multi-D) constellation. The constellation is then rotated by a precoding matrix and sent over different channel fades. In general, this technique can be applied over time, frequency or space.

To quantify the effectiveness of diversity techniques, the diversity order is defined as

$$\delta = - \lim_{P_s \to \infty} \frac{\log(P_b)}{\log(P_s)},$$

which implies that the BER decreases as the $\delta$-th power of the transmitted power. The
BER can then be approximated in high power regions as

\[ P_b \approx \mathcal{F} \cdot P_s^{-\delta}, \]

where \( P_s^{-\delta} \) is the diversity gain function, and \( \mathcal{F}^{-1} \) is the coding gain. For example, \( \mathcal{F} = a_M N_0 / \phi \) and \( \delta = 1 \) in (2.4).

### 2.3 Achievable Rates and Capacity

The concept of reliable communication was introduced by Claude Shannon in his 1948 seminal work [42]. In particular, Shannon’s theory showed that it is possible to communicate at a strictly positive transmission rate with an arbitrary small error probability. The maximum rate at which this can be done is called the capacity of the channel, and any rate below the capacity is an achievable rate. Additionally, Shannon showed that it is impossible to drive the error probability to zero for communication rates higher than the capacity.

#### 2.3.1 Single-input Single-output Channel

In general, the capacity of the Gaussian channel introduced in (2.1) depends on the channel conditions. In the static scenario, the power-constrained capacity \( C \) of (2.1) can be obtained by solving the following optimization problem

\[
C = \max_{f_X(\cdot)} I(x, y) \quad \text{s.t.} \quad \mathbb{E}[|x|^2] \leq 1, \tag{2.5}
\]

where \( I(x, y) \) is the mutual information between the input \( x \) and the output \( y \) of the channel; \( P_s \) is the average power constraint from (2.1); and the maximization is carried over all input probability density functions (PDFs) \( f_X(\cdot) \) that satisfy \( \mathbb{E}[|x|^2] = \int |x|^2 f_X(x) dx \leq 1 \). As proved in [42], the above capacity is achieved by transmitting with full power a randomly selected codeword from a Gaussian codebook, i.e., a codebook composed of independent realizations of \( \mathcal{CN}(0, P_s) \). The capacity of (2.5) in b/s/Hz is then given by the well-known formula

\[
C = \log \left(1 + \frac{P_s |h|^2}{N_0}\right). \tag{2.6}
\]
In high power regions, it is easy to show that the capacity can be approximated as

$$C \approx m \cdot \log(P_s) + G,$$

where $m$ is known as the multiplexing gain and $G$ as the power gain. From (2.6), $m = 1$ and the capacity increases logarithmically with $P_s$. In low power regions, using the fact that $\ln(1 + x) \approx x$ for small $x > 0$,

$$C \approx \frac{P_s|h|^2}{N_0 \ln(2)},$$

and the capacity decreases linearly with $P_s$. Note that in this static scenario, the channel gain $h$ is constant and it can be safely assumed to be known at the transmitter and receiver.

In a fast fading environment where the transmitted codeword spans several coherence periods, different cases of channel state information (CSI) must be considered depending on which nodes track the channel values. For instance, when the receiver is able to perfectly track $h$ but cannot feed back this information to the transmitter, the knowledge of $h$ can be considered as a channel output and the mutual information in (2.1) must be replaced by $I(x, \{y, h\})$. By the chain rule of mutual information and from the independence of $x$ and $h$,

$$I(x, \{y, h\}) = I(x, h) + I(x, y|h) = I(x, y|h) = \mathbb{E}_h[I(x, y|h = \hat{h})],$$

where $I(x, y|h = \hat{h})$ is the conditional mutual information given a realization $\hat{h}$ of the random variable $h$. The capacity problem in this case can be written as

$$C = \max_{f_X(\cdot)} \mathbb{E}_h[I(x, y|h = \hat{h})] \quad \text{s.t.} \quad \mathbb{E}[|x|^2] \leq 1. \quad (2.7)$$

Similar to (2.5), the above capacity is achieved by Gaussian inputs and hence

$$C = \mathbb{E}_h \left[ \log \left( 1 + \frac{P_s|h|^2}{N_0} \right) \right].$$

The capacity in (2.7) is referred to as the ergodic capacity as the codeword must be large enough to capture the ergodic behavior of the channel.

When the CSI can be acquired by both the receiver and transmitter, the latter can use this information to adapt to the channel conditions. Consider the case in which the
transmitted codeword spans over $N$ coherence intervals, and let $q_i P_s$ be the power that the transmitter allocates to the $i$-th coherence period. This can be modeled as a parallel Gaussian channel and the capacity under a long-term power constraint (i.e., the power used over the $N$ coherence intervals) is the solution to

$$C = \max_{q_i \geq 0} \frac{1}{N} \sum_{i=1}^{N} \log \left( 1 + \frac{q_i P_s |h_i|^2}{N_0} \right) \quad \text{s.t.} \quad \frac{1}{N} \sum_{i=1}^{N} q_i \leq 1,$$

where $h_i$ is the $i$-th channel realization. The capacity in (2.8) is achieved by the well-known water-filling policy [43]:

$$q_i^* = \left( \mu^* - \frac{N_0}{P_s |h_i|} \right)^+, \quad (2.9)$$

where $[x]^+ = \max\{0, x\}$, and $\mu^*$ is a constant that satisfies the power constraint

$$\frac{1}{N} \sum_{i=1}^{N} \left( \mu^* - \frac{N_0}{P_s |h_i|} \right)^+ = 1.$$

The solution in (2.9) holds for any parallel channel as long as the transmitter has knowledge of $h_i$ for all $1 \leq i \leq N$. However, when $N$ is large enough to capture the ergodicity of the channel,

$$\frac{1}{N} \sum_{i=1}^{N} \log \left( 1 + \frac{q_i^* P_s |h_i|^2}{N_0} \right) \to \mathbb{E}_{h_i} \left[ \log \left( 1 + \frac{q_i^* P_s |h_i|^2}{N_0} \right) \right],$$

$$\frac{1}{N} \sum_{i=1}^{N} \left( \mu^* - \frac{N_0}{P_s |h_i|} \right)^+ \to \mathbb{E}_{h_i} \left[ \left( \mu^* - \frac{N_0}{P_s |h_i|} \right)^+ \right] = 1.$$

In this case, $\mu^*$ depends only on the channel statistics rather than on future values of $h_i$, and $q_i^*$ becomes a function of the present value of $h_i$ and its distribution [39].

In a slow fading environment where the transmitted codeword spans a single coherence interval, the ergodic capacity is theoretically zero. This is because for a given channel realization, the maximum rate supported by the channel is given as in the static case by $\log(1 + P_s |h|^2 / N_0)$. Given that $\log(1 + P_s |h|^2 / N_0)$ is a function of the random variable $h$, it is also random and there is a non-zero probability that it falls below the required rate $I$. 
The system is then said to be in \textit{outage} and the outage probability is defined as

\[
p_{\text{out}} = \mathbb{P} \left[ \log \left( 1 + \frac{P_s |h|^2}{N_0} \right) < I \right].
\] \hspace{1cm} (2.10)

It is important to note from (2.10) that minimizing the outage with full CSI at the transmitter and receiver is equivalent to maximizing the achievable rate of a static channel under the same CSI assumptions [44]. The solution of the static case is also equivalent to maximizing the rate in (2.8) under a \textit{short-term} power constraint, i.e., a power constraint over a single coherence period with \( q_i \leq 1 \).

\subsection*{2.3.2 MIMO Channel}

Consider now a system where the transmitter and receiver have multiple antennas. Specifically, let the transmitter and receiver have \( M_t \) and \( M_r \) antennas, respectively. The input-output relation of such a multiple-input multiple-output (MIMO) system can be written as

\[
y = \sqrt{P_s} H x + n,
\] \hspace{1cm} (2.11)

where \( y = [y_1, \ldots, y_{M_r}]^\top \) is the received signal; \( x = [x_1, \ldots, x_{M_t}] \) is the transmitted signal; \( H \) is the \( M_r \times M_t \) channel matrix with channel gains \( h_{ij} \) from transmit antenna \( j \) to receive antenna \( i \); and the noise vector \( n \sim \mathcal{CN}(0, K) \) with covariance matrix \( K = N_0 I_{M_r} \).

The capacity of the MIMO system in (2.11) was analyzed by Telatar in his seminal work in [45]. Specifically, under a static environment where \( H \) is deterministic, it was shown in [45] that the mutual information between \( x \) and \( y \) is maximized by Gaussian inputs and can be written as

\[
I(x, y) = \log \det \left( I_{M_t} + P_s H^\dagger K^{-1} H Q \right),
\] \hspace{1cm} (2.12)

where \( Q = \mathbb{E}[x x^\dagger] \) is the positive semidefinite (psd) covariance matrix of \( x \). The power-constrained capacity is then

\[
C = \max_{Q \succeq 0} \log \det \left( I_{M_t} + \frac{P_s}{N_0} H^\dagger H Q \right) \quad \text{s.t.} \quad \text{tr}(Q) \leq 1.
\] \hspace{1cm} (2.13)
Let the eigenvalue decomposition of $H^\dagger H = U^\dagger \Lambda U$, where $U$ is unitary and $\Lambda = \text{diag}(\Lambda_1, \ldots, \Lambda_M_t)$ is diagonal. Since the diagonal components $\Lambda_j$ are the square of the singular values of $H$, they are real and non-negative. As proved in [45], the capacity in (2.13) is achieved by diagonalizing the channel by means of $Q^* = U^\dagger \tilde{Q}^* U$, where $\tilde{Q}^* = \text{diag}(\tilde{q}_1^*, \ldots, \tilde{q}_{M_t}^*)$ and the diagonal entries $\tilde{q}_j^*$ are the solution to

$$C = \max_{\tilde{q}_j \geq 0} \sum_{j=1}^{M_t} \log \left( 1 + \frac{\tilde{q}_j P_s \Lambda_j}{N_0} \right) \quad \text{s.t.} \quad \sum_{j=1}^{M_t} \tilde{q}_j \leq 1. \quad (2.14)$$

The problem in (2.14) is the same as the one for the parallel channel in (2.8) and thus water-filling over the $\Lambda_j$ values is optimal. As noted before, this solution also minimizes the outage probability in slow fading environments with full CSI at transmitter and receiver.

For the fast fading scenario with CSI at the receiver only, the capacity can be obtained similar to (2.7) by solving the following problem:

$$C = \max_{Q \succeq 0} \mathbb{E}_H \left[ \log \det \left( I_{M_t} + \frac{P_s}{N_0} H^\dagger H Q \right) \right] \quad \text{s.t.} \quad \text{tr}(Q) \leq 1.$$ 

As demonstrated by Telatar [45], when each component of the channel matrix $h_{ij} \sim \mathcal{CN}(0, \phi)$, transmitting independent streams from each antenna with full-power is optimal, that is $Q^* = (1/M_t) \cdot I_{M_t}$. Furthermore, when both transmitter and receiver have full CSI, the system can be modeled as a parallel channel similar to (2.14) and water-filling over both space and time becomes the capacity achieving scheme under a long-term power constraint. In the case of a short-term power constraint, the solution to the static case in (2.14) remains optimal.

### 2.3.3 Two-way Channel

All of the above concepts apply for OW communications, i.e., when one user wants to communicate to another. However, these concepts can be extended to TW schemes where two nodes want to exchange information. Consider a system with two single-antenna nodes that alternate in transmission. Let the signal received at node $i \in \{1, 2\}$ be given by

$$y_i = \sqrt{P_s} h x_k + n_i,$$
where $x_k$ is the signal transmitted by user $k$ with power $P_{sk}$ ($k \in \{1, 2\} \neq i$), and $n_i$ is the noise sample at node $i$. Analogous to the capacity concept, the capacity region is defined as the set of all achievable rates $\{I(x_1, y_2), I(x_2, y_1)\}$ such that the two users can communicate reliably. Given that the two users share the same channel, this results in a trade-off between the rates of the users, i.e., an increase in the achievable rate of one user is attained at the expense of the other. Several scalar performance measures can be derived from the capacity region. For instance, the static and ergodic sum capacity are respectively defined as

$$C_{\text{sum}} = \max_{f_{X_1}(\cdot), f_{X_2}(\cdot)} I(x_1, y_2) + I(x_2, y_1),$$

$$C_{\text{sum}} = \max_{f_{X_1}(\cdot), f_{X_2}(\cdot)} \mathbb{E}_h[I(x_1, y_2|h = \hat{h}) + I(x_2, y_1|h = \hat{h})],$$  \hspace{1cm} (2.15)$$

where $f_{X_k}(\cdot)$ is the PDF of the input $x_k$; and $I(x_1, y_2) + I(x_2, y_1)$ and $\mathbb{E}_h[I(x_1, y_2|h = \hat{h}) + I(x_2, y_1|h = \hat{h})]$ are achievable sum rates.

Before closing this section, it should be noted that for practical systems using finite inputs (such as QAM or QPSK), a strong and long channel code is needed to approach the capacity with Gaussian inputs. As such, the analysis with Gaussian codebooks implicitly assumes strong channel coding with sufficiently long codewords.

### 2.4 BICM, BICM-ID and EXIT Charts

Among different coding techniques, BICM (originally introduced by Zehavi in [46]) has been considered as a state-of-the-art coded modulation structure to extract time diversity over fading channels [46–52]. The basic idea of BICM is to concatenate a channel encoder and a random bit interleaver that spans over several coherence periods. A diversity equal to the minimum Hamming distance between codewords of the channel code can then be obtained. In BICM, the channel encoder and the modulator can be designed separately, allowing more flexibility in the system design. Furthermore, the performance of BICM systems over fading channels can be significantly improved through iterative decoding (ID) [48,49] (first proposed in [48]).

Extrinsic information transfer (EXIT) charts [53] provide a method to analyze the convergence behavior of iterative decoding systems such as BICM-ID. Specifically, they
can be used to predict important features in the BER curve. For example, these charts are useful to predict the signal-to-noise ratio (SNR) $P_s/N_0$ range at which a sudden BER drop is attained over iterations. This range is known as the pinch-off or waterfall region. The region for SNRs higher than the pinch-off area is known as the error-floor region. Besides the waterfall region, EXIT charts can also predict the number of iterations required for convergence and whether the system will converge to a high or low BER level.

In this thesis, we do not provide a comprehensive survey on BICM, BICM-ID and EXIT charts and the interested reader is referred to [54] for further details.

### 2.5 Input-Output Relations

We now present the input-output relations for the considered relaying systems. These relations will be used throughout the thesis.

The considered system, shown in Fig. 2.1, consists of three single-antenna nodes: a relay node $R$, and two HD source nodes $S_1$ and $S_2$. The relay node $R$ assists in the transmission between $S_1$ and $S_2$, and might operate in either HD or FD mode. In the FD mode, the residual self-interference is explicitly taken into account. The channel gain from $S_1$ to $S_2$ is denoted by $h_0$, whereas the gains from $S_1$ to $R$ and $S_2$ to $R$ are given respectively by $h_1$ and $h_2$. These channel gains are assumed to be reciprocal so that the gains from node $A$ to node $B$ and from $B$ to $A$ are the same, $A,B \in \{S_1, S_2, R\}$ ($A \neq B$). For simplicity, we denote $h_l = \sqrt{\alpha_l}e^{j\theta_l}$, where $\alpha_l$ and $\theta_l$ are the magnitude squared and phase of $h_l$, $l \in \{0,1,2\}$.

In this thesis, we consider frequency-flat Rayleigh fading so that the channel gains are independent zero-mean circularly Gaussian distributed which is denoted as $h_l \sim \mathcal{CN}(0, \phi_l)$, i.e., with Rayleigh distributed magnitude. Furthermore, we adopt the block fading model such that the channel gains $\mathbf{h} = [h_0, h_1, h_2]$ remain constant for a coherence interval $T_c$ (given in number of symbol periods) and change independently after. We consider both the static scenario (where transmitted codeword spans over a time-invariant $\mathbf{h}$) and the fast fading one (where it spans over multiple realizations of $\mathbf{h}$). We first introduce the HD protocols of interest before describing the FD ones.

#### 2.5.1 HD Protocols

The considered HD protocols are summarized in Table 2.1 and are described in the following.
Fig. 2.1 Channel gains.

Table 2.1 HD relay protocols.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Phase 1</th>
<th>Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-orthogonal</td>
<td>$S_1 \rightarrow {R, S_2}$</td>
<td>$S_1 \rightarrow S_2$</td>
</tr>
<tr>
<td>DT</td>
<td>$S_1 \rightarrow S_2$</td>
<td>$S_1 \rightarrow S_2$</td>
</tr>
<tr>
<td>Orthogonal</td>
<td>$S_1 \rightarrow {R, S_2}$</td>
<td>$R \rightarrow S_2$</td>
</tr>
<tr>
<td>Dual-hop</td>
<td>$S_1 \rightarrow R$</td>
<td>$R \rightarrow S_2$</td>
</tr>
<tr>
<td>Two-way</td>
<td>$S_1 \rightarrow R$</td>
<td>$R \rightarrow {S_1, S_2}$</td>
</tr>
</tbody>
</table>

NAF Protocol

We first introduce the OW non-orthogonal AF (NAF) protocol, which is obtained by combining non-orthogonal transmission in Fig. 1.1 with the AF relaying scheme. Specifically, we consider the NAF scheme studied in [20–22], which has been shown to be optimal in terms of diversity-multiplexing tradeoff over the single-relay AF channel [21]. Without loss of generality, we assume $S_1$ is the source node $S$ and $S_2$ the destination node $D$. In this protocol, transmission is carried in cooperative frames composed of two unit-time phases as shown in Table 2.1 (where $T_c \geq 2$ is a multiple of two). In the first phase, denoted as the broadcasting phase, $S$ sends the first signal $x_1$ to both $R$ and $D$. The received signals at $R$ and $D$ can be written respectively as

$$y_{r,1} = \sqrt{P_s h_1 x_1} + n_{r,1}, \quad \text{and} \quad y_1 = \sqrt{P_s h_0 x_1} + n_{d,1},$$

where $P_s$ is a constant related to the transmission power at $S$; and $n_{r,1}$ and $n_{d,1}$ are the independent noise samples with respective variance $N_r$ and $N_d$, denoted as $n_{r,1} \sim \mathcal{CN}(0, N_r)$.
and $n_{d,1} \sim \mathcal{CN}(0, N_d)$. In the second or cooperative phase, $S$ sends the signal $x_2$ to $D$ while $R$ amplifies and forwards the symbol received during the first phase to $D$ using an amplification coefficient $b$. The received signal at $D$ in the second phase is expressed as

$$y_2 = \sqrt{P_s h_0 x_2} + \sqrt{P_r b h_2 y_{r,1}} + n_{d,2} = \sqrt{P_s h_0 x_2} + \sqrt{P_r b h_2 (\sqrt{P_s h_1 x_1} + n_{r,1})} + n_{d,2},$$

where $P_r$ is a constant related to the transmission power at $R$ and $n_{d,2} \sim \mathcal{CN}(0, N_d)$.

The signal received at $D$ over these two phases, $y = [y_1, y_2]^\top$, can be written in matrix form as in [22]:

$$y = \sqrt{P_s} H_{NAF} x + n,$$

where $x = [x_1, x_2]^\top$ is the input vector; $B = \begin{pmatrix} 0 & 0 \\ b & 0 \end{pmatrix}$; $n_r = [n_{r,1}, n_{r,2}]^\top \sim \mathcal{CN}(0, N_r I_2)$ is the noise vector at $R$; and $n_d = [n_{d,1}, n_{d,2}]^\top \sim \mathcal{CN}(0, N_d I_2)$ is the noise vector at $D$. The equivalent $2 \times 2$ channel matrix in (2.16) is thus

$$H_{NAF} = \begin{pmatrix} h_0 & 0 \\ \sqrt{P_r b h_2} & h_0 \end{pmatrix},$$

and the equivalent noise vector $n = [n_1, n_2]^\top = [n_{d,1}, \sqrt{P_r b h_2 n_{r,1}} + n_{d,2}]^\top \sim \mathcal{CN}(0, K)$ with

$$K = \begin{pmatrix} N_d & 0 \\ 0 & N_d + P_r b^2 \alpha_2 N_r \end{pmatrix}.$$

Let the temporal covariance matrix of the input vector $x$ be denoted as

$$Q = \mathbb{E}[xx^\top] = \begin{pmatrix} q_1 & q_{12} \\ q_{12} & q_2 \end{pmatrix} \succeq 0.$$  

From the covariance matrix and (2.16), it is straightforward to see that $S$ allocates a power of $q_1 P_s$ to $x_1$ in the first phase and $q_2 P_s$ to $x_2$ in the second phase. The average transmitted power at $S$ is then $P_s \cdot \text{tr}(Q) = P_s (q_1 + q_2)$ per transmission frame or $(P_s/2) \cdot \text{tr}(Q) =$
\[ P_s(q_1 + q_2)/2 \] per symbol period. Now, let \( z_2 P_r \) be the average power constraint at \( R \) in the second phase, i.e., \( \mathbb{E}[b^2|\sqrt{P_s}h_1x_1 + n_{r,1}|^2] = z_2 \). The amplification coefficient can then be expressed as

\[
b = \sqrt{\frac{z_2}{q_1 P_s \Xi_1 + N_r}}, \tag{2.19}
\]

where the parameter \( \Xi_1 \) depends on the CSI available at the relay. For instance, when the relay only knows the second order statistics of the \( S-R \) channel, the FG CDI coefficient is obtained by setting \( \Xi_1 = \phi_1 \). On the other hand, when the relay has instantaneous knowledge of the \( S-R \) link, the VG CI coefficient is given with \( \Xi_1 = \alpha_1 \). Note that for the FG coefficient, the power transmitted by the relay during a coherence interval \( T_c \) might be larger or smaller than \( z_2 P_r \) depending on the value of \( h \). However, the power averaged over several realizations of \( h \) remains \( z_2 P_r \).

As a shorthand notation, let the mutual information between the input \( x \) and output \( y \) vectors of (2.16) conditioned on a realization of \( h \) be given by \( I|h = I(x, y|h = \hat{h}) \). Assuming Gaussian codebooks at the source, the conditional achievable rate can be written from (2.12) and (2.16) as

\[
I|h = \frac{1}{2} \log \det\left( I_2 + P_s H_{\text{NAF}}^\dagger K^{-1} H_{\text{NAF}} Q \right), \tag{2.20}
\]

with \( K \) as in (2.18) and \( H_{\text{NAF}} \) as in (2.17). In (2.20), the \( 1/2 \) pre-log factor is due to the fact that the transmission protocol is carried over two phases. The unconditional mutual information or ergodic achievable rate can then be obtained by averaging (2.20) over the channel gains, i.e., \( I = \mathbb{E}[I|h] \).

Note that the NAF protocol introduced above is general, in the sense that it includes the direct transmission scheme and other OW relay protocols as explained below.

**DT Protocol**

In the direct transmission (DT) protocol, as shown in Table 2.1, the relay is not used for transmission \( b = 0 \) and the source simply communicates its information through the direct \( S-D \) link. Although the DT scheme is not a relay protocol, it is an important alternative in cooperative systems where using the relay might not always be beneficial. By substituting
In (2.16), the input-output relation for the DT scheme is given by

\[ y = \sqrt{P_s} \begin{pmatrix} h_0 & 0 \\ 0 & h_0 \end{pmatrix} x + n_d. \]  

(2.21)

The conditional rate in (2.20) when \( x_1 \) and \( x_2 \) are independent (\( q_{12} = 0 \)) and \( z_2 = 0 \) reduces to

\[ I|h = \frac{1}{2} \left[ \log \left( 1 + \frac{q_1 \alpha_0 P_s}{N_d} \right) + \log \left( 1 + \frac{q_2 \alpha_0 P_s}{N_d} \right) \right]. \]  

(2.22)

**OAF Protocol**

We now describe the orthogonal AF (OAF) in which the source and relay alternate for transmission as shown in Table 2.1. This protocol is given by combining orthogonal transmission in Fig. 1.1 with AF relaying and can be obtained by setting \( q_2 = 0 \) in the NAF protocol. Specifically, the input-output relation in (2.16) simplifies when \( q_2 = 0 \) to

\[ y = \sqrt{P_s} h_{OAF} x_1 + n, \]  

(2.23)

where \( h_{OAF} = [h_0, \sqrt{P_r} b h_1 h_2]^\top \) is the first column of (2.17) and the covariance matrix of \( n \) is still given as in (2.18). The conditional achievable rate of the OAF system is then given from (2.23) as

\[ I|h = \frac{1}{2} \log \left( 1 + \frac{q_1 \alpha_0 P_s}{N_d} + \frac{\alpha_1 \alpha_2 q_1 b^2 P_s P_r}{N_d + \alpha_2 b^2 P_r N_r} \right). \]  

(2.24)

**DHAF Protocol**

In the dual-hop AF (DHAF) protocol in Fig. 1.1 or Table 2.1, the direct link from source to destination is either ignored or heavily shadowed so that it cannot be used for communication. The input-output relation in this case can be obtained by setting \( q_2 = 0 \) and \( h_0 = 0 \) in (2.16) as

\[ y = \sqrt{P_s P_r} b h_1 h_2 x_1 + n_2, \]  

(2.25)
where \( n_2 \sim \mathcal{CN}(0, N_d + b^2 \alpha_2 P_r N_r) \). The conditional rate of the DHAF system is then given by

\[
I | h = \frac{1}{2} \log \left( 1 + \frac{\alpha_1 \alpha_2 q_1 b^2 P_s P_r}{N_d + \alpha_2 b^2 P_r N_r} \right)
\]  

(2.26)

**TW AF Protocol**

We now turn our attention to the two-way AF (TWAF) protocol in which \( S_1 \) and \( S_2 \) want to exchange information as in Fig. 1.1. In particular, we consider the two-phase TW system in [23,24] as it provides a higher multiplexing gain than the three-phase scheme [55]. Transmission in this protocol is again carried in frames divided in two unit-time phases as shown in Table 2.1 (where \( T_c \geq 2 \) is a multiple of two). In the first phase, also referred to as the multiple-access phase, \( S_1 \) and \( S_2 \) transmit simultaneously their respective symbols \( x_1 \) and \( x_2 \) to \( R \). The received signal at \( R \) is given by

\[
y_{r,1} = \sqrt{P_{s1} h_1 x_1} + \sqrt{P_{s2} h_2 x_2} + n_{r,1},
\]

where \( n_{r,1} \sim \mathcal{CN}(0, N_r) \), and \( P_{si} \) is constant related to the power transmitted at node \( S_i \) (\( i \in \{1, 2\} \)). In the second phase, denoted as the broadcasting phase, \( R \) simply amplifies the received signal and broadcasts it to \( S_1 \) and \( S_2 \). The signal received at \( S_i \) can be written as

\[
y_i' = \sqrt{P_r b h_i y_{r,1}} + n_{d,i} = \sqrt{P_r b h_i (\sqrt{P_{s1} h_1 x_1} + \sqrt{P_{s2} h_2 x_2} + n_{r,1})} + n_{d,i},
\]

where \( n_{d,i} \sim \mathcal{CN}(0, N_{di}) \) is the noise sample at node \( i \). Assuming perfect knowledge of \( h \) and \( b \) at \( S_i \), the self-interference term created by its own transmitted signal \( x_i \) can be removed as \( y_i = y_i' - \sqrt{P_{si} P_r b h_i^2 x_i} \). The received signal at \( S_i \) after removing the self-interference can then be written as

\[
y_i = \sqrt{P_{sk} P_r b h_1 h_2 x_k} + \sqrt{P_r b h_i n_r} + n_{d,i}
= \sqrt{P_{sk} P_r b h_1 h_2 x_k} + n_i,
\]

(2.27)

where \( k = 3 - i \) and the equivalent noise \( n_i \sim \mathcal{CN}(0, N_{di} + \alpha_i b^2 P_r N_r) \).

Let \( q_i = \mathbb{E}[|x_i|^2] \) so that the average power allocated at \( S_i \) in the first phase is \( q_i P_{si} \). As before, two types of amplification coefficient can be used at relay depending on the
availability of channel knowledge:

\[ b = \sqrt{q_1 P_{s1} \Xi_1 + q_2 P_{s2} \Xi_2 + N_r}. \]  

(2.28)

In the case of the FG CDI technique \( \Xi = \phi \), whereas \( \Xi = \alpha \) for the VG CI coefficient.

The conditional achievable rate in the \( S_k \to S_i \) direction can be expressed from (2.27) as [23, 24]:

\[ I_i|h = \frac{1}{2} \log \left( 1 + q_k \alpha_1 \alpha_2 b^2 P_{sk} P_r \frac{N_{d_i}}{N_r} + \alpha_i b^2 P_r N_r \right). \]  

(2.29)

The conditional sum rate is then given as \( I_{\text{sum}}|h = I_1|h + I_2|h \) and the unconditional achievable sum rate is \( I_{\text{sum}} = \mathbb{E}[I_{\text{sum}}|h] \). It should be noted that the OW schemes introduced above can also be seen as TW protocols carried over four transmission phases, i.e., \( S_1 \to S_2 \) over the first two phases and \( S_2 \to S_1 \) over the remaining two.

2.5.2 FD Protocols

We now introduce the FD protocols of interest. As mentioned before, the residual self-interference due to FD transmission is taken into consideration.

LR Protocol

First, we describe the linear relaying (LR) protocol introduced in [56], which is obtained by combining cooperative relaying in Fig. 1.2 with AF transmission. The LR protocol can be considered as a generalization of NAF relaying to FD operation and has only been studied in terms of achievable rate under no self-interference [56, 57]. As before, we assume that \( S_1 \) is the source \( S \) and \( S_2 \) the destination \( D \). The transmission in LR is carried in frames composed of \( L \) symbol periods as shown in Table 2.2 (where \( T_c \geq L \) is a multiple of \( L \)). In each symbol period \( i \) (\( 1 \leq i \leq L \)), \( S \) transmits the information signal \( x_i \) to both the relay \( R \) and the destination \( D \). At the same time, \( R \) transmits a linear combination of symbols previously received in the same frame to \( D \).

Specifically, the signal received at the relay at time \( i \) can be expressed as

\[ y_{r,i} = \sqrt{P_s} h_1 x_i + n_{r,i} + v_i, \]  

(2.30)
Table 2.2  LR protocol.

<table>
<thead>
<tr>
<th>Phase</th>
<th>$S_1$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$S_1 \rightarrow {R, S_2}$</td>
<td>$-$</td>
</tr>
<tr>
<td>2</td>
<td>$S_1 \rightarrow {R, S_2}$</td>
<td>$R \rightarrow S_2$</td>
</tr>
<tr>
<td></td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$i$</td>
<td>$S_1 \rightarrow {R, S_2}$</td>
<td>$R \rightarrow S_2$</td>
</tr>
<tr>
<td></td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$L-1$</td>
<td>$S_1 \rightarrow {R, S_2}$</td>
<td>$R \rightarrow S_2$</td>
</tr>
<tr>
<td>$L$</td>
<td>$S_1 \rightarrow S_2$</td>
<td>$R \rightarrow S_2$</td>
</tr>
</tbody>
</table>

where $n_{r,i} \sim \mathcal{CN}(0, N_r)$; and $v_i$ is the residual self-interference term due to FD operation and after self-interference cancellation. In the following, we assume that this term is Gaussian as $v_i \sim \mathcal{CN}(0, V)$. Detailed discussions about the distribution and variance of $v_i$ are provided in Chapter 7. The signal transmitted from the relay at time $i \geq 2$ can be written as

$$t_i = \sum_{k=1}^{i-1} b_{i,k} y_{r,k},$$

where $b_{i,k}$ is the coefficient used at time $i$ to amplify the signal received at time $k$, $y_{r,k}$, $1 \leq k \leq i - 1$. The signal received at the destination is thus given by

$$y_i = \sqrt{P_s h_0 x_i} + \sqrt{P_r h_2 t_i} + n_{d,i}$$

$$= \sqrt{P_s h_0 x_i} + \sqrt{P_r h_2} \left( \sum_{k=1}^{i-1} b_{i,k} y_{r,k} \right) + n_{d,i}$$

$$= \sqrt{P_s h_0 x_i} + \sqrt{P_s P_r h_2} \left( \sum_{k=1}^{i-1} b_{i,k} x_k \right) + \sqrt{P_r h_2} \left( \sum_{k=1}^{i-1} b_{i,k} (n_{r,k} + v_k) \right) + n_{d,i},$$

where $n_{d,i} \sim \mathcal{CN}(0, N_d)$. It should be noted from Table 2.2 that the relay operates in HD mode in the first time slot, i.e., it receives $y_{r,1}$ but does not transmit ($t_1 = 0$). As such, $y_{r,1}$ in (2.30) does not have the extra residual self-interference term and $v_1 = 0$.

The amplification coefficients $b_{i,k}$ can be grouped into a strictly lower triangular $L \times L$
amplification matrix $B$ as

$$B = \begin{cases} b_{i,k}, & 2 \leq i \leq L, \ 1 \leq k \leq i - 1 \\ 0, & \text{o.w.} \end{cases}$$

The signal transmitted by the relay, $t = [0, t_2, \ldots, t_L]^{\top}$, can then be written in vector form as

$$t = By_r = B(\sqrt{P_s}h_1 x + n_r + v), \quad (2.31)$$

where $y_r = [y_{r,1}, \ldots, y_{r,L}]^{\top}$ is the vector received at the relay; $x = [x_1, \ldots, x_L]^{\top}$; $n_r = [n_{r,1}, \ldots, n_{r,L}]^{\top} \sim \mathcal{CN}(0, N_r I_L)$; and $v = [0, v_2, \ldots, v_L]^{\top} \sim \mathcal{CN}(0, K_v)$ is the self-interference vector due to FD operation with

$$K_v = \sqrt{P_r} \cdot \begin{pmatrix} 0 & 0 \\ 0 & I_{L-1} \end{pmatrix}. \quad (2.32)$$

The signal received at the destination over the $L$ time slots can be similarly expressed in matrix form as

$$y = \sqrt{P_s} \left( h_0 I_L + \sqrt{P_r} h_1 h_2 B \right) x + \sqrt{P_r} h_2 B(n_r + v) + n_d$$

$$= \sqrt{P_s} H_{LR} x + n, \quad (2.33)$$

where $n_d = [n_{d,1}, \ldots, n_{d,L}]^{\top} \sim \mathcal{CN}(0, N_d I_L)$. The equivalent $L \times L$ channel matrix in (2.33) is then given by

$$H_{LR} = h_0 I_L + \sqrt{P_r} h_1 h_2 B, \quad (2.34)$$

and the equivalent $L \times 1$ noise vector $n \sim \mathcal{CN}(0, K)$ with

$$K = P_r \alpha_2 B (N_r I_L + K_v) B^{\dagger} + N_d I_L. \quad (2.35)$$

Interestingly, the LR protocol in (2.33) reduces to the NAF in (2.16) for $L = 2$.

Let $Q = \mathbb{E}[xx^{\top}]$ with diagonal elements $q_i$ ($1 \leq i \leq L$). The source then transmits an
average power of \( P_s \cdot \text{tr}(Q) \) per transmission frame or \((P_s/L) \cdot \text{tr}(Q)\) per symbol period. Similarly, let the covariance matrix of the signal transmitted by the relay be denoted as

\[
Z = \mathbb{E}[tt^\dagger] = B(\alpha_1 P_s Q + N_r I_L + K_v)B^\dagger,
\]

(2.36)

with diagonal elements \( z_i \), \( (1 \leq i \leq L) \). Specifically, the diagonal values of (2.36) can be written for \( i \geq 2 \) as

\[
\begin{align*}
z_i &= \mathbb{E}[|t_i|^2] \equiv \mathbb{E} \left[ \sum_{k=1}^{i-1} b_{i,k} (h_1 \sqrt{P_s x_k + n_{r,k} + v_k})^2 \right] \\
&= N_r \left( \sum_{k=1}^{i-1} b_{i,k}^2 \right) + V \left( \sum_{k=2}^{i-1} b_{i,k}^2 \right) + P_s \alpha_1 \left( \sum_{k=1}^{i-1} \sum_{l=1}^{i-1} b_{i,k} b_{i,l} \mathbb{E}[x_k x_l^\dagger] \right).
\end{align*}
\]

The relay then transmits an average power of \( P_r \cdot \text{tr}(Z) \) per transmission frame or \((P_r/L) \cdot \text{tr}(Z)\) per symbol period.

Depending on the CSI available at the relay and on how the symbols \( y_{r,k} \) are superimposed, different amplification matrices \( B \) can be considered as discussed before. For instance, one can assume that the relay only amplifies the signal received in the previous period similar to [57]. In this case, \( b_{i,k} = b \) when \( k = i - 1 \) and \( b_{i,k} = 0 \) otherwise \((2 \leq i \leq L)\), where

\[
b = \sqrt{\frac{\text{tr}(Z)}{N_r(L-1) + V(L-2) + P_s \Xi_1 \sum_{i=2}^{L} q_{i-1}}}.
\]

(2.37)

As before, \( \Xi = \alpha \) when the relay has instantaneous knowledge of the \( S-R \) link and \( \Xi = \phi \) when only second order statistics are known.

Assuming Gaussian codebooks at the source, the conditional achievable rate can be written from (2.33) and (2.12) as

\[
I|h = \frac{1}{L} \log \det(I_L + H_{LR}^\dagger K^{-1} H_{LR} Q),
\]

(2.38)

with \( K \) as in (2.35) and \( H_{LR} \) as in (2.34).
2 Background and Relay Protocols

DHAFF Protocol

Next, we consider the FD DHAFF protocol in Fig. 1.2. In this protocol, the direct $S$-$D$ link is either non-existent $\alpha_0 = 0$ due to heavy shadowing or treated as a source of interference. In particular, the source continuously transmits its information symbol $x_i$ while the relay amplifies and forwards the signal received in the previous time slot with a power amplification $b$. Specifically, the signal received at $R$ at time $i$ can be written as in (2.30), whereas the signal transmitted by $R$ is now $t_i = by_{r,i-1}$ The signal received at $D$ can then be written as

$$y_i = \sqrt{P_sP_r}h_1h_2bx_{i-1} + \sqrt{P_r}h_2(b(n_{r,i-1} + v_{i-1}) + \sqrt{P_s}h_0x_i + n_{d,i}, \quad (2.39)$$

where $x_i$ is treated as interference when $\alpha_0 > 0$ and $x_{i-1}$ is the desired symbol.

Let $E[|x_i|^2] = q_1$ so that an average power of $q_1P_s$ is spent at the source. Assuming an average power of $z_2P_r$ at the relay, the amplification coefficient in (2.39) can be expressed as

$$b = \sqrt{\frac{z_2}{q_1P_s\Xi_1 + N_r + V}}. \quad (2.40)$$

Note that besides normalizing the power of the FD relay, the amplification coefficient in (2.40) is also needed prevent oscillation [58]. Assuming Gaussian codebooks at $S$, the conditional achievable rate can now be written from (2.39) as

$$I|h = \left(1 + \frac{\alpha_1\alpha_2q_1P_sP_r^2b^2}{\alpha_2P_r^2b^2[N_r + V] + N_d + \alpha_0q_1P_s}\right). \quad (2.41)$$

It can be observed that similar to the LR protocol, the relay of the DH protocol operates in HD mode in the first time slot. However, assuming that the DH transmission is carried continuously over several time slots, the effect of this transient state is negligible and can thus be ignored in (2.39).

2.6 Chapter Summary

In this chapter, background material related to the problems addressed in this thesis was presented. The wireless fading channel was first introduced. Different error and achiev-
able rate performance metrics were then discussed along with important concepts such as diversity and capacity. A brief description of BICM and BICM-ID was also provided. The input-output relations for the AF relaying protocols of interest were finally presented. The achievable rate and error performance of these protocols are analyzed in the following chapters.
Chapter 3

Capacity and Power Allocation of the Static Non-orthogonal AF Relay Channel\(^1\)

During the past few years, much attention has been paid to the capacity, together with optimal power strategies, of various OW and TW HD-AF protocols over static channel links. Extensive research has been devoted to dual- or multi-hop AF systems that do not take the direct source-destination link into account. For instance, power allocation strategies that maximize the conditional achievable rate have been addressed for different configurations in [61–65] and references within. In particular, optimal power sharing strategies between the source and the relay to maximize the achievable rate in (2.26) were derived in [61]. Significant research has also been devoted to power allocation strategies for orthogonal AF schemes that take the direct link into consideration [61, 66–68]. Specifically, the maximization of (2.24) over the powers at the source and relay has been carried in [61]. For TWAF relaying, power allocation strategies for static channels have been also investigated in [69–72]. Specifically, the optimal power sharing scheme that maximizes the conditional sum rate in (2.29) has been derived in [69].

On the other hand, the analysis and optimization of OW non-orthogonal networks are more challenging and have received less attention [20–22]. For instance, the power-

---

\(^1\)Parts of this chapter have been presented at the 2012 IEEE International Conference on Communications in Ottawa, Canada [59]; and have been published in the IEEE Transactions on Wireless Communications [60].
constrained capacity of a single-relay NAF channel was investigated in the fast fading scenario in [22]. Under the assumption of full CSI available only at the destination node, it was shown in [22] that the input distribution must be Gaussian with a diagonal covariance matrix to achieve the ergodic capacity. Due to the complexity of the problem, the diagonal elements of the covariance matrix can only be obtained numerically in [22] by solving optimization problems of multi-dimensional integrals. In addition, it was shown in [22] that the OAF and DT schemes are optimal in low and high SNR regimes, respectively. However, the maximization of the conditional rate for the static NAF channel in (2.20) has not been addressed in the literature.

As discussed in Section 2.3, the capacity of the static MIMO system has been well established in [45], where it was shown that the optimal signaling method is to water-fill over the square of the singular values of the channel matrix. Although the maximization for the NAF channel in (2.20) resembles that in static MIMO systems in (2.13), there are significant differences that make the relay problem more challenging. Specifically, observe from (2.20) that different from MIMO systems, the channel and noise covariance matrices are functions of the powers transmitted at the source and relay. Thus, the technique proposed in [45] for MIMO systems is not applicable to the static NAF channel. More importantly, the mutual information of the NAF system in (2.20) is no longer concave and conventional optimization methods cannot be applied to find the optimal input covariance matrix, the power allocation, and ultimately, the capacity. Due to these challenging issues, closed-form solutions on the capacity limit of the general static NAF channel remain unknown, to the best of our knowledge.

Motivated by the above observations, this chapter attempts to provide some new results on the fundamental limit of the static cooperative AF channel. In particular, this chapter analyzes the capacity and characterizes the optimal input covariance matrix at the source and power allocation scheme at the relay for a half-duplex single-relay NAF system with static channel gains. Both individual and global power constraints are considered. Specifically, by focusing on the per-node power constrained system, we first derive all locally optimal solutions of the mutual information. By comparing these solutions, it is then shown that the capacity is achieved by using either a DT scheme, a NAF beamforming (BF) protocol with a specific unit-rank covariance matrix, or a NAF protocol using a specific full-rank (FR) covariance matrix. For the system under the joint power constraint, it is shown following a similar approach that the capacity can be achieved by using either a DT
or a NAF-BF scheme. Thus, different from the fading scenario in [22], the OAF protocol is strictly suboptimal under both constraints and a diagonal covariance matrix is not always optimal. By further analyzing the asymptotic behavior in high and low transmission powers, it is demonstrated that NAF relaying is advantageous when the relay has large power and the source does not. However, the NAF protocol can provide significant gains over DT in medium transmission power regions as illustrated by different network configurations.

The rest of the chapter is organized as follows. Section 3.1 first introduces the problem formulation. The optimal covariance matrix at the source and power allocation scheme at the relay for individual and global power constraints are then derived in Sections 3.2 and 3.3, respectively. The behavior of the system in high and low transmission powers is discussed in Section 3.4. The capacities of the AF channel are illustrated in Section 3.5 for several network configurations. Finally, Section 3.6 concludes the chapter.

### 3.1 Problem Formulation

In this chapter, we consider the static scenario similar to [45]. In this scenario, the coherence time $T_c$ is assumed to be large enough so that the channel gains $\mathbf{h} = [h_0, h_1, h_2]$ can be considered as time-invariant constants for the entire duration of the transmitted codeword. For this static system, a good estimate of the channel gains can be obtained and hence it is assumed that $S$, $R$ and $D$ have full knowledge of $\mathbf{h}$. Since the relay has full CSI, the CI coefficient $b = \sqrt{z_2/[q_1\alpha_1 + N_r]}$ in (2.19) is adopted. The input-output relation of the considered NAF system was explained in Section 2.5 and is given by (2.16). It is important to note from Section 2.5 that the NAF protocol is general in the sense that it includes not only the OAF scheme in (2.23), but also the DT scheme in (2.21). Specifically, the OAF scheme corresponds to $q_2 = 0$, while the DT protocol is obtained by setting $z_2 = 0$. When $q_1, q_2, z_2 > 0$, one has the NAF protocol.

It can be shown from (2.16) that for a given transmit covariance matrix $\mathbf{Q} = \begin{pmatrix} q_1 & q_{12} \\ \hat{q}_{12} & q_2 \end{pmatrix}$ and power at the relay $z_2$, a Gaussian codebook must be used at the source to maximize the mutual information between the input and output vectors of the NAF channel [22, 45]. The mutual information between the input and output of the NAF channel is then given by (2.20). The objective of this chapter is to maximize the mutual information in (2.20) under both individual and global power constraints. First, for the per-node constraint scenario,
we assume that \( \text{tr}(Q) = q_1 + q_2 \leq q_s \) and \( z_2 \leq z_r \). With these constraints, the source and relay have respectively an average power constraint of \( q_s P_s \) and \( z_r P_r \) per cooperative frame, or \( q_s P_s/2 \) and \( z_r P_r/2 \) per symbol period. The constants \( q_s \) and \( z_r \) might be set to two to have constraints of \( P_s \) and \( P_r \) per symbol period. The per-node power-constrained capacity of the static NAF system can then be calculated as

\[
C_{\text{indiv}} = \max_{Q \succeq 0, z_2 \geq 0} I|\mathbf{h} \quad \text{s.t.} \quad \text{tr}(Q) \leq q_s, \quad z_2 \leq z_r.
\]  

(3.1)

To provide a fair comparison among different protocols and similar to [15, 61, 62, 64, 66–68], we also consider the global power constraint scenario. A network with such a constraint will provide an upper bound to any system with individual constraints. Furthermore, this constraint is a reasonable design criterion for static systems where the overall network power consumption is of interest. In this scenario, we assume that \( P_s = P_r = P_t \) and \( \text{tr}(Q) + z_2 = q_1 + q_2 + z_2 \leq q_t \). The system is then allowed to spend an average power of up to \( q_t P_t \) per cooperative frame or \( q_t P_t/2 \) per symbol period. The global power-constrained capacity of the static NAF system is then calculated as

\[
C_{\text{joint}} = \max_{Q \succeq 0, z_2 \geq 0} I|\mathbf{h} \quad \text{s.t.} \quad \text{tr}(Q) + z_2 \leq q_t.
\]  

(3.2)

Observe from (2.20) that, different from the static MIMO system in (2.13), the channel and noise covariance matrices are functions of the powers transmitted at the source and relay, i.e., \( \mathbf{H}_{\text{NAF}} = \mathbf{H}_{\text{NAF}}(q_1, z_2) \) and \( \mathbf{K} = \mathbf{K}(q_1, z_2) \). As a consequence, the problems in (3.1) and (3.2) are in general not concave with respect to \( z_2 \) and the psd matrix \( Q \), as shall be shown in subsequent sections. Hence, water-filling over the eigenvalues of \( \mathbf{H}_{\text{NAF}}^\dagger \mathbf{K}^{-1} \mathbf{H}_{\text{NAF}} \) is no longer optimal and the Karush-Kuhn-Tucker (KKT) conditions cannot be used to find the global maximizer. In the following, we first derive all local maximizers of the mutual information in (2.20). The capacities in (3.1) and (3.2) can then be calculated by comparing these local solutions.

### 3.2 Capacity under Individual Power Constraints

In this section, we derive the optimal covariance matrix at the source and optimal power allocation scheme at the relay to achieve the capacity with the per-node power constraints.
in (3.1). Let \( q_{12} = q e^{j\theta} \), where \( q \) and \( \theta \) denote the magnitude and phase of \( q_{12} \), respectively. To simplify the notation, let also \( \gamma_0 = (P_s \alpha_0)/N_d \), \( \gamma_1 = (P_s \alpha_1)/N_r \) and \( \gamma_2 = (P_r \alpha_2)/N_d \). In the remainder of the chapter, we assume that the values of \( \gamma_l \) are arbitrary but strictly positive, i.e., \(|h_l| = \sqrt{\alpha_l} > 0\).

Substituting the amplification coefficient \( b = \sqrt{z_2/(q_1 \alpha_1 + N_r)} \), the mutual information in (2.20) can be written as

\[
I(h_{12}) = \frac{1}{2} \log \det\left( I_2 + P_s H_{NAF}^\dagger K^{-1} H_{NAF} Q \right) = \frac{1}{2} \log \left[ f_C(q, q_{12}) \right], \tag{3.3}
\]

where \( q = [q_1, q_2, z_2] \) and

\[
f_C(q, q_{12}) = 1 + f_d(q) + f_o(q, q_{12}), \tag{3.4}
\]

with

\[
f_d(q) = \frac{q_1 \gamma_0 + q_2 \gamma_0 + q_1^2 q_2 \gamma_1 + q_1 q_2 \gamma_1^2 + q_1 q_2 \gamma_1 \gamma_0 + q_1 z_2 \gamma_2 \gamma_0 + q_1 z_2 \gamma_1 \gamma_2}{q_1 \gamma_1 + z_2 \gamma_2 + 1}, \tag{3.5}
\]

and

\[
f_o(q, q_{12}) = \frac{2q \sqrt{z_2 \gamma_0 \gamma_1 \gamma_2 (q_1 \gamma_1 + 1)} \cos(\theta_0 - \theta_1 - \theta_2 - \theta) - q^2 \gamma_0^2 (q_1 \gamma_1 + 1)}{q_1 \gamma_1 + z_2 \gamma_2 + 1}. \tag{3.6}
\]

Given that \( \log(\cdot) \) is a monotonically increasing function, the capacity can be obtained by maximizing \( f_C(q, q_{12}) \) in (3.4). For the system with per-node power constraints, the capacity in (3.1) can be obtained by solving the following optimization problem:

\[
\max_{q, q_{12}} f_C(q, q_{12}) \quad \text{s.t.} \quad \begin{cases} 
q, q_1, q_2, z_2 \geq 0, \\
q_1 + q_2 \leq q_s, \quad z_2 \leq z_r, \\
q^2 \leq q_1 q_2, 
\end{cases} \tag{3.7}
\]

where the last constraint guarantees that \( Q \) remains psd. The feasible region of (3.7) is illustrated in Fig. 3.1.

First, observe from (3.4) that \( f_C(\cdot) \) is an increasing function of \( q_2 \). At the optimal solution, the source must then use all its power, i.e., \( q_1 + q_2 = q_s \). Define the plane \( \mathcal{R}_1 = \{q \in \mathbb{R}^3 \mid q_k, z_2 \geq 0, \quad z_2 \leq z_r, \quad q_1 + q_2 = q_s\} \). The solution to (3.7) then lies on \( \mathcal{R}_1 \),
i.e., on the front surface of the feasible region in Fig. 3.1. Furthermore, it can be seen from (3.6) that $f_o(q, qe^{j\theta}) \leq f_o(q, qe^{j[\theta_0-\theta_1-\theta_2]})$ for any feasible $q$ and $q = |q_{12}|$. The equality is achieved when $\theta^* = \theta_0 - \theta_1 - \theta_2$, which becomes the optimal angle$^2$. Substituting this optimal angle, $f_o(\cdot)$ can be written as $f_o(q, qe^{j\theta^*}) = Aq^2 + Bq$ with $A$ and $B$ as in (3.6). Since $A < 0$, $f_o(q, qe^{j\theta^*})$ is a concave function of $q$. Given the concavity of $f_o(q, qe^{j\theta^*})$ in $q$ and that the maximum value of $q$ is $\sqrt{q_1 q_2}$ due to the psd constraint in (3.7), the optimal $q$ is either at the stationary point $q = -B/2A$ or at this maximum value, i.e.,

$$q^* = \min \left\{ \sqrt{\frac{z_2 \gamma_1 \gamma_2}{\gamma_0 (q_1 \gamma_1 + 1)}}, \sqrt{q_1 q_2} \right\}. \quad (3.8)$$

Then,

$$f_C(q, qe^{j\theta}) \leq f_C(q, q_{12}^*) = 1 + f_d(q) + f_o(q, q_{12}^*), \quad (3.9)$$

where the equality is achieved when $q_{12}^* = q^*e^{j\theta^*}$. For any given $q$, the optimal value of $q_{12}$ is thus $q_{12}^*$ with $\theta^* = \theta_0 - \theta_1 - \theta_2$ and $q^*$ as in (3.8).

Depending on whether $\sqrt{\frac{z_2 \gamma_1 \gamma_2}{\gamma_0 (q_1 \gamma_1 + 1)}}$ in (3.8) is smaller than $\sqrt{q_1 q_2}$, $R_1$ can

Note that for the general $2 \times 2$ static MIMO channel with components $h_{m,n} = \sqrt{\alpha_{m,n}} e^{j\theta_{m,n}}$ ($1 \leq m, n \leq 2$), the optimal angle is $\theta^* = \theta_{2,2} - \theta_{2,1}$ when $\alpha_{1,2} = 0$ as in (2.17).
Proof. See Appendix A.1.

From Lemma 3.1, neither $f_C(q, q_{12}^*)$ nor consequently $f_C(q, q_{12})$ in (3.4) are quasiconcave in $\mathcal{R}_1$ [73, Ch.3.4]. Thus, $f_C(q, q_{12})$ cannot be log-concave, i.e., $\log[f_C(q, q_{12})]$ is not concave [73]. As such, (3.1) and (3.7) are non-concave optimization problems. Given that $f_C(q, q_{12})$ does not have any local maximizers in the interior of $\mathcal{R}_1$, the solution to (3.7) must lie on the boundary. As illustrated in Fig. 3.1, let the line segments that form the perimeter of the rectangle $\mathcal{R}_1$ be defined as $\ell_1 = \{q \in \mathbb{R}^3 \mid q_1 = 0, q_2 = q_s, 0 \leq z_2 \leq z_r \}$, $\ell_{II} = \{q \in \mathbb{R}^3 \mid z_2 = 0, q_1 + q_2 = q_s, q_1, q_2 \geq 0 \}$, $\ell_{III} = \{q \in \mathbb{R}^3 \mid q_2 = 0, q_1 = q_s, 0 \leq z_2 \leq z_r \}$, and $\ell_{IV} = \{q \in \mathbb{R}^3 \mid z_2 = z_r, q_1 + q_2 = q_s, q_1, q_2 \geq 0 \}$. Note that $\ell_{II}$, $\ell_{III}$ and $\ell_{IV}$ correspond to the DT, OAF, and NAF protocols, respectively, while $\ell_1$ is the trivial case in which the source transmits only in the second time slot and the relay amplifies noise. We first analyze the local solutions in $\ell_1$, $\ell_{II}$ and $\ell_{III}$.

**Lemma 3.2.** The maximizer in $\ell_{II}$, denoted as $C_{DT}$, is achieved by the DT scheme with
equal power allocation:

\[ q_{1,DT} = q_{2,DT} = q_s/2, \quad z_{2,DT} = q_{12,DT} = 0, \] (3.13)

and outperforms the maximizer in \( \ell_1 \). The maximizer in \( \ell_{III} \), denoted as \( C_{OAF} \), is achieved by the OAF scheme with full power at both nodes:

\[ q_{1,OAF} = q_s, \quad z_{2,OAF} = z_r, \quad q_{2,OAF} = q_{12,OAF} = 0. \] (3.14)

Proof. When \( q \in \ell_1 \subset \mathcal{P}_2 \), \( q_1 = 0 \) and \( q^* = 0 \) from (3.8). Since \( f_C([0, q_s, z_2], 0) = f_{P_2}([0, q_s, z_2]) \) is a decreasing function of \( z_2 \), it is maximized when \( z_2 = 0 \). When \( q \in \ell_{II} \subset \mathcal{P}_1 \), \( z_2 = 0 \) and \( q^* = 0 \) from (3.8). The function \( f_C([q_1, q_2, 0], 0) = f_{P_1}([q_1, q_2, 0]) \) is maximized when \( q_1 = q_2 = q_s/2 \) from the geometric/arithmetic mean inequality. Furthermore, \( f_C([q_s/2, q_s/2, 0], 0) > f_C([0, q_s, 0], 0) \) for any \( \gamma > 0 \). Thus, the maximizer in \( \ell_{II} \) is greater than the one in \( \ell_1 \). When \( q \in \ell_{III} \subset \mathcal{P}_2 \), \( q_2 = 0 \) and \( q^* = 0 \) from (3.8). Since \( f_C([q_s, 0, z_2], 0) = f_{P_2}([q_s, 0, z_2]) \) is increasing with \( z_2 \), it is maximized when \( z_2 = z_r \).

We now analyze the maximizer in the remaining segment \( \ell_{IV} \).

**Lemma 3.3.** The maximizer in \( \ell_{IV} \), denoted as \( C_{NAF} \), is given by

\[ C_{NAF} = \max\{C_{BF}, C_{FR}\}, \] (3.15)

and outperforms the maximizer in \( \ell_{III} \) since \( C_{BF} > C_{OAF} \). In (3.15), \( C_{BF} \) is achieved by

\[ q_{1,BF} = r_6, \quad q_{2,BF} = q_s - r_6, \quad z_{2,BF} = z_r, \quad q_{12,BF} = \sqrt{q_{1,BF}q_{2,BF}}e^{j\theta^*}, \] (3.16)

where \( r_6 \) is the only root in \( 0 < r_5 < r_6 < q_s \) of the 6th order polynomial \( P_6(q_1) \) in (A.4), and \( r_5 \) is the unique positive root of the cubic polynomial \( P_5(q_1) \) in (A.3); whereas \( C_{FR} \) is achieved by

\[ q_{1,FR} = r_3, \quad q_{2,FR} = q_s - r_3, \quad z_{2,FR} = z_r, \quad q_{12,FR} = \sqrt{\frac{z_{2,FR}\gamma_1\gamma_2}{\gamma_0(q_{1,FR}\gamma_1 + 1)}}e^{j\theta^*}, \] (3.17)

as long as \( P(r_3) - z_r > 0 \), where \( P(q_1) \) is the cubic polynomial in (3.10) and \( r_3 \) is the largest
real root of the cubic polynomial $P_3(q_1)$ in (A.6).

Proof. When $q \in \ell_{IV}$, we must consider two subcases depending on whether $P(q_1)$ is smaller or greater than $z_r$ for $0 \leq q_1 \leq q_s$. Denote the line subsegments $\ell_{IV}^1(z_r) = \{q \in \ell_{IV} \cap \mathcal{P}_1\} = \{q \in \ell_{IV} \mid z_r \leq P(q_1)\}$ and $\ell_{IV}^2(z_r) = \{q \in \ell_{IV} \cap \mathcal{P}_2\} = \{q \in \ell_{IV} \mid z_r \geq P(q_1)\}$. When $P(q_1) < z_r$, $\ell_{IV}$ is completely in $\mathcal{P}_2$ as shown in the left side of Fig. 3.2, i.e., $\ell_{IV} = \ell_{IV}^2(z_r)$ and $\ell_{IV}^1(z_r) = \emptyset$. On the other hand, when $P(q_1) \geq z_r$, the two extremes of $\ell_{IV}$ are in $\mathcal{P}_2$ while the mid-section is in $\mathcal{P}_1$ as shown in the right side of Fig. 3.2. Note that in this case $\ell_{IV}^1(z_r)$ is always closed and convex as $P(q_1)$ is quasiconcave and never includes the end-points $q_1 = 0$ and $q_1 = q_s$, i.e., $\ell_{IV}^1(z_r)$ is never empty. The solutions for these two subcases are derived in Appendix A.2.

![Fig. 3.2 Feasible region when $P(q_1) < z_r$ (left) and $P(q_1) > z_r$ (right) for $0 \leq q_1 \leq q_s$.](image)

First, observe from Lemma 3.3 that $\ell_{IV}$ has at most two maximizers. Analogous to MIMO systems [74, Ch.2.3], we have denoted the solution in (3.16) as the NAF beamforming (NAF-BF) scheme. This is because the rank of $Q$ is unity when $|q_{12,BF}| = \sqrt{q_{1,BF}q_{2,BF}}$ ($\det(Q) = 0$), which corresponds to the case where the source transmits the same information symbol in both time slots, i.e., the transmitted vector in (2.16) is $x = wx$ where $w$ is a $2 \times 1$ beamforming vector and $x$ is the information bearing signal. However, different from MIMO systems, the beamforming in (2.16) is applied over space and time rather than over space only. On the other hand, the solution in (3.17) corresponds to the NAF protocol with a full-rank covariance matrix, NAF-FR. Although the NAF-BF maximizer in (3.16) is always feasible, note from Lemma 3.3 that the NAF-FR one in (3.17) might not depending on the channel gains, i.e., when $P(r_3) - z_r < 0$. More importantly, it can be implied from Lemma 3.3 that the maximizer in $\ell_{III}$ cannot be the solution to (3.7), i.e., the OAF protocol is not optimal under individual constraints. This is because the NAF-BF
maximizer is always feasible and outperforms the OAF one. Finally, it is also important to note that different from static MIMO systems [45], the NAF solutions in (3.16) and (3.17) do not diagonalize $H_{NAF}^H K^{-1} H_{NAF}$ in (3.3).

From the above lemmas, the maximizer of (3.7) and hence the capacity in (3.1) is finally given in the following theorem.

**Theorem 3.1.** The capacity of the system under individual power constraints $C_{\text{ indiv}}$ in (3.1) is given by

$$C_{\text{ indiv}} = \max\{C_{\text{ DT}}, C_{\text{ BF}}, C_{\text{ FR}}\}.$$  \hspace{1cm} (3.18)

**Proof.** This follows directly from Lemmas 3.1–3.3. \hfill \Box

Although the capacity in Theorem 3.1 is expressed as the maximum rate among these three schemes (due to the non-concavity of the problem), the solution in (3.18) provides several important insights. For instance, among all possible transmission protocols, the capacity in (3.18) can only be achieved by the DT scheme with equal power allocation as in (3.13), the NAF-BF scheme with the power allocation in (3.16), or the NAF-FR scheme with a full-rank covariance matrix as in (3.17). The choice of protocol depends on the power available at the source and relay, as well as the network configuration. It can also be seen from (3.18) that at the optimal solution, the relay uses either all its available power or no power at all. This is in contrast to the behavior of AF systems over fading channels in [22], where it was shown that the relay power can take on any value between zero and $z_r$. However, similar to [22], using full power at the relay is in fact not beneficial under some channel conditions and DT is optimal. This is despite the fact that the overall power consumption is reduced. Observe also from (3.18) that different from the fading channel in [22], a diagonal covariance matrix is only optimal for the DT protocol. Finally, note from Theorem 3.1 that although the OAF scheme cannot be the solution to (3.18), it can outperform the DT scheme in some scenarios, as will be shown in Section 3.5.

Further discussions on the capacity in (3.18) regarding different transmission power regimes and network configurations shall be discussed in subsequent sections.
3.3 Capacity under Joint Power Constraints

In the previous section, we consider the system under individual power constraints. We now analyze the capacity under a joint power constraint. For the system with a global power constraint, the capacity in (3.2) can be obtained by solving the following optimization problem:

$$\max_{q, q_{12}} \quad f_C(q, q_{12}) \quad \text{s.t.} \quad \begin{cases} q, q_1, q_2, z_2 \geq 0, \\ q_1 + q_2 + z_2 \leq q_t, \\ q^2 \leq q_1q_2, \end{cases}$$

with $f_C(q, q_{12})$ as in (3.4). The feasible region of (3.19) is illustrated in Fig. 3.3.

First, at the optimal solution, it is easy to show that the system must use all its power $q_1 + q_2 + z_2 = q_t$. This follows again from the fact that $f_C(\cdot)$ is an increasing function of $q_2$. The solution to (3.7) then must lie on the front plane in Fig. 3.3, i.e., on the plane $\mathcal{R}_2 = \{q \in \mathbb{R}^3 | 0 \leq q_k, z_2 \leq q_t, q_1 + q_2 + z_2 = q_t\}$. Define the line segments that form the perimeter of the triangle $\mathcal{R}_2$ as $\ell_I = \{q \in \mathbb{R}^3 | q_1 = 0, q_2 + z_2 = q_t, q_2, z_2 \geq 0\}$, $\ell_{II} = \{q \in \mathbb{R}^3 | z_2 = 0, q_1 + q_2 = q_t, q_1, q_2 \geq 0\}$, and $\ell_{III} = \{q \in \mathbb{R}^3 | q_2 = 0, q_1 + z_2 = q_t, q_1, z_2 \geq 0\}$. As before, $\ell_{II}$, $\ell_{III}$, and $\ell_I$ represent the DT, OAF and trivial protocols, respectively, while the interior of $\mathcal{R}_2$ now corresponds to the NAF scheme.

Similar to (3.9) in the previous section, $f_C(q, q_{12}) \leq f_C(q, q^*_{12})$ for any feasible $q$, where the equality is achieved when $q^*_{12} = q^* e^{i\theta^*}$ with $q^*$ as in (3.8) and $\theta^* = \theta_0 - \theta_1 - \theta_2$. As
shown in Fig. 3.3, $\mathcal{R}_2$ can then be divided into two regions as

$$\mathcal{P}_1 = \{ q \in \mathcal{R}_2 \mid q_2 \geq P(q_1) \}, \quad \text{and} \quad \mathcal{P}_2 = \{ q \in \mathcal{R}_2 \mid q_2 \leq P(q_1) \}, \quad (3.20)$$

where now

$$P(q_1) = \frac{\gamma_1 \gamma_2 (q_t - q_1)}{\gamma_0 q_1 \gamma_1 + 1 + \gamma_1 \gamma_2}. \quad (3.21)$$

Note that $P(q_1)$ is a decreasing function of $q_1$ with $P(0) = q_t$ and $P(q_t) = 0$. Hence, $f_C(q, q_1)$ is again given as in (3.12) but now using the regions defined by (3.20) and (3.21). We first analyze $f_C(q, q_1)$ the interior of $\mathcal{R}_2$.

**Lemma 3.4.** $f_C(q, q_1)$ is not quasiconcave in $\mathcal{R}_2$ and has a single local maximizer in the interior of $\mathcal{R}_2$, denoted as $C_{BF}$, that is achieved by

$$q_{1,BF} = \begin{cases} \frac{q_t/2}{-q_1 + \sqrt{q_1 q_2 + 1}}, & \gamma_1 = \gamma_2 \\ \frac{-q_2 + \sqrt{q_1 q_2 + 1}}{\gamma_1 - \gamma_2}, & \gamma_1 \neq \gamma_2, \end{cases}$$

$$q_{2,BF}(q_{1,BF}) = \frac{\gamma_0 (q_t - q_{1,BF}) [q_{1,BF} \gamma_1 + q_{1,BF} \gamma_2 + 1]^2}{\gamma_0 [q_{1,BF} \gamma_1 + (q_t - q_{1,BF}) \gamma_1 + 1]^2 + q_{1,BF} \gamma_1 q_{1,BF} \gamma_2 (q_{1,BF} \gamma_1 + 1)},$$

$$z_{2,BF} = q_t - q_{1,BF} - q_{2,BF}, \quad q_{12,BF} = \sqrt{q_{1,BF} q_{2,BF} e^{j\theta^*}}. \quad (3.22)$$

**Proof.** See Appendix A.3. \qed

Similar to the per-node power constraint scenario in Lemma 3.1, observe from Lemma 3.4 that neither (3.19) nor (3.2) are concave problems [73]. Note also that the local solution in (3.22) corresponds to the NAF beamforming scheme and is always feasible. As before, the NAF-BF solution in (3.22) does not diagonalize $H_{N_{AF}}^\dagger K^{-1} H_{N_{AF}}$. We now analyze the perimeter of $\mathcal{R}_2$.

**Lemma 3.5.** The maximizer in $\ell_H$, $C_{DT}$, is achieved by

$$q_{1,DT} = q_{2,DT} = q_t/2, \quad z_{2,DT} = q_{12,DT} = 0, \quad (3.23)$$
and outperforms the maximizer in \( \ell_1 \). The maximizer in \( \ell_{III} \), \( C_{OAF} \), is achieved by

\[
q_{1,OAF} = \begin{cases}
q_t, & b_1 \geq 0 \\
\min \left\{ q_t, -c_1/b_1 \right\}, & \gamma_1 = \gamma_2 \\
\min \left\{ q_t, -b_1 - \sqrt{b_1^2 - 4a_1c_1} \right\}, & b_1 < 0, \gamma_1 \neq \gamma_2
\end{cases}
\]

\[z_{2,OAF} = q_t - q_{1,OAF}; \quad q_{2,OAF} = q_{12,OAF} = 0, \]

(3.24)

where \( a_1 = (\gamma_1 - \gamma_2)(\gamma_0\gamma_1 - \gamma_0\gamma_2 - \gamma_1\gamma_2), \quad b_1 = 2(q_t\gamma_2 + 1)(\gamma_0\gamma_1 - \gamma_0\gamma_2 - \gamma_1\gamma_2) \) and \( c_1 = (q_t\gamma_2 + 1)(\gamma_0 + q_t\gamma_0\gamma_2 + q_t\gamma_1\gamma_2) \).

Proof. The proof for \( \ell_1 \) and \( \ell_{II} \) follows from Lemma 3.2 by replacing \( q_s \) by \( q_t \). The proof for \( \ell_{III} \) is given in [61, 67].

The maximizer of (3.19) and thus the capacity in (3.2) are finally given in the following theorem.

**Theorem 3.2.** The capacity of the system under joint power constraints \( C_{joint} \) in (3.2) is given by

\[
C_{joint} = \max \{ C_{DT}, C_{BF} \}. \tag{3.25}
\]

Proof. This follows from Lemmas 3.4–3.5, and the fact that \( C_{BF} > C_{OAF} \) (see Appendix A.4).
Although (3.25) is again given as the maximum rate between two schemes, the capacity can be established in closed-form for some network topologies, as shall be shown shortly.

3.4 High and Low Power Regions

In the previous section, the capacity of the static AF system was derived and expressions for the optimal covariance matrix at the source and power allocation at the relay were provided. In this section, we provide further insights on the behavior of the capacities in (3.18) and (3.25) along with their respective power allocations by focusing on two asymptotic regions, namely the low and high transmission powers. Although the OAF protocol has been shown to be suboptimal under individual and joint power constraints, its capacity will also be discussed here due to its popularity in the literature.

3.4.1 Per-node Power Constraints

As shown in Theorem 3.1, the capacity for the individual power constraint scenario is achieved by either the DT, the NAF-FR or the NAF-BF protocol. Hence, one must evaluate and compare $C_{\text{DT}}$, $C_{\text{BF}}$ and $C_{\text{FR}}$ in (3.18). Consider the following asymptotic cases.

High Source Power, Fixed Relay Power

In this case, we assume that $P_s \to \infty$ while $P_r$ remains fixed. Substituting (3.13) in (3.3), the capacity of the DT scheme is given as $C_{\text{DT}} = \log(1 + [q_s \gamma_0]/2) = \log(P_s) + O(1)$, which grows logarithmically with $P_s$ as already well-known. For the NAF-FR scheme, note from (A.6) that the coefficients of $P_3(\cdot)$, $B_3, C_3, D_3$, become positive for large values of $P_s$. Since $A_3 < 0$, $P_3(\cdot)$ has then a single positive root $r_3$. Furthermore, applying the method of dominant balance [75, Ch.3.4], it can be shown that $r_3 = O(1)$. $P_3(\cdot)$ can then be approximated as $P_3(q_1) = A_3q_1^3 + B_3q_2 + O(P_s^4)$ and its root $r_3 \to q_s/2$. It can be easily shown that $P(q_s/2) - z > 0$ for large $P_s$ and thus $q_{\text{FR}} \to q_s/2$ from Lemma 3.3. Substituting (3.17) in (3.4) with $q_{\text{FR}} = q_s/2$, $f_C(q_{\text{FR}}, q_{12, \text{FR}}) \approx ([q_s \gamma_0]/2)^2$ and the capacity of the NAF-FR scheme $C_{\text{FR}} \approx \log([q_s \gamma_0]/2) \approx \log(P_s)$ also grows logarithmically with $P_s$. However, by comparing $(1 + [q_s \gamma_0]/2)^2$ to (3.4) with $q_{\text{FR}} = q_s/2$ in (3.17), it can be easily shown that $C_{\text{DT}} > C_{\text{FR}}$ for sufficiently large $P_s$. The NAF-FR system then approaches the DT scheme from below. Moreover, it can be seen from Appendix A.2 that the NAF-FR
technique outperforms the NAF-BF one when $B_3$, $C_3$, and $D_3$ are positive. Specifically, following a similar approach as above, it can be shown that $r_6$ in (3.16) grows as $O(1)$ and $C_{\text{BF}} \approx [1/2] \log(P_s)$, i.e., the capacity of the NAF-BF scheme only increases logarithmically with $\sqrt{P_s}$. This is due to the fact that a single information symbol is sent in two time slots. Due to the same reason, it is easy to show that $C_{\text{OAF}} \approx [1/2] \log(P_s)$, i.e., the capacity of the NAF-BF scheme only increases logarithmically with $\sqrt{P_s}$. This means that the capacity of the NAF-BF protocol approaches that of the NAF-BF protocol from below, i.e., $C_{\text{OAF}} \rightarrow C_{\text{BF}}$. Substituting (3.14) in (3.3), the capacity of the OAF system can then be approximated as $C_{\text{OAF}} \approx [1/2] \log(1+q_s\gamma_0)$, which outperforms $C_{\text{DT}} = \log(1+q_s\gamma_0)/2$ as long as $4\gamma_1 > q_s\gamma_0^2$. Hence, for large relay powers, $C_{\text{OAF}} \rightarrow C_{\text{BF}}$ and $C_{\text{BF}} > C_{\text{OAF}} > C_{\text{DT}}$ when $q_s\gamma_0^2 < 4\gamma_1$, or $C_{\text{DT}} > C_{\text{BF}} > C_{\text{OAF}}$ when $q_s\gamma_0^2 > 4\gamma_1$. The capacity in (3.18) is then achieved by either the NAF-BF or the DT scheme depending on the values of $q_s$, $\gamma_0$ and $\gamma_1$. In this case, relaying is useful for low source powers, i.e., $q_s\gamma_0^2 < 4\gamma_1$ when $P_s \rightarrow 0$.

High Source and Relay Power

We now consider the scenario in which both $P_s, P_r \rightarrow \infty$ but their ratio $P_r/P_s$ remains constant. As in first case, it can be easily shown that $r_3$ in (3.17) and $r_6$ in (3.16) grow as $O(1)$. Hence, $C_{\text{FR}} \approx \log(P_s)$ and $C_{\text{BF}} \approx [1/2] \log(P_s)$. Similarly, it can be shown that $C_{\text{DT}} \approx \log(P_s)$ and $C_{\text{OAF}} \approx [1/2] \log(P_s)$. By comparing the function values in (3.4) with (3.17) and (3.13), it can further be shown that $C_{\text{DT}} > C_{\text{FR}}$. Therefore, $C_{\text{DT}} > C_{\text{FR}} > C_{\text{BF}} > C_{\text{OAF}}$. Thus, as in the first case, DT is optimal when both the source and relay have high powers. A similar observation was made in [22] for the AF fading channel.
Low Source and Relay Power

We finally consider the scenario in which $P_s, P_r \to 0$ with $P_r/P_s$ fixed. For the DT scheme, $C_{DT} = \log(1 + [q_s\gamma_0]/2) \approx [q_s\gamma_0]/[2\ln(2)]$, where we have used the fact that $\ln(1 + x) \approx x$ for small values of $x$. The capacity of the DT scheme then decreases linearly with $P_s$ as already-known. For the NAF-FR protocol, the maximizer in (3.17) is not feasible as $P_s(\gamma_1) - z_r < 0$ when $P_s, P_r \to 0$. For the NAF-BF scheme, $P_6(q_1) \approx \gamma_0\gamma_1(2q_1 - q_s) + O(P_s^3)$ and $q_{1,BF} = r_6 \to q_s/2$. Substituting (3.16) in (3.4) with $q_{1,BF} = q_s/2$, $C_{BF} \approx [1/2]\log(1 + q_s\gamma_0) \approx [q_s\gamma_0]/[2\ln(2)]$ also decreases linearly with $P_s$. However, by comparing the function values, it can be easily shown that $C_{BF} > C_{DT}$. In addition, $C_{OAF} \approx [1/2]\log(1 + q_s\gamma_0) \approx [q_s\gamma_0]/[2\ln(2)]$ for low power regions. Therefore, $C_{BF} > C_{DT}$ for low source and relay powers and the capacity in (3.18) is achieved by the NAF-BF scheme. Although the NAF-BF scheme is optimal, it does not provide great advantages in this scenario as $C_{BF} \approx C_{OAF} \approx C_{DT}$.

3.4.2 Global Power Constraint

In the joint power constraint scenario, it can be seen from Theorem 3.2 that one must compare $C_{DT}$ and $C_{BF}$ in (3.25). We again include the OAF protocol in this comparison. Consider the following two cases.

High Global Power

In this case, we assume that $P_t \to \infty$. Substituting (3.23) in (3.3), the capacity of the DT scheme is given as $C_{DT} = \log(1 + [q_t\gamma_0]/2) \approx \log(P_t)$, which grows logarithmically with $P_t$. For the NAF-BF scheme, it can be shown from (3.22) that $q_{1,BF}$ and $q_{2,BF}$ grow as $O(1)$ and hence $C_{BF} \approx [1/2]\log(P_t)$. This is again due to the fact that only one information symbol is sent in each cooperative frame. Likewise, it is easy to show that $q_{1,OAF}$ in (3.24) grow as $O(1)$ for any network configuration and $C_{OAF} \approx [1/2]\log(q_t)$. Since $C_{BF} > C_{OAF}$ from Theorem 3.2, $C_{DT} > C_{BF} > C_{OAF}$ for high values of $q_t$ and the DT scheme achieves the capacity in (3.25).
Low Global Power

We then consider the case with $P_t \to 0$. In low transmission powers, $C_{DT} = \log(1 + [q_t \gamma_0]/[2 \ln(2)])$. For the NAF-BF scheme, it can be shown from (3.22) that $q_{1,BF}, q_{2,BF} \to q_t/2$. Substituting (3.22) in (3.3) with $q_{1,BF} = q_{2,BF} = q_t/2$, the capacity of the NAF-BF scheme can be approximated as $C_{BF} \approx \log(1 + q_t \gamma_0) \approx [q_t \gamma_0]/[2 \ln(2)]$. For the OAF scheme, it can be shown that $q_{1,OAF} = q_t$ as $-c_1/b_1$ and $[-b_1 - \sqrt{b_1^2 - 4a_1 c_1}]/[2a_1]$ in (3.24) approach infinity for low values of $P_t$. Substituting (3.24) in (3.3) with $q_{1,OAF} = q_t$, $C_{OAF} = [1/2] \log(1 + q_t \gamma_0) \approx [q_t \gamma_0]/[2 \ln(2)]$. Given that $\log(1+[q_t \gamma_0]/2) > [1/2] \log(1+q_t \gamma_0)$, $C_{DT} > C_{BF}$. Therefore, $C_{DT} > C_{BF} > C_{OAF}$ for low values of $q_t$ and the capacity in (3.25) is again achieved by the DT scheme. The gain provided by DT is however not significant as $C_{DT} \approx C_{BF} \approx C_{OAF}$ in low power regions.

The advantages of the considered relaying strategies subject to different power constraints are summarized in Table 3.1. Under individual power constraints, AF relaying is advantageous only when the relay is allocated a large power and the source is not. In addition, it can be seen that under a global constraint, AF relaying does not provide any advantage over DT in the considered asymptotic regions. However, as shall be shown in the following section, AF relaying is useful in medium power ranges and might provide significant gains.

### Table 3.1  Capacity achieving protocols in high and low power regions.

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Scenario</th>
<th>Capacity achieving protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiv.</td>
<td>High $P_s$, fixed $P_r$</td>
<td>DT</td>
</tr>
<tr>
<td></td>
<td>High $P_r$, fixed $P_s$</td>
<td>NAF-BF ($q_s \gamma_0^2 &lt; 4\gamma_1$) or DT ($q_s \gamma_0^2 &gt; 4\gamma_1$)</td>
</tr>
<tr>
<td></td>
<td>High $P_s, P_r$</td>
<td>DT</td>
</tr>
<tr>
<td></td>
<td>Low $P_s, P_r$</td>
<td>NAF-BF $\approx$ DT $\approx$ OAF</td>
</tr>
<tr>
<td>Joint</td>
<td>High $P_t$</td>
<td>DT</td>
</tr>
<tr>
<td></td>
<td>Low $P_t$</td>
<td>DT $\approx$ NAF-BF $\approx$ OAF</td>
</tr>
</tbody>
</table>

### 3.5 Illustrative Examples

Based on analysis in Sections 3.2, 3.3 and 3.4, this section illustrates the capacity and optimal power allocation for some specific network configurations. In particular, we shall consider two network models, namely the symmetric and the linear network model. Without
loss of generality, we set \(q_s = 2\), \(z_r = 2\), and \(q_t = 2\). The per-node average power constraints at the source and relay are then \(P_s\) and \(P_r\) per symbol period, respectively, whereas \(P_t\) is the joint average power constraint, also per symbol period. Furthermore, we assume that the noise variances at both receiving nodes are unity, i.e., \(N_d = N_r = 1\).

### 3.5.1 Symmetric Network Model

Consider the symmetric network configuration where the source, relay and destination are all equidistant from each other. In this configuration, the nodes form an equilateral triangle and thus \(\alpha_0 = \alpha_1 = \alpha_2 = \alpha\) for some \(\alpha > 0\). Fig. 3.4 shows the achievable rates of various AF protocols for the individual power constraint scenario with \(\alpha = 1\). In particular, we consider the rates of the DT, NAF-FR, NAF-BF and OAF schemes using the power allocations in (3.13), (3.17), (3.16), and (3.14), respectively. For comparison purposes, we also consider the rate of the NAF system using a diagonal covariance matrix \(q = 0\) and with full power at source and relay \(q = [q_s/2, q_s/2, z_r]\), denoted as the NAF-D protocol. In Fig. 3.4, the rates are normalized by the rate of the DT scheme (3.13), and are plotted against source power \(P_s\) (in dB) for a fixed relay power \(P_r = 0\) dB. First, observe from Fig. 3.4 that the NAF-BF scheme is optimal for \(P_s < 0.2\) dB, the NAF-FR scheme is optimal for \(0.3\) dB < \(P_s\) < 1.1 dB, and the DT scheme is optimal for \(P_s > 1.2\) dB. This is in agreement with the capacity of the per-node power constrained system in (3.18), shown with square markers in Fig. 3.4 and all subsequent figures. As discussed in Section 3.4.1, it can be seen from Fig. 3.4 that AF relaying is not useful in high source power ranges as \(C_{FR}/C_{DT} \rightarrow 1\) from below, \(C_{BF}/C_{DT} \rightarrow 1/2\) and \(C_{OAF}/C_{DT} \rightarrow 1/2\).

Fig. 3.5 shows the rates of the considered AF protocols against relay power \(P_r\) (in dB) for a fixed \(P_s\) and still for the individual power constraint scenario. Two source powers are considered in Fig. 3.5, \(P_s = 0\) dB and \(P_s = 4.77\) dB. The rates in Fig. 3.5 are normalized by the rate of the DT scheme with \(P_s = 0\) dB, i.e., \(C_{DT} = 1\). Note from Fig. 3.5 that for \(P_s = 0\) dB, the capacity in (3.18) is achieved by the NAF-FR scheme when \(P_r < -0.6\) dB and by the NAF-BF protocol when \(P_r > -0.5\) dB. On the other hand, the capacity for \(P_s = 4.77\) dB is achieved by the DT scheme in the entire range of \(P_r\) in Fig. 3.5. In both cases, the rate achieved by the OAF scheme approaches that of the NAF-BF protocol for high relay powers. This is in agreement with the large relay power analysis in Section 3.4.1 as \(C_{BF} > C_{DT}\) for \(P_s = 0\) dB (since \(q_s\gamma_0^2 = 2P_s^2\alpha_0^2 = 2 < 4\gamma_1 = 4P_s\alpha_1 = 4\) and \(C_{DT} > C_{BF}\) in this case).
for $P_s = 4.77$ dB (since $q_s \gamma_0^2 = 18 > 4\gamma_1 = 12$). It is also important to note from Figs. 3.4 and 3.5 that under certain configurations, using the relay might in fact reduce the achievable rate of the system.

The achievable rates over the symmetric network model are shown in Fig. 3.6 for the global constraint scenario. In this scenario, we consider the DT, NAF-BF, and OAF protocols with the power allocations in (3.23), (3.22), and (3.24), respectively. We also consider the OAF scheme with equal power allocation ($q = 0$ and $q = [q_t/2, 0, q_t/2]$), denoted as OAF-EQ, and the NAF system using a diagonal covariance and equal power allocation ($q = 0$ and $q = [q_t/3, q_t/3, q_t/3]$), denoted as NAF-EQ. As before, the rates in Fig. 3.6 are normalized by the rate of the DT scheme (3.23) and are plotted against the total transmitted power $P_t$ (in dB). Observe from Fig. 3.6 that for high transmission powers, $C_{BF}/C_{DT} \rightarrow 1/2$ and $C_{OAF}/C_{DT} \rightarrow 1/2$, which agrees with Section 3.4.2. Note also from Fig. 3.6 and Section 3.4.2 that $C_{DT} \approx C_{BF} \approx C_{OAF}$ in low power regions. More importantly, the DT scheme can be seen to be optimal for the entire range of $P_t$ in Fig. 3.6. This is because over the symmetric network configuration, it can be easily shown that $C_{DT} > C_{BF}$ regardless of the value of $\alpha$. 

**Fig. 3.4** Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against $P_s$ ($P_r = 0$ dB).
Fig. 3.5 Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against $P_r$ ($P_s = 0, 4.77$ dB).

Fig. 3.6 Normalized rates of the DT, NAF-BF, OAF, OAF-EQ and NAF-EQ protocols against $P_t$. 
3.5.2 Linear Network Model

We now consider the more practical linear network model which captures path-loss across transmission links. In this configuration, it is assumed that the relay is placed in the line between the source and the destination, the S-D distance is normalized to one, while the S-R and R-D distances are given by $d$ and $(1-d)$, respectively. As a result, $\alpha_0 = 1$, $\alpha_1 = 1/d^\nu$ and $\alpha_2 = 1/(1-d)^\nu$, where $\nu$ is the path-loss exponent, which is set to 3 in this section for convenience. Fig. 3.7 shows the normalized rates of the per-node constrained systems for $d = 0.5$ and $P_s = P_r$. First, note from Fig. 3.7 that the capacity in Theorem 3.1 is achieved by the NAF-BF scheme for $P_s < 6.3$ dB, by the NAF-FR scheme for $6.4 \text{ dB} < P_s < 10.1$ dB, and by the DT scheme for $P_s > 10.2$ dB. As explained in Section 3.4.1, it can be seen from Fig. 3.7 that $C_{\text{FR}}/C_{\text{DT}} \to 1$ and $C_{\text{BF}}/C_{\text{DT}}, C_{\text{OAF}}/C_{\text{DT}} \to 1/2$ when $P_s, P_r \to \infty$, whereas $C_{\text{BF}}/C_{\text{DT}}, C_{\text{OAF}}/C_{\text{DT}} \to 1$ when $P_s, P_r \to 0$, i.e., AF relaying does not provide significant gains in low and high power regions. However, in medium power ranges, the NAF-BF system can be seen to provide impressive gains over the DT scheme. Specifically, the rate of the NAF-BF protocol in Fig. 3.7 is 3.5 times that of the DT scheme at $P_s = -10$ dB.

![Fig. 3.7 Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against $P_s = P_r$ ($d = 0.5$).](image-url)
Fig. 3.8 shows the rates of the individual power constrained systems against $P_r$ for fixed $P_s = 0$ dB. Three values of $d$ are considered: $d = 0.1, 0.5$ and $0.9$. Note that the rate of the DT scheme is equal for these values as $\alpha_0 = 1$. Observe from Fig. 3.8 that the rate of the OAF system approaches that of the NAF-BF scheme for high relay powers, as expected. Furthermore, $C_{BF} > C_{DT}$ for all values of $d$ in Fig. 3.8 since $q_s \gamma_0^2 < 4 \gamma_1$. More importantly, the asymptotic gain offered by the NAF-BF scheme over the DT system increases as the relay gets closer to the source. Specifically, the NAF-BF system presents asymptotic gains of 1.2, 2.1 and 5.4 times the rate of the DT protocol for $d = 0.9, 0.5$ and 0.1, respectively.

![Fig. 3.8](image)

**Fig. 3.8** Normalized rates of the DT, NAF-FR, NAF-BF, OAF and NAF-D protocols against $P_r$ ($P_s = 0$ dB, $d = 0.1, 0.5, 0.9$).

The rates of the global power constrained systems are shown in Fig. 3.9 for $d = 0.6$. Note that the asymptotic behavior described in Section 3.4.2 holds for all systems in Fig. 3.9. Although AF relaying does not provide any advantage over DT for high and low transmission powers, it can be seen from Fig. 3.9 that significant gains can be achieved in medium power levels. For instance, an almost two-fold increase in rate over the DT protocol is attained by the NAF-BF system at $P_t = -10$ dB.

Overall, it can be seen from the results in this section that AF relaying might provide significant rate improvements under the right power regime and network topology, despite
the half-duplex constraint. It can also be seen that the capacity of the per-node and joint power constrained system is only achieved by the protocols described in Theorems 3.1 and 3.2, respectively. This means that neither the OAF protocol nor the NAF system using a diagonal covariance matrix are optimal, as illustrated by the figures in this section. Finally, note that all capacity results presented above have been verified through exhaustive search.

3.6 Chapter Summary

In this chapter, we derived the capacity of the static cooperative NAF system under individual and global power constraints. By analyzing all local maximizers of the mutual information, expressions for the optimal covariance matrix at the source and power allocation at relay were obtained. Specifically, it was demonstrated that for the per-node power constrained system, the capacity is achieved by either a DT, a NAF-FR, or a NAF-BF scheme depending on the transmission power and network configuration. For the joint power constraint scenario, it was shown that only a DT or NAF-BF protocol can achieve the capacity. By further studying the system in high and low transmission powers, we
showed that NAF relaying is useful only when the relay has large power and the source does not. However, in medium power regions, the NAF protocol can provide significant gains over the DT scheme as illustrated by the symmetric and linear network models.
Chapter 4

Achievable Rate and Power Allocation of AF Relay Systems over Fading Channels

In the previous chapter, we studied the capacity of the static NAF system in which the transmitted codeword spans over a single channel realization. Such analysis involved the maximization of the conditional mutual information. As discussed in the introduction of Chapter 3, the capacities for the DHAF, OAF and TWAF protocols over a static environment have all been addressed in the literature. On the other hand, when the transmitted codeword spans over several realizations, the ergodic achievable rate becomes the criterion of interest. In this fading environment, the capacity study involves two steps. At first, one needs to compute the ergodic achievable rate by averaging the conditional mutual information in (2.20), (2.24), (2.26), and (2.29) over the channel realizations $h$. This expectation will be different depending on the selected amplification coefficient, e.g., for either the FG or the CI coefficient. Optimization must then be applied on the derived expressions to find the optimal average power allocation schemes that maximize the achievable rate or sum rate for OW or TW AF systems, respectively. In this regard, the following works have investigated achievable rate expressions and power allocation schemes for HD AF systems over fast fading channels:

1Parts of this chapter have been presented at the 2012 IEEE Vehicular Technology Conference in Québec City, Canada [76], and the 2013 IEEE International Conference on Communications in Budapest, Hungary [77]; and have been published in the IEEE Transactions on Vehicular Technology [78].
For **DHAF** systems using the **CI** coefficient, an infinite series representation of the achievable rate and two upper bounds in terms of the Meijer-G function were derived in [79] for Rayleigh channels. A single integral expression for the rate of the system over Nakagami-m was proposed in [80] by using a scaled version of the harmonic mean to approximate the instantaneous end-end SNR. Another single integral expression was proposed in [81] using a probability density function (PDF) approach, where the PDF is given as a finite/infinite sum of v-th order modified Bessel functions of the second kind for Rayleigh, Nakagami (integer m) and Rician fading channels. Considering generalized fading, an expression for the rate was given in [82] as a truncated series of Meijer-G terms. Also for general fading channels, a double integral expression was presented in [83] by applying a novel generalized transformed characteristic function.

For **DHAF** systems using the **FG** coefficient, a closed-form expression of the achievable rate was provided in [84] as a finite sum of integrals based on random matrix theory and assuming a multi-antenna system with unit-variance Rayleigh fading. For the single-antenna scheme over Nakagami-m, a single-integral expression of the rate was given in [85] along with a Taylor approximation. Upper and lower bounds for such system were also provided in [86] in terms of the Lommel and Meijer-G functions. Finally, tight upper and lower bounds over different fading distributions were presented in [87] in terms of the Meijer-G or sum of exponential integral functions of order n.

For **OAF** systems, bounds on the achievable rate were first derived in [88] for CI and FG coefficients over Rayleigh channels. A moment-generating function approach was proposed in [89] to calculate the rate of the OAF-FG system over generalized fading channels. For the OAF-FG system with multiple antennas, a closed-form expression was derived in [90] following a similar approach as in [84].

As discussed in Chapter 3, the capacity of the **FG- and CI-NAF** channel was investigated in [22] over Rayleigh channels. In [22], it was shown that the input distribution must be Gaussian with a diagonal covariance matrix to maximize the achievable rate. However, the computation of the rate and power allocation that achieves the capacity was performed numerically in [22]. Interestingly, by examining the rate in high SNR regimes, it was demonstrated in [22] that NAF relaying is not useful compared to
the DT scheme. This result, however, is only applicable to the FG system over unit variance channels.

- For TWAF systems, bounds on the sum rate were derived in [91, 92] for Rayleigh fading. Power allocation schemes to maximize such bounds were also proposed in [91,92] assuming equal power at the source nodes. Based on an approximation to the instantaneous SNR, the works in [93,94] presented achievable rate expressions for the TW system over Rayleigh and Nakagami-\(m\) channels, respectively. The expressions in [93,94] are given in terms of Fox’s \(H\)-function and equal power allocation among all the nodes is assumed. Note that only the CI scheme was considered in [91–94].

In the above works, the achievable rates are obtained in terms of complicated mathematical functions. Although many of these functions provide an accurate representation so that Monte Carlo simulations are no longer needed, they do not provide insightful expressions of the achievable rate. In particular, most of the above works resort to plotting when discussing the behavior of the considered AF system. As a result, it is very difficult to obtain optimal average power allocation strategies based on channel statistics (rather than instantaneous values) to further improve the achievable rates or sum rates of OW and TW AF systems. To the best of our knowledge, analyses on optimal average power allocation schemes among the nodes over fading channels are scarce in the literature and numerical optimization methods are usually needed. In addition, the techniques proposed in the current literature are applicable to either OW or TW relaying using either the CI or the FG amplification coefficient. As such, it is not clear what the effect of using the FG and CI coefficients is to the performance of different AF protocols, and how these protocols compare to one another.

Inspired by the above observations, this chapter proposes a general approach to study ergodic achievable rates and power allocation schemes for HD single-relay AF systems over Rayleigh fading channels. Our main idea is to rely on the capacity of a two-branch maximal-ratio combining (MRC) system [95,96] and on a simple approximation to the logarithm to derive tight approximations to the achievable rate in high and low transmission power regions. Different from previous works, the proposed approach is applicable to OW and TW protocols using either the CI or the FG coefficient. Furthermore, the proposed approach results in insightful expressions from which comparisons among different protocols/coefficients, and optimal average power allocations can be obtained. Specifically,
approximations are first derived for the DHAF, NAF and two-phase TW relaying schemes using both the CI and FG coefficients. The derived approximations are tight in their respective power ranges and easy to analyze, as they are given in terms of the well-known exponential integral. The systems using the FG and CI coefficients are then compared at high and low power regimes. While the CI technique is better for DH and TW schemes at high powers, the FG scheme is superior in low powers for OW systems and in high powers for the NAF protocol. Asymptotic power allocation schemes based on the derived approximations are then proposed to maximize the rate or sum rate for OW and TW systems, respectively. For the specific case of the DH scheme with CI, a closed-form expression of the rate is derived using the MRC approach and bisection on the closed-form derivative is proposed to obtain the optimal power allocation. Finally, the AF protocols of interest are compared to one another and to the DT scheme. Although OW relaying is shown to be inferior to the DT scheme in terms of capacity and sum rate in high and low power regimes, TW relaying might be advantageous at high powers under the right channel conditions.

The rest of the chapter is organized as follows. The problem formulation is introduced in Section 4.1. Approximations to the achievable rates of the considered systems are presented in Section 4.2 along with comparisons between the CI and FG techniques. Optimal power allocation schemes to maximize the rate and sum rate are derived in Section 4.3. The special case of the DH system using the CI coefficient is discussed in Section 4.4. Rate comparisons among different protocols are studied in Section 4.5. To verify our theoretical work, simulation results are then presented in Section 4.6. Finally, Section 4.7 concludes the chapter.

4.1 Problem Formulation

In this chapter, we consider the fast fading scenario in which the coherence time of the channel is small enough so that the transmitted codeword spans several realizations. Specifically, Rayleigh block fading is assumed and thus the channel gains among the nodes are distributed as \( h_l \sim \mathcal{CN}(0, \phi_l) \), \( l \in \{0, 1, 2\} \). The variances \( \Phi = [\phi_0, \phi_1, \phi_2] \) account for different pathloss effects over the links. Recall that \( \alpha_l = |h_l|^2 \). Since all gains are complex circular Gaussian, \( \alpha_l \) is exponentially distributed with mean \( \phi_l \). In this ergodic environment, the channel gains are rapidly changing and thus only receiver CSI is considered. In particular, the destination node has full CSI, whereas the relay has either statistical...
or full knowledge of the incoming links. Two types of amplification coefficients can then be used at the relay, namely, the FG and the CI coefficients with $\Xi = \phi$ and $\Xi = \alpha$ in (2.19) and (2.28). Both the OW and the TW relay protocols described in Section 2.5 are considered. For OW relaying, the input-output relation of the NAF channel is given by (2.16). This protocol includes the OAF with $q_2 = 0$ in (2.23) and the DT scheme in (2.21) with $z_2 = 0$ as special cases. Moreover, the DHAF scheme in (2.25) is obtained when the direct link is blocked $\phi_0 \approx 0$. For TW relaying, we consider the two-phase TW protocol with input-output relation given by (2.27).

For a given amplification coefficient and power allocation scheme $q = [q_1, q_2, z_2]$, it was shown in [22] that the unconditional mutual information between the input and output of the NAF channel $I = \mathbb{E}[I|h]$ is maximized by using Gaussian inputs with a diagonal covariance matrix $Q$. The achievable rate conditioned on $h$ is then given as (2.20) with $q_{12} = 0$. For TW relaying, the conditional rate in the $S_k \rightarrow S_i$ direction ($i \in \{1, 2\}, k = 3-i$) is given by (2.29). The unconditional achievable rate is then written as $I_i = \mathbb{E}[I_i|h]$, and the unconditional achievable sum rate is $I_{\text{sum}} = I_1 + I_2$. The objective of this chapter is to maximize the unconditional achievable rate or sum rate for different OW or TW relay protocols, respectively. To provide a fair comparison among different schemes, we focus on the global constraint scenario similar to [91,92]. Let $P_s = P_r = P_t$ in (2.16), $P_{s1} = P_{s2} = P_t$ in (2.27) and $q_1 + q_2 + z_2 \leq q_t$. The system is then allowed to spend an average power of up to $q_t P_t$ per transmission frame. To have an average power constraint of $P_t$ per symbol period, the value of $q_t$ can be set to the number of phases required to communicate from $S_1$ to $S_2$ in OW relaying, or to exchange symbols between the two source nodes in TW schemes. In the case of OW relaying, we consider the maximum ergodic achievable rate, i.e., the power-constrained capacity, as the performance criterion. Specifically, the capacity can be calculated as

$$C = \max_{q_1, q_2, z_2 \geq 0} \mathbb{E}[I|h] \quad \text{s.t.} \quad q_1 + q_2 + z_2 \leq q_t,$$

(4.1)

where the conditional mutual information $I|h$ is given by (2.20), (2.24), or (2.26) depending on the protocol under consideration. For TW systems, our objective is to maximize the achievable sum rate, which can be expressed as

$$C_{\text{sum}} = \max_{q_1, q_2, z_2 \geq 0} \mathbb{E}[I_1|h + I_2|h] \quad \text{s.t.} \quad q_1 + q_2 + z_2 \leq q_t,$$

(4.2)
where $I_i|h$ is given in (2.29). For both problems in (4.1) and (4.2), the CI and FG techniques will be addressed.

Finding the solution to the optimization problems in (4.1) and (4.2) involves two steps. At first, one needs to compute the unconditional achievable rates for a given power allocation scheme. Such rates then need to be optimized with respect to $q$. In the following, we present solutions to these problems by focusing on the high and low transmission power regimes, i.e., high and low values of $P_t$. By first considering a fixed average power allocation, we derive closed-form yet tight approximations to the achievable rates. These closed-form solutions will then be used to find the optimal power allocation schemes. The impact of the selected amplification coefficient, comparisons among different protocols, and insights on (4.1) and (4.2) shall also be analyzed and discussed.

4.2 Achievable Rates and Closed-form Approximations

In this section, we derive approximations to the achievable rates and provide comparisons between the CI and FG techniques. As shall be shown in Sections 4.3 and 4.6, the approximations are tight in their respective power ranges and provide important insights on the problems in (4.1) and (4.2). To this end, we first outline a general approach that can be applied to analyze the rates of the considered AF protocols in high and low power regions.

At high power regions, our approach is to exploit the capacity of a two-branch MRC system. MRC is a receiver combining scheme that maximizes the instantaneous SNR at the output of the combiner [40]. After co-phasing and weighting each branch by the optimal MRC coefficients, the combiner SNR is given as the sum of the instantaneous SNRs in each branch. When the branches experience independent Rayleigh fading, the combiner SNR is given as a sum of independent exponentially distributed random variables. The following proposition, which is based on the capacity of such a MRC system with two branches having unequal average SNRs [95,96], will be used to examine the rates of the considered systems.

**Proposition 4.1.** Let $a_0$ be a non-negative constant, and $\omega_1$ and $\omega_2$ be independent exponentially distributed random variables with means $\mu_1$ and $\mu_2$, respectively. Define

$$J(x) = \exp(x)E_1(x),$$  \hspace{1cm} (4.3)
where $E_1(\cdot)$ is the exponential integral [97]:

$$E_1(x) = \int_x^\infty \frac{e^{-u}}{u} \, du = -\left( \gamma + \ln(x) + \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n!} \right), \quad (4.4)$$

and $\gamma$ is the Euler constant. Then, for $a_0 > 0$, one has

$$\mathbb{E}[\ln(a_0 + \omega_1)] = \ln(a_0) + \mathcal{J}(a_0/\mu_1), \quad (4.5a)$$

$$\mathbb{E}[\ln(a_0 + \omega_1 + \omega_2)] = \begin{cases} 1 + \ln(a_0) + \left( 1 - \frac{a_0}{\mu_1} \right) \mathcal{J}\left( \frac{a_0}{\mu_1} \right), & \mu_1 = \mu_2 \\ \ln(a_0) + \frac{\mu_1 \mathcal{J}(a_0/\mu_1) - \mu_2 \mathcal{J}(a_0/\mu_2)}{\mu_1 - \mu_2}, & \mu_1 \neq \mu_2, \end{cases} \quad (4.5b)$$

and for $a_0 = 0$,

$$\mathbb{E}[\ln(\omega_1)] = \ln(\mu_1) - \gamma, \quad (4.6a)$$

$$\mathbb{E}[\ln(\omega_1 + \omega_2)] = \begin{cases} 1 - \gamma + \ln(\mu_1), & \mu_1 = \mu_2 \\ \frac{\mu_1 \ln(\mu_1) - \mu_2 \ln(\mu_2)}{\mu_1 - \mu_2} - \gamma, & \mu_1 \neq \mu_2. \end{cases} \quad (4.6b)$$

**Proof.** See Appendix B.1. 

From the above proposition, approximations to the achievable rate in high power regions can be derived as follows. First, the conditional rate of the considered AF systems can be written in general form as $\log(1 + f(h))$, where $f(h)$ can be represented as a fraction of two functions, i.e., $f(h) = f_1(h)/f_2(h)$. The achievable rate can then be written as $\mathbb{E}[\log(1 + f(h))] = \mathbb{E}[\log(f_1(h) + f_2(h))] - \mathbb{E}[\log(f_2(h))]$. At high transmission powers, one can ignore the lower order terms of $f_1(h) + f_2(h)$ and $f_2(h)$. The expectations will have a form similar to either (4.5) or (4.6). The approximation to the achievable rate at high powers is then obtained by ignoring the lower order terms and using Proposition 4.1 to take the expectation of the remaining terms.

At low transmission powers, the conditional rate can be approximated as $\mathbb{E}[\log(1 + f(h))] \approx \mathbb{E}[f(h)]/\ln(2)$, where the fact that $\ln(1 + x) \approx x$ for small $x > 0$ has been used. One can then easily take the expectation of $f(h)$ by ignoring the higher order terms.

Based on the above approaches, in the following, we shall analyze the achievable rates for each of the considered AF systems in further detail. Without loss of generality, we assume unit noise power at all nodes $N_d = N_r = N_{di} = 1$. Note that this normalization
does not affect the generality of the analysis as the noise variances can be easily included in the channel variances $\phi_i$.

### 4.2.1 OW DHAF Systems

To illustrate the proposed approach, consider first the DH system using the coefficients in (2.19). As previously mentioned, the direct $S_1-S_2$ is under heavy shadowing in the DH protocol. Substituting (2.19) into (2.26), the conditional achievable rate for this system simplifies to

$$I_{DH} | h = \frac{1}{2} \log \left( 1 + \frac{P^2_t q_1 z_2 \alpha_1 \alpha_2}{P_t (q_1 \Xi_1 + z_2 \alpha_2) + 1} \right)$$

$$= \frac{1}{2} \log [1 + f_{DH}(h)], \quad (4.7)$$

where $\Xi_1 = \phi_1$ for the FG coefficient and $\Xi_1 = \alpha_1$ for the CI one. As outlined before, in this case

$$f_{DH}(h) = f_{1,DH}(h)/f_{2,DH}(h) \quad \text{with} \quad f_{1,DH}(h) = P^2_t q_1 z_2 \alpha_1 \alpha_2 \quad \text{and} \quad f_{2,DH}(h) = P_t (q_1 \Xi_1 + z_2 \alpha_2) + 1.$$ At high powers,

$$f_{1,DH}(h) + f_{2,DH}(h) = P^2_t q_1 z_2 \alpha_1 \alpha_2 + P_t (q_1 \Xi_1 + z_2 \alpha_2) + 1$$

$$= P^2_t q_1 z_2 \alpha_1 \alpha_2 + O(P_t),$$

$$f_{2,DH}(h) = P_t (q_1 \Xi_1 + z_2 \alpha_2) + 1$$

$$= P_t (q_1 \Xi_1 + z_2 \alpha_2) + O(1),$$

where $O(\cdot)$ denotes the big-O notation as $P_t \rightarrow \infty$. The achievable rate in (4.7) can then be written in high power regions as

$$I_{DH} | h = \frac{1}{2} \log \left( \frac{f_{1,DH}(h) + f_{2,DH}(h)}{f_{2,DH}(h)} \right)$$

$$= \frac{1}{2} \log \left( \frac{P^2_t q_1 z_2 \alpha_1 \alpha_2 + O(P_t)}{P_t (q_1 \Xi_1 + z_2 \alpha_2) + O(1)} \right). \quad (4.8)$$
The rate for the FG system can be approximated from (4.8) by ignoring the lower order terms and substituting $\Xi_1 = \phi_1$ as

$$I_{DH}^{FG} \approx \frac{1}{2} \log(P_t q_1 z_2) + \frac{1}{2 \ln(2)} \left\{ E[\ln(\alpha_1)] + E[\ln(\alpha_2)] - E[\ln(q_1 \phi_1 + z_2 \alpha_2)] \right\}$$

$$= \frac{1}{2} \log(P_t) + \log(\sqrt{z_2 \phi_2}) - \frac{1}{2 \ln(2)} J \left( \frac{q_1 \phi_1}{z_2 \phi_2} \right) - \frac{\gamma}{\ln(2)}, \quad (4.9)$$

where the first two expectations are solved using (4.6a), and the last one using (4.5a) in Proposition 4.1. Similarly, by substituting $\Xi_1 = \alpha_1$, the rate of the CI system can be approximated at high powers from (4.8) and Proposition 4.1 as

$$I_{DH}^{CI} \approx \frac{1}{2} \log(P_t q_1 z_2) + \frac{1}{2 \ln(2)} \left\{ E[\ln(\alpha_1)] + E[\ln(\alpha_2)] - E[\ln(q_1 \alpha_1 + z_2 \alpha_2)] \right\}$$

$$= \frac{1}{2} \log(P_t) + \log(\sqrt{z_2 \phi_2}) - \frac{1}{2 \ln(2)} \log(\frac{q_1 \phi_1}{z_2 \phi_2}) - \frac{\gamma}{\ln(2)}, \quad q_1 \phi_1 = z_2 \phi_2, \quad (4.10)$$

where (4.6a) and (4.6b) have been used. From (4.8), the rate difference between the FG and CI techniques can be approximated as

$$I_{DH}^{FG} - I_{DH}^{CI} \approx \frac{1}{2} E \left[ \log \left( \frac{q_1 \alpha_1 + z_2 \alpha_2}{q_1 \phi_1 + z_2 \phi_2} \right) \right]. \quad (4.11)$$

Applying Jensen’s inequality to $\alpha_1$ in (4.11), it can be easily shown that $I_{DH}^{FG} - I_{DH}^{CI} \leq 0$. Therefore, CI is more beneficial than FG in high power regions.

In low power regions, as outlined before, the conditional achievable rate can be approximated from (4.7) as

$$I_{DH|h} \approx \frac{f_{DH}(h)}{2 \ln(2)} = \frac{P_t^2 q_1 z_2 \alpha_1 \alpha_2}{2 \ln(2)[P_t(q_1 \Xi_1 + z_2 \alpha_2) + 1]},$$

where the fact that $\ln(1 + x) \approx x$ for small $x > 0$ has been used. For both FG and CI systems, the unconditional achievable rate can then be approximated as

$$I_{DH} \approx E \left[ \frac{P_t^2 q_1 z_2 \alpha_1 \alpha_2}{2 \ln(2)[1 + O(P_t)]} \right] = \frac{P_t^2 q_1 z_2 \phi_1 \phi_2}{2 \ln(2)}, \quad (4.12)$$

where now $O(\cdot)$ denotes the big-$O$ notation as $P_t \rightarrow 0$. Hence, the rates of the CI and FG systems are approximated as

$$I_{DH}^{CI} \approx \frac{1}{2} \log(P_t) + \log(\sqrt{z_2 \phi_2}) - \frac{1}{2 \ln(2)} \log(\frac{q_1 \phi_1}{z_2 \phi_2}) - \frac{\gamma}{\ln(2)}, \quad q_1 \phi_1 = z_2 \phi_2, \quad (4.10)$$

where (4.6a) and (4.6b) have been used. From (4.8), the rate difference between the FG and CI techniques can be approximated as

$$I_{DH}^{FG} - I_{DH}^{CI} \approx \frac{1}{2} E \left[ \log \left( \frac{q_1 \alpha_1 + z_2 \alpha_2}{q_1 \phi_1 + z_2 \phi_2} \right) \right]. \quad (4.11)$$

Applying Jensen’s inequality to $\alpha_1$ in (4.11), it can be easily shown that $I_{DH}^{FG} - I_{DH}^{CI} \leq 0$. Therefore, CI is more beneficial than FG in high power regions.

In low power regions, as outlined before, the conditional achievable rate can be approximated from (4.7) as

$$I_{DH|h} \approx \frac{f_{DH}(h)}{2 \ln(2)} = \frac{P_t^2 q_1 z_2 \alpha_1 \alpha_2}{2 \ln(2)[P_t(q_1 \Xi_1 + z_2 \alpha_2) + 1]},$$

where the fact that $\ln(1 + x) \approx x$ for small $x > 0$ has been used. For both FG and CI systems, the unconditional achievable rate can then be approximated as

$$I_{DH} \approx E \left[ \frac{P_t^2 q_1 z_2 \alpha_1 \alpha_2}{2 \ln(2)[1 + O(P_t)]} \right] = \frac{P_t^2 q_1 z_2 \phi_1 \phi_2}{2 \ln(2)}, \quad (4.12)$$

where now $O(\cdot)$ denotes the big-$O$ notation as $P_t \rightarrow 0$. Hence, the rates of the CI and FG systems are approximated as
systems converge in low power regions and the performance is asymptotically independent of the selected coefficient. However, for slightly higher powers, it is straightforward to show that

\[
I_{\text{FG}} - I_{\text{CI}} = \frac{1}{2} \cdot \mathbb{E} \left[ \log(1 + f_{\text{FG}}(h)) - \log(1 + f_{\text{CI}}(h)) \right]
\]

\[
\approx \frac{1}{2 \ln(2)} \cdot \mathbb{E} \left[ f_{\text{FG}}(h) - f_{\text{CI}}(h) \right]
\]

\[
= \frac{1}{2 \ln(2)} \cdot \mathbb{E} \left[ \frac{P_t^3 q_1^2 z_2 \alpha_1 \alpha_2 (\alpha_1 - \phi_1)}{(P_t z_2 \alpha_2 + P_t q_1 \alpha_1 + 1)(P_t q_1 \phi_1 + P_t z_2 \alpha_2 + 1)} \right] \geq 0, \quad (4.13)
\]

as \( f_{\text{FG}}(h) - f_{\text{CI}}(h) \) is convex with respect to \( \alpha_1 \). Therefore, FG is better than CI, and the CI system approaches the FG one from below at sufficiently low powers.

### 4.2.2 OW Cooperative Systems

Consider now the general cooperative NAF system in (2.20) with \( q_{12} = 0 \). As noted before, the OAF and DT schemes are special cases of this protocol. First, consider the NAF system with \( q_1, q_2, z_2 > 0 \). The conditional achievable rate of this system using the coefficients in (2.19) can be written as

\[
I_{\text{NAF}}|h = \frac{1}{2} \log \left[ 1 + f_{\text{NAF}}(h) \right], \quad (4.14)
\]

where \( f_{\text{NAF}}(h) = f_{1,\text{NAF}}(h)/f_{2,\text{NAF}}(h) \) with

\[
f_{1,\text{NAF}}(h) = P_t^3 (q_1^2 q_2 \alpha_0^2 \Xi_1) + P_t^2 q_1 (q_2 \alpha_0^2 + q_1 \alpha_0 \Xi_1 + z_2 \alpha_0 \alpha_2 + z_2 \alpha_1 \alpha_2 + q_2 \alpha_0 \Xi_1)
\]

\[
+ P_t \alpha_0 (q_1 + q_2),
\]

\[
f_{2,\text{NAF}}(h) = P_t (q_1 \Xi_1 + z_2 \alpha_2) + 1.
\]

As in the DH case, high power approximations to the achievable rate of the FG and CI systems can be obtained from (4.14) and Proposition 4.1 by ignoring the lower order order terms.
Achievable Rate and Power Allocation of AF over Fading Channels

4 Achievable Rate and Power Allocation of AF over Fading Channels

Thus, by following the same procedure, the achievable rate for the FG system is approximated by setting $\Xi_1 = \phi_1$ and solving the expectations as

$$I_{\text{FG}} \approx \frac{1}{2} \log(p^2 q_1 q_2 + \phi_1) + \frac{1}{2 \ln(2)} \left\{ 2 \cdot E[\ln(\alpha_0)] - E[\ln(q_1 \phi_1 + z_2 \alpha_2)] \right\}$$

$$= \log(P_t) + \log(\phi_0 \sqrt{q_1 q_2}) - \frac{1}{2 \ln(2)} J \left( \frac{q_1 \phi_1}{z_2 \phi_2} \right) - \gamma, \quad (4.15)$$

where (4.6a) and (4.5a) have been used. The rate of the CI system with $\Xi_1 = \alpha_1$ can be similarly approximated in high power regions as

$$I_{\text{CI}} \approx \frac{1}{2} \log(P_t q_1^2) + \frac{1}{2 \ln(2)} \left\{ 2 \cdot E[\ln(\alpha_0)] + E[\ln(\alpha_1)] - E[\ln(q_1 \alpha_1 + z_2 \alpha_2)] \right\}$$

$$= \log(P_t) + \log(\phi_0) - \frac{\gamma}{\ln(2)}, \quad (4.16)$$

by applying (4.6a) and (4.6b). The rate difference between the FG and CI system for the NAF protocol can be written from (4.14) and Proposition 4.1 as

$$I_{\text{FG}} - I_{\text{CI}} \approx \frac{1}{2 \ln(2)} \times \left\{ 1 - J(1), \quad z_2 \phi_2 \left[ \frac{\ln(q_1 \phi_1) - \ln(z_2 \phi_2)}{q_1 \phi_1 - z_2 \phi_2} \right] - J \left( \frac{q_1 \phi_1}{z_2 \phi_2} \right), \quad q_1 \phi_1 \neq z_2 \phi_2 \right\} \quad (4.17)$$

As shown in Appendix B.2, (4.17) is strictly positive and hence $I_{\text{FG}} - I_{\text{NAF}} \geq I_{\text{CI}}$ in high power regions. Interestingly and contrary to DH schemes, the FG coefficient outperforms the CI one for NAF systems operating in high transmission powers. This holds despite the fact that the FG technique requires less channel knowledge at the relay than the CI one. In low power regions, the rate for the FG and CI systems can be approximated following the
proposed approach as
\[
I_{\text{NAF}} \approx \mathbb{E} \left[ \frac{f_{\text{NAF}}(h)}{2\ln(2)} \right] = \mathbb{E} \left[ \frac{P_t \alpha_0[q_1 + q_2] + O(P_t^2)}{2\ln(2)[1 + O(P_t)]} \right] \approx \frac{P_t (q_1 + q_2) \phi_0}{2\ln(2)}. \tag{4.18}
\]

Furthermore, \(f_{\text{FG}}^{\text{NAF}}(h) - f_{\text{CI}}^{\text{NAF}}(h)\) can be shown to be a convex function with respect to \(\alpha_1\) and thus \(I_{\text{FG}}^{\text{NAF}} \geq I_{\text{CI}}^{\text{NAF}}\). As in DH systems, the FG system then outperforms the CI one in low powers, and the difference between these two schemes decreases as the power decreases.

Consider now the OAF protocol in which the source remains silent in the second phase, i.e., \(q_2 = 0\). In this case, the conditional achievable rate in (2.24) simplifies to
\[
I_{\text{OAF}}(h) = \frac{1}{2} \log \left( 1 + P_t q_1 \alpha_0 + \frac{P_t^2 q_1 z_2 \alpha_1 \alpha_2}{P_t(q_1 \Xi_1 + z_2 \alpha_2) + 1} \right)
= \frac{1}{2} \log [1 + f_{\text{OAF}}(h)], \tag{4.19}
\]
where \(f_{\text{OAF}}(h) = f_{1,\text{OAF}}(h)/f_{2,\text{OAF}}(h)\) with
\[
f_{1,\text{OAF}}(h) = P_t^2 q_1 (q_1 \alpha_0 \Xi_1 + z_2 \alpha_0 \alpha_2 + z_2 \alpha_1 \alpha_2) + P_t q_1 \alpha_0
= P_t^2 q_1 (q_1 \alpha_0 \Xi_1 + z_2 \alpha_0 \alpha_2 + z_2 \alpha_1 \alpha_2) + O(P_t),
\]
\[
f_{2,\text{OAF}}(h) = P_t (q_1 \Xi_1 + z_2 \alpha_2) + 1
= P_t (q_1 \Xi_1 + z_2 \alpha_2) + O(1).
\]

At high transmission powers, the unconditional rate can be approximated by ignoring the lower order terms as
\[
I_{\text{OAF}} \approx \frac{1}{2} \log(P_t) + \frac{1}{2} \mathbb{E} \left[ \log \left( \frac{q_1 (q_1 \alpha_0 \Xi_1 + z_2 \alpha_0 \alpha_2 + z_2 \alpha_1 \alpha_2)}{q_1 \Xi_1 + z_2 \alpha_2} \right) \right].
\]

Unfortunately, due to the cross terms in the numerator, the proposed MRC approach cannot be used to solve for the above expectation. Since the second term is independent of \(P_t\), we can only say that \(I_{\text{OAF}} = \frac{1}{2} \log(P_t) + O(1)\) in high power regimes. In low powers, following a procedure similar to the NAF system, the achievable rate for both FG and CI systems
can be approximated from (4.19) as

\[ I_{OAF} \approx \mathbb{E} \left[ \frac{f_{OAF}(\mathbf{h})}{2 \ln(2)} \right] = \mathbb{E} \left[ \frac{P_t q_1 \alpha_0 + O(P_t^2)}{2 \ln(2) [1 + O(P_t)]} \right] = \frac{P_t q_1 \phi_0}{2 \ln(2)}. \tag{4.20} \]

As in the NAF scheme, \( I_{OAF}^{FG} \geq I_{OAF}^{CI} \) at low powers and the CI system approaches the FG one from below.

Finally, consider the DT scheme in which the relay is not used for transmission with \( z_2 = 0 \). The unconditional rate is simply obtained by taking the expectation of (2.21) with respect to \( \alpha_0 \). The rate in this case can be written in closed-form as

\[ I_{DT} = \frac{1}{2 \ln(2)} \left[ J \left( \frac{1}{q_1 P_t \phi_0} \right) + J \left( \frac{1}{q_2 P_t \phi_0} \right) \right], \tag{4.21} \]

which can be respectively approximated at high and low powers as

\[ I_{DT} \approx \log(P_t \phi_0 \sqrt{q_1 q_2}) - \frac{\gamma}{\ln(2)}, \quad \text{and} \quad I_{DT} \approx \frac{P_t \phi_0 (q_1 + q_2)}{2 \ln(2)}. \tag{4.22} \]

### 4.2.3 TW Systems

For the two-phase TWAF system using the amplification coefficients in (2.28), the conditional achievable rate in (2.29) can be written as

\[ I_{i,\text{TW}}|\mathbf{h} = \frac{1}{2} \log \left[ 1 + \frac{P_t^2 q_k z_2 \alpha_1 \alpha_2}{P_t (q_1 \Xi_1 + q_2 \Xi_2 + z_2 \alpha_i) + 1} \right], \tag{4.23} \]

where \( \Xi_1 = \phi_1 \) and \( \Xi_2 = \phi_2 \) for FG, and \( \Xi_1 = \alpha_1 \) and \( \Xi_2 = \alpha_2 \) for CI. Following the same approach as in the previous subsections, the expectation of (4.23) can be respectively approximated in high power regions for FG and CI systems as

\[ I_{i,\text{TW}}^{FG} \approx \frac{1}{2} \log(P_t q_k z_2) + \frac{1}{2 \ln(2)} \cdot \left\{ \mathbb{E}[\ln(\alpha_1)] + \mathbb{E}[\ln(\alpha_2)] - \mathbb{E}[\ln(q_1 \phi_1 + q_2 \phi_2 + z_2 \alpha_i)] \right\} = \frac{1}{2} \log(P_t) + \log \left( \sqrt{\frac{q_k z_2 \phi_1 \phi_2}{q_1 \phi_1 + q_2 \phi_2}} \right) - \frac{1}{2 \ln(2)} J \left( \frac{q_1 \phi_1 + q_2 \phi_2}{z_2 \phi_i} \right) - \frac{\gamma}{\ln(2)}. \tag{4.24} \]
and

\[ I_{i,TW}^{CI} \approx \frac{1}{2} \log(P_i q_i z_d) + \frac{1}{2 \ln(2)} \times \left\{ \mathbb{E}[\ln(\alpha)] + \mathbb{E}[\ln(\alpha_2)] - \mathbb{E}[\ln(q_1 \alpha_1 + q_2 \alpha_2 + z_d \alpha_i)] \right\} \]

\[ = \frac{1}{2} \log(P_i) + \left\{ \frac{\log(\sqrt{z_d \phi_i}) - \frac{\gamma}{2 \ln(2)} - \frac{1}{2 \ln(2)}}{(q_i + z_d) \phi_i} + q_k \phi_k \log(\frac{q_i + z_d}{z_d \phi_i} + q_k \phi_k \log(z_d \phi_i)} - \frac{\gamma}{2 \ln(2)}; \quad q_k \phi_k = (q_i + z_d) \phi_i \right\} \]

\[ \text{(4.25)} \]

where we have applied (4.6a) and (4.5a) to (4.24), and (4.6) to (4.25). Using Jensen’s inequality, it can be easily shown that \( I_{i,TW}^{CI} \geq I_{i,TW}^{FG} \) at high powers. Thus, similar to the DH scheme, the CI system outperforms the FG one in high power regions.

In low power regions, the achievable rates for both FG and CI systems converge and can simply be approximated from (4.23) as

\[ I_{i,TW} \approx P_i^2 q_i z_d \phi_1 \phi_2. \]  

\[ \text{(4.26)} \]

In this case, \( f_i^{FG}(h) - f_i^{CI}(h) \) is neither convex nor concave and thus one cannot say which of the two coefficients is dominant at low powers.

### 4.2.4 Remarks

First, observe from (4.9), (4.10), (4.15), (4.16), (4.24) and (4.25) that the derived high power approximations are easy to evaluate as they involve only the exponential, logarithmic and exponential integral functions. In general, the high power approximations derived in this section can be written as

\[ I \approx m \cdot \log(P_i) + G_{\text{high}}(q, \Phi), \]  

\[ \text{(4.27)} \]

where \( m \in \{1/2, 1\} \) is the multiplexing gain and \( G_{\text{high}}(\cdot, \cdot) \) is the power gain. As shown in the previous subsections, the multiplexing gain depends only on the transmission protocol, whereas the power gain also depends on the power allocation \( q \), the network configuration \( \Phi \), and the selected amplification coefficient (FG or CI). As shall be shown in Section 4.6, the derived bounds are tight in medium to high power regions. The impact of the power gain \( G_{\text{high}}(\cdot, \cdot) \) can thus be significant in these power regimes, as will be illustrated shortly.
In low power regions, it can be seen from previous subsections that the derived approximations can be written as

\[ I \approx P^n_t \cdot G_{\text{low}}(q, \Phi). \quad (4.28) \]

Similar to (4.27), \( n \in \{1, 2\} \) only depends on the transmission protocol, whereas \( G_{\text{low}}(\cdot, \cdot) \) also depends on the power allocation, network configuration and amplification coefficient.

For ease of reference, Tables 4.1 and 4.2 summarize the approximations derived in this section in terms of \( n, m, G_{\text{high}}(\cdot, \cdot) \) and \( G_{\text{low}}(\cdot, \cdot) \). In these tables, a star (*) has been placed next to the dominant amplification coefficient when known. As shall be illustrated in Section 4.6, the approximations in Tables 4.1 and 4.2 are tight in their respective power regions. The problems in (4.1) and (4.2) can thus be solved by maximizing the approximations in these two tables without the need of lengthy Monte Carlo simulations.

**Table 4.1** High power approximations for different AF protocols.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Type</th>
<th>( m )</th>
<th>( G_{\text{high}}(q, \Phi) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>CI*</td>
<td>1/2</td>
<td>( \log(\sqrt{q_1^2 q_2^2}) - \frac{\gamma}{2 \ln(2)} - \frac{1}{2 \ln(2)}, q_1 \phi_1 = z_2 \phi_2 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \log(q_1^2 q_2^2) - \frac{\gamma}{2 \ln(2)} + \frac{1}{2 \ln(2)}, q_1 \phi_1 \neq z_2 \phi_2 )</td>
</tr>
<tr>
<td></td>
<td>FG</td>
<td>1/2</td>
<td>( \log(\sqrt{z_2^2 \phi_2^2}) = \frac{\gamma}{2 \ln(2)} - \frac{\gamma}{2 \ln(2)} )</td>
</tr>
<tr>
<td>NAF</td>
<td>CI</td>
<td>1</td>
<td>( \log(\sqrt{q_1^2 q_2^2}) - \frac{\gamma}{2 \ln(2)} - \frac{1}{2 \ln(2)}, q_1 \phi_1 = z_2 \phi_2 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \log(q_1^2 q_2^2) - \frac{\gamma}{2 \ln(2)} + \frac{1}{2 \ln(2)}, q_1 \phi_1 \neq z_2 \phi_2 )</td>
</tr>
<tr>
<td>OAF</td>
<td>CI/FG</td>
<td>1/2</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>DT</td>
<td></td>
<td>1</td>
<td>( \log(\sqrt{q_1^2 q_2^2}) - \frac{\gamma}{\ln(2)} )</td>
</tr>
<tr>
<td>TW</td>
<td>CI*</td>
<td>1/2</td>
<td>( \log(\sqrt{z_2^2 \phi_2^2}) = \frac{\gamma}{2 \ln(2)} - \frac{1}{2 \ln(2)}, q_k \phi_k = (q_i + z_2) \phi_i )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \log(q_1^2 q_2^2) - \frac{\gamma}{2 \ln(2)} + \frac{1}{2 \ln(2)}, q_k \phi_k \neq (q_i + z_2) \phi_i )</td>
</tr>
<tr>
<td>FG</td>
<td>1/2</td>
<td>( \log(\sqrt{g_k z_2^2 \phi_2^2}) = \frac{\gamma}{2 \ln(2)} - \frac{1}{2 \ln(2)} )</td>
<td></td>
</tr>
</tbody>
</table>
### Table 4.2 Low power approximations for different AF protocols.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Type</th>
<th>$n$</th>
<th>$G_{\text{low}}(\mathbf{q}, \Phi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>CI / FG</td>
<td>2</td>
<td>$\frac{q_1 z_2 \phi_1 \phi_2}{2 \ln(2)}$</td>
</tr>
<tr>
<td>NAF</td>
<td>CI / FG</td>
<td>1</td>
<td>$\frac{(q_1 + q_2) \phi_0}{2 \ln(2)}$</td>
</tr>
<tr>
<td>OAF</td>
<td>CI / FG</td>
<td>1</td>
<td>$\frac{q_1 \phi_0}{2 \ln(2)}$</td>
</tr>
<tr>
<td>DT</td>
<td></td>
<td>1</td>
<td>$\frac{\phi_0 (q_1 + q_2)}{2 \ln(2)}$</td>
</tr>
<tr>
<td>TW</td>
<td>CI / FG</td>
<td>2</td>
<td>$\frac{q_k z_2 \phi_1 \phi_2}{2 \ln(2)}$</td>
</tr>
</tbody>
</table>

### 4.3 Optimal Power Allocation

In this section, optimal power allocation schemes that maximize the achievable rate in (4.1) or the sum rate in (4.2) are investigated. As before, we focus on high and low power regions.

#### 4.3.1 DHAF Systems

The capacity of the DH system, denoted as $C_{\text{DH}}$, using either the FG or the CI coefficient in (2.19) can be obtained by maximizing the expectation of (4.7) over the channel gains. Since (4.7) is not a function of $q_2$, the feasible region in (4.1) reduces to $q_1 + z_2 \leq q_t$. By taking the first derivatives, it can be easily shown that (4.7) is strictly increasing with $q_1$ and $z_2$. Hence, the power constraint in (4.1) must be tight $q_1 + z_2 = q_t$. Furthermore, the second derivative of (4.7) over the line segment $z_2 = q_t - q_1$ ($0 \leq q_1 \leq q_t$) can be shown to be negative. The maximization of the achievable rate is thus a concave optimization problem for DH systems. The optimal power allocation $\mathbf{q}_{\text{DH}} = [q_{1,\text{DH}}, z_{2,\text{DH}}] = [q_{1,\text{DH}}, q_t - q_{1,\text{DH}}]$ is then unique and can be easily obtained by finding a stationary point in the rate.

At high enough transmission powers, the optimal power allocation that achieves the capacity in (4.1) with CI or FG can be obtained by maximizing the approximations in (4.9) or (4.10) over the line $z_2 = q_t - q_1$. Since (4.9) and (4.10) are non-linear, finding a closed-form expression for the unique stationary point $q_{1,\text{DH}}$ appears difficult for any $\Phi$. From the concavity of the problem, numerical methods can be used to find the global maximizer. However, a suboptimal power allocation in high power regions can be derived as follows. Using (4.11) and the fact that $J(x) < \ln (1 + [1/x])$ [97], the achievable rates
in (4.9) and (4.10) can be lower bounded as
\[ I_{DH}^{CI} \geq I_{DH}^{FG} \geq \frac{\log(P_t)}{2} + \frac{1}{2} \log \left( \frac{q_1 z_2 \phi_1 \phi_2}{q_1 \phi_1 + z_2 \phi_2} \right) - \frac{\gamma}{\ln(2)}. \] (4.29)

The maximization of the bound in (4.29) can also be shown to be a concave problem. By setting \( z_2 = q_t - q_1 \) and equating the derivative of (4.29) to zero, the suboptimal power allocation \( q_{DH}^{sub} = [q_{1,DH}^{sub}, q_t - q_{1,DH}^{sub}] \) at high powers can be calculated as
\[ q_{1,DH}^{sub} = \begin{cases} \frac{q_t (\phi_2 - \sqrt{\phi_1 \phi_2})}{\phi_2 - \phi_1}, & \phi_1 \neq \phi_2 \\ q_t/2, & \phi_1 = \phi_2. \end{cases} \] (4.30)

As shall be illustrated in Section 4.6, the system using the suboptimal power allocation in (4.30) performs closely to the one using the optimal power allocation that achieves the capacity in (4.1) for different network configurations.

In low power regions, it is easy to see from (4.12) and the arithmetic/geometric mean inequality that uniform power allocation \( q_{1,DH} = z_{2,DH} = q_t/2 \) is optimal for CI and FG systems. This holds regardless of the value of \( \Phi \).

### 4.3.2 Cooperative Systems

For the cooperative systems, one must find the power allocation \( q \) that maximizes the expectation of (4.14) as in (4.1). As observed in [22] and Chapter 3, the maximization in (4.1) for cooperative systems is in general not a concave problem. Given that a diagonal covariance matrix \( Q \) is optimal [22] for fading channels, it is easy to show following a similar approach as in Chapter 3 that the capacity can be written as
\[ C = \max \{ C_{DT}, C_{OAF}, C_{NAF} \}, \]
where \( C_{DT}, C_{OAF} \) and \( C_{NAF} \) are the capacities of the DT, OAF and NAF protocols, respectively.

First, it can be easily shown that the mutual information of the DT system in (4.21) is maximized when \( q_{1,DT} = q_{2,DT} = q_t/2 \) for any value of \( P_t \). The capacity of the DT scheme is then given as \( C_{DT} = \frac{1}{\ln(2)} : J(2/[P_t \phi_0 q_t]) \). At high powers, the capacity of the NAF channel can be calculated by maximizing the approximations in (4.15) and (4.16) for the FG and CI system, respectively. Observe from (4.15) that \( I_{NAF}^{FG} \) is a decreasing function of \( z_2 \) since \( J(x) > 0 \) is a decreasing function of \( x > 0 \). This means that using the relay in fact reduces the rate in high power regions. From this and (4.17), the following inequality
Achievable Rate and Power Allocation of AF over Fading Channels

holds:

\[ I_{DT} \geq I_{NAF}^{FG} \geq I_{NAF}^{CI} \]

where the equality is achieved when \( z_{2,NAF} = 0 \). Given that \( I_{DT} \) in (4.21) is maximized when \( q_{1,DT} = q_{2,DT} = q_t / 2 \), we have the following inequality

\[ C_{DT} \geq I_{DT} \geq I_{NAF}^{FG} \geq I_{NAF}^{CI} \]

(4.31)

where the equality is achieved when \( q_{1,NAF} = q_{2,NAF} = q_t / 2 \) and \( z_{2,NAF} = 0 \). The optimal power allocation for both FG and CI NAF systems is thus \( q_{NAF} = [q_t / 2, q_t / 2, 0] \). The capacity of the NAF system is then \( C_{NAF} = C_{DT} \) and is achieved when the relay is inactive. This observation is similar to that shown in [22] for symmetric \( \phi_0 = \phi_1 = \phi_2 \) NAF channels using the FG coefficient. Furthermore, observe from Table 4.1 that the rate of the OAF system only grows as \( \frac{1}{2} \log(P_t) \) rather than logarithmically with \( P_t \) as the rate of the DT scheme. Hence, \( C_{DT} = C_{NAF} > C_{OAF} \) at sufficiently high powers. Therefore, the capacity in (4.1) for cooperative systems operating in high power regions is achieved by the DT scheme.

In low powers, it can be seen from Table 4.2 that the achievable rate of the NAF and OAF systems is independent of the power allocated to the relay \( z_2 \). As a result, the maximum rate is achieved when \( z_2 = 0 \). Therefore, at sufficiently low powers, the relay is inactive and the optimal power allocations are \( q_{NAF} = [q_t / 2, q_t / 2, 0] \) and \( q_{OAF} = [q_t, 0, 0] \). Given that \( I_{DT|h} > I_{OAF|h} \) when \( q_{1,DT} = q_{2,DT} = q_t / 2 \) and \( q_{OAF} = [q_t, 0, 0] \), \( C_{DT} = C_{NAF} > C_{OAF} \) and the capacity in (4.1) is then also achieved by the DT scheme in low power regimes. Thus, the capacity of the cooperative system operating in high and low power regions is achieved by the DT scheme for both FG and CI coefficients.

4.3.3 TW Systems

For TWAF systems, our objective is to maximize the sum rate as in (4.2). For the TW system operating in high power regions, the approximations in (4.24) and (4.25) must be used. Similar to DH systems, this is a non-linear problem and numerical methods are required. A suboptimal power allocation can, however, be obtained as in (4.30). Similar
to (4.29), the sum rate of the FG and CI systems can be lower bounded as

$$I_{\text{sum}}^{\text{TW, CI}} \geq I_{\text{sum}}^{\text{TW, FG}} \geq \frac{1}{2} \log \left( \frac{P_t^2 q_1 q_2 z_2^2 \phi_1^2 \phi_2^2}{(q_1 \phi_1 + q_2 \phi_2 + z_2 \phi_1)(q_1 \phi_1 + q_2 \phi_2 + z_2 \phi_2)} \right) - \frac{2\gamma}{\ln(2)}. \quad (4.32)$$

Although the above bound is not concave, a global maximizer can still be found using the approach in [69]. From (4.30) and [69], it can be shown that the power allocation that maximizes (4.32), $q_{\text{TW}}^{\text{sub}} = [q_{1,\text{TW}}^{\text{sub}}, q_{2,\text{TW}}^{\text{sub}}, z_{2,\text{TW}}^{\text{sub}}]$, is given by

$$q_{\text{TW}}^{\text{sub}} = \begin{cases} \left[ \frac{q_1(\phi_2 - \sqrt{\phi_1 \phi_2})}{2(\phi_2 - \phi_1)}, \frac{q_1(\sqrt{\phi_1 \phi_2} - \phi_1)}{2(\phi_2 - \phi_1)}, q_t/2 \right], & \phi_1 \neq \phi_2 \\ [q_t/4, q_t/4, q_t/2], & \phi_1 = \phi_2. \end{cases} \quad (4.33)$$

As indicated by (4.30) and (4.33), there is a strong relation between the OW and TW protocols that do not take into account the direct $S_1$-$S_2$ link. As will be shown in Section 4.6, the power allocation in (4.33) approaches closely to the optimal one and provides great advantages over the schemes previously proposed in [91, 92].

In low power regions, the sum rate of the TW system using either CI or FG can be approximated from (4.26) as

$$I_{\text{sum}}^{\text{TW}} \approx \frac{P_t^2 z_2^2 (q_1 + q_2) \phi_1 \phi_2}{2 \ln(2)}. \quad (4.34)$$

From the inequality of the geometric and arithmetic means, it is not difficult to see that the sum rate in (4.34) is maximized when 50% of the power is given to the relay, while the other 50% is equally distributed between the two sources. The optimal power allocation and the solution to (4.2) at low powers is then $q_{\text{TW}} = [q_t/4, q_t/4, q_t/2]$.

### 4.4 Closed-form Solution for DHAF System with CI Coefficient

For the specific case of the DH system with the CI coefficient, closed-form expressions of the unconditional achievable rate can be obtained using the MRC approach in Proposition 4.1. Specifically, factoring $f_{1,\text{DH}}(h) + f_{2,\text{DH}}(h) = (1 + P_t q_1 \alpha_1)(1 + P_t z_2 \alpha_2)$ in (4.7), the achievable
rate of the CI system can be expressed as

\[ I_{\text{DH}}^{\text{CI}} = 0.5 \cdot \mathbb{E} [\log(1 + P_t q_1 \alpha_1)] + 0.5 \cdot \mathbb{E} [\log(1 + P_t z_2 \alpha_2)] - 0.5 \cdot \mathbb{E} [\log(1 + P_t q_1 \alpha_1 + P_t z_2 \alpha_2)]. \]  

(4.35)

From (4.5) in Proposition 4.1, the expectations in (4.35) can be easily solved as

\[ I_{\text{DH}}^{\text{CI}} = \begin{cases} \frac{1}{2 \ln(2)} \left[ \left(1 + \frac{1}{P_t q_1 \phi_1} \right) J \left(\frac{1}{P_t q_1 \phi_1}\right) \right] - 1, & q_1 \phi_1 = z_2 \phi_2 \\ \frac{1}{2 \ln(2)} \left[ \left(1 + \frac{1}{P_t q_1 \phi_1} \right) J \left(\frac{1}{P_t q_1 \phi_1}\right) \right], & q_1 \phi_1 \neq z_2 \phi_2. \end{cases} \]  

(4.36)

Observe that (4.36) is in closed-form and holds for any value of \( P_t \).

Given that a closed-form expression is available for this system, (4.36) can also be used to derive optimal power allocations as in Section 4.3. Recall from Section 4.3.1 that the mutual information of the DH system is concave along the line \( z_2 = q_t - q_1 \). The optimal allocation can then simply be obtained by finding the unique stationary point of (4.36) along this line. With a slight abuse of notation, let \( I([q_1, z_2]) = I_{\text{DH}}^{\text{CI}} \) in (4.36). The derivative of \( I([q_1, q_t - q_1]) \) can be written when \( q_1 \phi_1 \neq (q_t - q_1) \phi_2 \) as

\[ \frac{dI([q_1, q_t - q_1])}{dq_1} = -A q_1^4 - B \phi_1^2 - E q_1 + H \phi_1^3 + F q_1^3 - G, \]  

(4.37)

where we have used the fact that \( dJ(x)/dx = J(x) - [1/x] \) [97]. In (4.37), \( A = J_2 \phi_2^3 (\phi_2 + \phi_1 - P_t q_t \phi_2^2) + J_1 \phi_1 (P_t q_t \phi_2 - \phi_1 - \phi_2) + P_t q_t \phi_2 (\phi_1^2 + 4 \phi_1^2 + 5 \phi_2 \phi_1), B = q_t^4 \phi_2^2 [P_t q_t \phi_1 (J_1 \phi_1 + \phi_1 + 4 \phi_2) + (6 \phi_2 + 3 \phi_1 - P_t q_t \phi_2^2) J_2], H = P_t \phi_2 \phi_1 (2 \phi_2 + \phi_1)^2, E = q_t^3 \phi_2^2 (4 J_2 \phi_2 + J_2 \phi_1 + P_t q_t \phi_2 \phi_1), F = q_t \phi_2 [J_1 \phi_1^2 (2 P_t q_t \phi_2 - 1) + 2 P_t q_t \phi_2 (2 \phi_1 + 3 \phi_2) + J_2 \phi_2 (4 \phi_2 + 3 \phi_1 - 2 P_t q_t \phi_1)], \) and \( G = J_2 q_t \phi_3^3, \) with \( J_1 = J (1/[P_t (q_t - q_1) \phi_2]) \) and \( J_2 = J (1/[P_t q_1 \phi_1]) \). When \( q_1 \phi_1 = (q_t - q_1) \phi_2, q_1 = (q_t \phi_2)/(\phi_2 + \phi_1) \) and (4.37) can be simplified to

\[ \frac{d}{dq_1} I \left( \frac{q_t \phi_2}{\phi_2 + \phi_1}, \frac{q_t \phi_1}{\phi_2 + \phi_1} \right) = \frac{(\phi_2 - \phi_1)^2 K}{2 q_t^3 \phi_3^3 \phi_1 \phi_2^3 P_t^2}, \]  

(4.38)

where \( K = [\phi_1 + \phi_2]([\phi_1 + \phi_2] J_3 + 2 P_t q_t \phi_2 \phi_1 J_3 - P_t q_t \phi_2 \phi_1 - [P_t q_t \phi_2 \phi_1]^2 \) and \( J_3 = J \phi_1)/(P_t q_t \phi_2 \phi_1) \). The optimal power allocation can then be obtained by finding the point \( 0 \leq q_1 \leq q_t \) such that the derivative in (4.37) is equal to zero. Given that (4.37) is highly non-linear, finding a closed-form expression for the stationary point \( q_1^{\text{CI}} \text{DH} \)
is not straightforward. From the concavity of (4.36), the optimal power allocation can be obtained numerically by performing bisection on (4.37) for $0 \leq q_1 \leq q_t$.

Observe from (4.38) that $I([q_1, q_t - q_1])$ has a stationary point at $q_1 = (q_t \phi_2) / (\phi_2 + \phi_1)$ when $\phi_2 = \phi_1$, i.e., when the channel is symmetric. Since $I([q_1, q_t - q_1])$ is concave, $q_{CI, DH}^* = \frac{q_t}{2}$ must be the global maximizer and thus uniform power allocation is optimal for the symmetric network.

### 4.5 Rate Comparisons

In this section, we provide comparisons among different OW and TW relay protocols. We consider both the maximum achievable individual and sum rate in high and low power regions with the optimal power allocations as in (4.1) and (4.2).

#### 4.5.1 Individual Rate Comparison

First, we compare the maximum achievable rate in (4.1), which can be understood as the rate attained by an individual node using any of the considered protocols. Without loss of generality, we focus on the forward direction $S_1 \rightarrow S_2$, in which $S_1$ acts as the source and $S_2$ as the destination.

Consider first the systems in high power scenarios. Observe from Table 4.1 that due to the half-duplex constraint of the relay, the multiplexing gain of the DH, OAF and TW systems is $m = 1/2$, i.e., the achievable rate of these systems grows as $\frac{1}{2} \log(P_t)$ at high powers. On the other hand, the rate of the DT and NAF systems increases as $\log(P_t)$ with $m = 1$. Note also from (4.7), (4.19) and (4.23) that for a given power allocation $q$, $I_{OAF|h} > I_{DH|h} \geq I_{2,TW|h}$, where the equality is achieved when $q_2 = 0$ in the TW system, i.e., when the other user has zero rate. From the multiplexing gain and Section 4.3.2, we obtain the following relation at sufficiently high powers for either the CI or the FG technique:

$$C_{DT} = C_{NAF} > C_{OAF} > C_{DH} \geq C_{2,TW},$$

(4.39)

where $C_{NAF}$ is achieved when $q_{NAF} = [q_t/2, q_t/2, 0]$. Therefore, at high powers and under a sum power constraint, the DT scheme is optimal and AF relaying is suboptimal in terms of the single-user achievable rate.
In low power regions, it can be seen from Table 4.2 that the achievable rate of the DH and TW systems decreases quadratically with $n = 2$, rather than linearly with $P_t$ as the rate of the DT, OAF and NAF schemes. From this fact, Table 4.2 and Section 4.3.2, it is easy to show that at sufficiently low powers

$$C_{DT} = C_{NAF} > C_{OAF} > C_{DH} \geq C_{2,TW}. \quad (4.40)$$

Since $C_{NAF}$ is attained when the relay is inactive, relaying is also suboptimal at low power regions. Thus, AF relaying does not provide any advantage over the DT scheme at high and low power ranges in terms of the achievable rate. Intuitively, (4.39) and (4.40) indicate that as the total system power approaches a large or a small value, it is more beneficial to allocate all resources to the source to try to reach the destination directly, rather than to split the power between source and relay. This is asymptotically true regardless of the values $\phi_l > 0$ as long as they remain constant (i.e., do not asymptotically approach zero or infinity as the system power does).

### 4.5.2 Sum Rate Comparison

We now compare the relaying protocols in terms of the maximum sum rate in (4.2) achieved by nodes $S_1$ and $S_2$. It is worth mentioning that the OW relaying schemes considered in this chapter can be seen as TW protocols spanning over two cooperative frames. In the first frame, $S_1$ communicates with $S_2$ using a power allocation $q^{(1)} = [q_1^{(1)}, q_2^{(1)}, z_2^{(1)}]$, while in the second frame, $S_2$ communicates to $S_1$ with a power allocation $q^{(2)} = [q_1^{(2)}, q_2^{(2)}, z_2^{(2)}]$. Given that a frame for OW relaying occupies two symbol periods, this exchange lasts four time slots and a pre-log $1/2$ factor must be added to the sum rate of OW schemes, i.e.,

$$I_{sum} = \frac{1}{2}[I_{1,OW} + I_{2,OW}].$$

Note that $I_{i,OW}$ can be found in Tables 4.1 and 4.2 for a given OW protocol by replacing $\mathbf{q}$ with $\mathbf{q}^{(i)}$ and interchanging the roles of $\phi_1$ and $\phi_2$ for the second term in the sum rate. The power constraint in (4.2) must also be changed accordingly to

$$\sum_{i=1}^2 q_1^{(i)} + q_2^{(i)} + z_2^{(i)} \leq 2q_t$$

for OW schemes.

Consider first the high power regime. Given that $I_{i,OW}$ follows the same trends described in the previous subsection, it can be shown that the ordering in (4.39) still applies to the sum rate of OW protocols at high powers. For the TW scheme, the extra $1/2$ factor is not
Achievable Rate and Power Allocation of AF over Fading Channels

needed and \( I_{\text{sum}}^{TW} \) grows as \( \log(P_t) \). The following ordering then holds in high power regions

\[
\{ C_{\text{sum}}^{TW}, C_{\text{sum}}^{DT} = C_{\text{sum}}^{NAF} \} > C_{\text{sum}}^{OAF} > C_{\text{sum}}^{DH}, \tag{4.41}
\]

where \( C_{\text{sum}}^{DT} \) and \( C_{\text{sum}}^{NAF} \) are achieved when the relay is inactive with \( q_1^{(i)} = q_2^{(i)} = q_t/2 \). Different from the single-user discussions, we have observed that the TW system in (4.41) might outperform the DT scheme at high powers depending on the channel characteristics, i.e., the values of \( \phi_l \). As an example, consider the FG system in a symmetric configuration \( \phi_1 = \phi_2 = \phi \) using the suboptimal uniform power allocation \( q_1 = q_2 = z_2 = q_t/3 \). The sum rate of this system can be simplified from Table 4.1 at high powers as

\[
I_{\text{sum}}^{TW} \approx \log(P_t) + \log \left( \frac{q_t\phi}{6} \right) - \frac{2\gamma + J(2)}{\ln(2)}.
\]

As mentioned before, the sum rate of the DT scheme is maximized when \( q_1^{(i)} = q_2^{(i)} = q_t/2 \) and thus \( C_{\text{sum}}^{DT} \approx \log(P_t) + \log(\phi_0 q_t/2) - [\gamma/\ln(2)] \). For this configuration, the TW system outperforms the DT one whenever \( I_{\text{sum}}^{TW} > C_{\text{sum}}^{DT} \), i.e.,

\[
\phi/\phi_0 > 3 \cdot \exp[\gamma + J(2)] \approx 7.6688 \tag{4.42}
\]

The relay is then worth using when the \( S_i-R \) link is at least 8 times stronger than the direct \( S_1-S_2 \) link. Therefore, the TW system has the potential to outperform the DT scheme and is the only protocol that might mitigate the impact of half-duplex relaying in high power regions.

Consider now the low power scenario. Since the sum rate of the DT scheme decreases linearly with \( P_t \), while those of the DH and TW decrease quadratically, the following ordering holds:

\[
C_{\text{sum}}^{DT} = C_{\text{sum}}^{NAF} > C_{\text{sum}}^{OAF} > \{ C_{\text{sum}}^{TW}, C_{\text{sum}}^{DH} \}, \tag{4.43}
\]

where \( C_{\text{sum}}^{DT} \) and \( C_{\text{sum}}^{NAF} \) are achieved when the relay is inactive. Thus, similar to the single-user case, relaying is suboptimal at low powers in terms of the sum rate performance.
4 Achievable Rate and Power Allocation of AF over Fading Channels

4.6 Illustrative Examples

In this section, simulation results are provided to confirm the analysis in this chapter. For all simulations, we adopt the linear network model similar to Chapter 3, where the relay is in the line between the source and destination, the $S_1$-$S_2$ distance is normalized to 1, and the $S_1$-$R$ and $R$-$S_2$ distances are $d$ and $(1 - d)$, respectively, ($0 \leq d \leq 1$). In this model, $\phi_0 = 1$, $\phi_1 = 1/d^\nu$ and $\phi_2 = 1/(1 - d)^\nu$, where $\nu = 3$ is the pathloss exponent. For convenience, the results in this section are plotted against $P_t$, and $q_t = 2$ so that $P_t$ is the power constraint per symbol period. In addition, three types of power allocation schemes are considered: i) uniform power allocation, i.e., $q = [q_t/3, q_t/3, q_t/3]$ for NAF/TW, or $q = [q_t/2, 0, q_t/2]$ for DH/OAF; ii) the optimal power allocation that maximizes the achievable rate or sum rate as in (4.1) or (4.2) obtained using exhaustive search; and iii) the power allocation schemes proposed in Section 4.3.

4.6.1 Tightness of the Proposed Approximations

To verify the high power approximations derived in Table 4.1, Fig. 4.1 shows the single-user achievable rates of the DH, TW and NAF systems using uniform power allocation at $d = 0.5$. Without loss of generality, we consider the forward $S_1 \rightarrow S_2$ direction and the rate of the NAF system has been pre-multiplied by a factor of $1/2$ for clarity of the figure. Observe from Fig. 4.1 that all approximations coincide with the Monte Carlo simulations in medium to high powers. Although all rates in Fig. 4.1 have the same slope, i.e., the same multiplexing gain, it can be seen that the power gains vary widely among the considered protocols. It can also be observed from Fig. 4.1 that at sufficiently high powers, the CI technique is superior for DH and TW systems, while the FG one is better for the NAF protocol, which confirms the analysis in Section 4.2.

Fig. 4.2 shows the achievable rates of the systems using uniform power allocation along with the low power approximations in Table 4.2. Observe again from Fig. 4.2 that the simple approximations in Table 4.2 match with the simulation results at a sufficiently low powers for all systems. As discussed in Section 4.2, the OW FG systems are slightly better than the CI ones in low power regimes. However, it can be seen in Fig. 4.2 that the two amplification coefficients present an almost identical performance as $P_t$ decreases.

Besides the derived approximations, the closed-form expression of the achievable rate for the DH CI system in (4.36) is also shown in Figs. 4.1 and 4.2. Note from these figures
Fig. 4.1 Achievable rates and high power approximations of the systems using uniform power allocation \((d = 0.5)\).

Fig. 4.2 Achievable rates and low power approximations of the systems using uniform power allocation \((d = 0.5)\).
that the expression derived in Section 4.4 matches with the Monte Carlo simulations for all values of $P_t$, as expected.

### 4.6.2 Optimal Power Allocation

To quantify the advantage of the power allocation schemes proposed in Section 4.3, in this subsection we compare the achievable rates of the considered systems using different power allocation strategies. As an example, we set $d = 0.3$ in this subsection. Also, we shall use solid lines to represent the CI technique and dotted ones to represent the FG technique.

Fig. 4.3 shows the achievable rates of the DH system in high power regimes. Besides the proposed power allocation in (4.30), we also consider the uniform and the optimal allocation that achieves the capacity in (4.1), which was obtained using exhaustive search. For the CI system, we further consider the allocation obtained by applying bisection to the closed-form derivative in (4.37). Fig. 4.3 indicates that the systems using the proposed allocation schemes outperform those using uniform allocation. Specifically, asymptotic gains of 0.8 dB and 1.2 dB can be observed for the CI and FG systems, respectively. More importantly, the proposed power allocations present a negligible loss over the optimal one for both amplification techniques. The rate difference between (4.30) and the bisection method in (4.37) is also negligible for the CI system. As expected, it can be seen from Fig. 4.3 that the CI technique outperforms the FG one in high power regions.

The achievable rate of the DH system in low power regions is shown in Fig. 4.4. Recall from Section 4.3 that uniform power allocation $q_{1,DH} = z_{2,DH} = 1$ is asymptotically optimal in low powers. Thus, only the uniform, optimal (exhaustive) and bisection (for CI) power allocations are examined in Fig. 4.4. As expected, the uniform power allocation scheme presents a negligible loss over the optimal one at sufficiently low powers. Furthermore, the FG system presents a better rate than the CI one, and the rate difference decreases as $P_t$ decreases. This is in agreement with the discussions in Section 4.2.

The single-user achievable rates of the NAF and OAF systems are shown in Figs. 4.5 and 4.6 in high and low power regions, respectively. Recall from Section 4.3 that for the NAF system, $q_{NAF} = [1, 1, 0]$ is optimal in high and low power scenarios, i.e., DT is optimal. As such, it can be seen from these figures that the NAF systems using the proposed and the exhaustive power allocation present an indistinguishable performance and clearly outperform the OAF/NAF schemes with uniform power allocation. Note from
Fig. 4.3 Achievable rates of the DH systems using different power allocation schemes in high power regions ($d = 0.3$).

Fig. 4.4 Achievable rates of the DH systems using different power allocation schemes in low power regions ($d = 0.3$).
Fig. 4.5 that different from the DH systems in Fig. 4.3, the FG NAF system outperforms the CI one at high powers.

![Achievable rates of the NAF and OAF systems using different power allocation schemes in high power regions (\(d = 0.3\)).](image)

Fig. 4.5  Achievable rates of the NAF and OAF systems using different power allocation schemes in high power regions (\(d = 0.3\)).

The sum rate performance of the TW system with different power allocation strategies is shown in Fig. 4.7 for high power regions. Besides the uniform and optimal allocations, the schemes proposed in [91, 92] for the CI system are also considered in Fig. 4.7. Observe from Fig. 4.7 that in high power regions, the allocation proposed in (4.33) performs closely to the optimal one and provides significant gains over the other schemes. In particular, the CI system using the proposed power allocation presents a 1.3 dB gain over the system with uniform power allocation and a 0.5 dB gain over the systems using the schemes previously proposed in [91,92]. Using the proposed power allocation solution, a significant gain of 1.9 dB can also be achieved over the uniform power allocation scheme in FG systems. As previously noted, the CI systems in Fig. 4.7 outperform the FG ones in high power regions for a given power allocation strategy.

Fig. 4.8 shows the sum rates of the TW systems using the CI technique in low power regions. The rates of the FG systems are omitted in Fig. 4.8 for clarity of the figure. Observe from Fig. 4.8 that as \(P_t\) decreases, the performance of the system using the proposed
Fig. 4.6  Achievable rates of the NAF and OAF systems using different power allocation schemes in low power regions ($d = 0.3$).

Fig. 4.7  Achievable sum rates of the TW systems using different power allocation schemes in high power regions ($d = 0.3$).
asymptotic power allocation \( q_{\text{TW}} = [1/2, 1/2, 1] \) closely approaches that of the system using the power allocation from exhaustive search. These two power allocations present a 0.5 dB gain over the uniform scheme. The allocations from [91, 92] also present a comparable performance. Given that the rates of the CI and FG techniques converge at low powers, a similar behavior can be obtained for the FG systems.

Before closing this subsection, it should be mentioned that the results with different values of \( d \) have also been investigated. For the DH and TW operating in high powers, we observe that the proposed schemes provide smaller gains over uniform power allocation when the relay is close to \( d = 0.5 \), but larger gains as the relay gets away from the midpoint in any direction. A similar behavior is observed for the TW system operating in low power regimes. In the case of the NAF system, the gain of the proposed power allocation scheme over the uniform one appears to be roughly increasing as the relay gets closer to the destination.
4.6.3 Rate Comparisons

To verify the rate comparisons discussed in Section 4.5, Fig. 4.9 shows the achievable rates and sum rates of the considered systems at \( d = 0.5 \). Uniform power allocation is assumed for the DH system, which is optimal for the CI system with \( d = 0.5 \) at all values of \( P_t \) (see Section 4.4) and asymptotically optimal for the FG system in high/low power scenarios (see Section 4.3). For the TW system, we apply \( \mathbf{q}_{TW} = [1/2, 1/2, 1] \), which is also our proposed power allocation for \( d = 0.5 \) in high and low power regions. Note that for the OW protocols at \( d = 0.5 \), the rate is equal to the sum rate. The rates in Fig. 4.9 are normalized by the capacity of the DT scheme, which is also the capacity of the NAF protocol, i.e., \( C_{DT} = C_{NAF} = C_{sum}^{DT} = C_{sum}^{NAF} \). Note from Fig. 4.9 that the normalized single-user rates of the DH and TW approach \( 1/2 \) at high powers and thus are outperformed by the DT scheme, as expected from Table 4.1. In terms of the achievable sum rate, it can be observed in Fig. 4.9 that the TW scheme is superior to DT in high power regions and so approaches \( C_{sum}^{DT} \) from above. This is in agreement with (4.39), (4.41) and (4.42) since \( \phi_1 = \phi_2 = 8 \) when \( \nu = 3 \). In low power regions, the normalized rates and sum rates of the DH and TW approach zero since their rates decrease quadratically rather than linearly with \( P_t \) (see Table 4.2). As discussed in (4.40) and (4.43), the DT scheme outperforms all relaying protocols at sufficiently low powers. Thus, these results agree with our analysis in Section 4.5.

4.7 Chapter Summary

In this chapter, achievable rates and power allocation schemes were investigated for OW and TW single-relay AF systems over Rayleigh fading channels. Tight high and low power approximations were first derived for several relaying protocols. The achievable rates of the systems using the CI and FG coefficients were then compared. From the derived approximations, asymptotic power allocation strategies were also proposed for the protocols of interest. For the specific case of DH transmission with a CI coefficient, a closed-form expression of the rate was derived and bisection on the closed-form derivative was proposed to find the optimal power allocation. Finally, comparisons among all relaying protocols were presented.
Fig. 4.9 Achievable rates and sum rates of the DH and TW systems normalized by $C_{DT} = C_{sum}^{DT} (d = 0.5)$. 
Chapter 5

Relay Adaptation Policies for AF Relay Systems over Fading Channels

In Chapter 4, we investigated the capacity of several AF protocols over fading channels. Similar to previous works in the literature, we focused on the conventional CI and FG amplification coefficients to maintain a long-term average power constraint at the relay. Although these two coefficients are widely used in the literature, they are simply power normalization factors and are thus not optimal in any sense. For the single-input single-output system consisting of a source and a destination, source adaptation schemes to maximize the achievable rate under an average power constraint were studied by Goldsmith and Varaiya in their seminal paper in [43]. In [43], the source is assumed to have full CSI of the source-destination link and adapts its output power over time according to the channel conditions. Several adaptation techniques were proposed in [43] such as channel inversion and the optimal water-filling policy in (2.9). Similar to the idea in [43], relay adaptation strategies have also been proposed in [101,102] assuming that the relay can acquire knowledge of its incoming and outgoing links. Specifically, the output power at the relay in [101,102] varies according to the CSI to maximize a given objective function while maintaining a long-term average power constraint. However, the works in [101,102] were limited to the dual-hop scenario that ignores the source-destination link. For NAF relaying, an optimal amplification coefficient was recently obtained in [103] using the uncoded

\footnote{Parts of this chapter have been presented at the 2012 IEEE International Conference on Communications in Ottawa, Canada [98], and the 2013 IEEE Vehicular Technology Conference in Dresden, Germany [99]; and have been published in the IEEE Transactions on Vehicular Technology [100].}
PEP criterion. This amplification coefficient is, however, just numerically optimized for the QPSK constellation and only the CDI knowledge of the channel is exploited in [103]. Although power allocation strategies among the nodes in AF networks have been previously investigated as highlighted in the previous two chapters, relay adaptation strategies for cooperative or TW AF schemes have not been addressed in the literature.

In this chapter, we study relay adaptation policies for half-duplex cooperative and TW AF systems in which the relay can exploit the CSI by means of a suitable amplification coefficient to enhance the performance. Specifically, assuming full CSI at the relay and Gaussian codebooks at the source nodes, the optimal relay functions are derived under a long-term average power constraint. We consider the achievable rate and sum rate as the performance criteria for OW and TW relaying, respectively. At first, focusing on the OAF protocol, the maximization of the achievable rate is shown to be a concave optimization problem. The KKT conditions are then used to obtain the optimal relay function in closed-form. For the NAF scheme, it turns out that the maximization of the rate results in a non-concave problem. Therefore, we develop a method from which a global solution in closed-form can be obtained by using an achievable upper bound on the rate. Similar to the OAF case, the maximization of the sum rate for the TW system is also shown to be a concave problem. Applying the KKT conditions, it is then shown that the relay adaptation solution is the root of a quartic polynomial. The derived optimal solutions for OAF, NAF and TW relaying can be understood as a multi-D water-filling like scheme implemented over time and space. As such, insights on the optimal relay functions with respect to the channel gains are finally discussed. Numerical results indicate that the derived schemes provide a significant rate improvement over the FG and CI techniques, especially in low power regions, thanks to the benefit of dynamic power allocation.

The rest of the chapter is organized as follows. Section 5.1 first introduces the problem formulation. The optimal power adaptation schemes for OAF, NAF and TW relaying are then derived in Section 5.2. Section 5.3 provides important insights on the derived adaptation solutions. To quantify the gain of the proposed schemes, rate performances are given in Section 5.4. Finally, Section 5.5 concludes the chapter.
5.1 Problem Formulation

In this chapter, we consider again the fast fading scenario in which the transmitted codeword spans several channel realizations. In particular, we assume that the codeword is sent over $N$ consecutive transmission frames as described by the protocols in Section 2.5. The OW OAF, OW NAF, and two-phase TW protocols with input-output relations (2.23), (2.16) and (2.27), respectively, are considered. The channel gains at frame $i$ ($1 \leq i \leq N$), $\mathbf{h}_i = [h_0[i], h_1[i], h_2[i]]$, are assumed to be arbitrarily distributed and remain constant during at least one frame. For this faded scenario, perfect knowledge of the channel gains $\mathbf{h}_i$ is considered at all receivers. More importantly, the relay also has full knowledge of $\mathbf{h}_i$. This knowledge can be acquired via feedback from the destination in OW systems, or can be readily estimated at the relay in TW schemes.

Recall from Section 2.5 that two types of amplification coefficients are usually considered in literature, i.e., the FG and the CI coefficients with $\Xi = \phi$ and $\Xi = \alpha$ in (2.19) and (2.28). Given that the relay is assumed to possess full CSI in this chapter, it may use this knowledge to adapt the amplification coefficient according to the channel variations. In particular, let $z_2[i]P_r$ denote the instantaneous power allocated to the relay in frame $i$. The relay can then vary $z_2[i]$ according to $\mathbf{h}_i$ while maintaining a long-term average power constraint $\mathbb{E}[z_2[i]] \leq z_2$, i.e., the average power constraint is $z_2P_r$. In this case, the relay adaptation (RA) amplification coefficient for OW and TW systems can be respectively written as

$$b[i] = \sqrt{\frac{z_2[i]}{q_1P_s\alpha_1[i] + N_r}},$$

and

$$b[i] = \sqrt{\frac{z_2[i]}{q_1P_s\alpha_1[i] + q_2P_{s_2}\alpha_2[i] + N_r}},$$

where $\alpha_l[i] = |h_l[i]|^2$. Note that the CI and FG coefficients in (2.19) and (2.28) can be considered as a special case of the general adaptation method in (5.1) and (5.2). For instance, $z_2[i] = z_2$ for CI, whereas $z_2[i] = z_2[q_1P_s\alpha_1[i] + N_r]/[q_1P_s\phi_1 + N_r]$ for OW FG and $z_2[i] = z_2[q_1P_s\alpha_1[i] + q_2P_{s_2}\alpha_2[i] + N_r]/[q_1P_s\phi_1 + q_2P_{s_2}\phi_2 + N_r]$ for TW FG. Note also that for the TW system, the CI and RA coefficients have the same CSI requirements.
In general, the relay might adapt the amplification coefficients in (5.1) and (5.2) in order to optimize a given objective function. Among different criteria, here we consider the achievable rate or sum rate for OW and TW systems, respectively, as the performance metric similar to Chapters 3 and 4. Certainly, one can use other design criteria to develop adaptation techniques following a similar approach. As can be seen from (5.1) and (5.2), the problem of finding the optimal relay adaptation scheme in terms of $h_i$, i.e., the optimal relay function $b[i]$, is equivalent to finding the optimal instantaneous power allocated to the relay at a given frame $z_2[i]$. Assuming Gaussian codebooks and that the transmitted codewords span over $N$ frames, the optimal $z_2[i]$ can then be obtained by maximizing the achievable rate or sum rate while maintaining the average power constraint. This can be modeled as a parallel channel and the optimization problem for OW systems can be written as

$$\max_{z_2[i] \geq 0} \left\{ \frac{1}{N} \sum_{i=1}^{N} I|h_i \right\} \text{ s.t. } \frac{1}{N} \sum_{i=1}^{N} z_2[i] \leq z_2, \quad (5.3)$$

where the conditional mutual information $I|h_i$ is given by either (2.20) or (2.24) depending on the OW protocol under consideration. Similarly, the optimization problem for TW relaying is given by

$$\max_{z_2[i] \geq 0} \left\{ \frac{1}{N} \sum_{i=1}^{N} [I_1|h_i + I_2|h_i] \right\} \text{ s.t. } \frac{1}{N} \sum_{i=1}^{N} z_2[i] \leq z_2, \quad (5.4)$$

with $I_1|h_i$ as in (2.29). The systems using the conventional FG and CI coefficients can be used as benchmarks to assess the power adaptation schemes from (5.3) and (5.4).

Before closing this section, it should be noted that the framework described here is general, in the sense that the average power allocation scheme among the nodes is left as the parameters $q = [q_1, q_2, z_2]$. In the following, the main objective is to derive an optimal power adaptation scheme by means of the amplification coefficient, which can be applied to any such parameters.
5.2 Optimal Power Adaptation Schemes

In this section, we shall develop optimal power adaptation schemes by solving the problems in (5.3) and (5.4). Let \( z_2[i] = z_{2,i} \) and denote the optimum value of \( z_{2,i} \) as \( z^*_{2,i} \). We first address the adaptation problem for the OAF system, followed by that of the NAF and TW schemes.

5.2.1 OAF System

Consider first the OAF system in which the source has a power allocation of \( q_1 > 0 \), whereas an average power constraint of \( z_2 > 0 \) is imposed on the relay. Recall from Section 2.5 that for the OAF scheme, the source is silent in the cooperative phase with \( q_2 = 0 \). Substituting the RA coefficient in (5.1), the conditional achievable rate in (2.24) can be simplified to

\[
I_{OAF}|h_i = \frac{1}{2} \log \left( 1 + q_1 \gamma_{0,i} + \frac{q_1 z_{2,i} \gamma_{1,i} \gamma_{2,i}}{q_1 \gamma_{1,i} + z_{2,i} \gamma_{1,i} + 1} \right) = \frac{1}{2} \log \left[ 1 + f_{OAF}(z_{2,i}, h_i) \right],
\]  

(5.5)

where \( \gamma_{0,i} = (P_s \alpha_0[i]) / N_d, \gamma_{1,i} = (P_s \alpha_1[i]) / N_r \) and \( \gamma_{2,i} = (P_r \alpha_2[i]) / N_d \). The optimal instantaneous power allocated to the relay for the OAF system can then be obtained by solving (5.3) with \( I|h_i \) as in (5.5). With a slight abuse of notation, let \( I|h_i = I(z_{2,i}|h_i) \). The first and second derivatives of (5.5) are given respectively as

\[
\frac{\partial I_{OAF}(z_{2,i}|h_i)}{\partial z_{2,i}} = \frac{a_i}{2 \ln(2)c_i z_{2,i}^2 + d_i z_{2,i} + e_i},
\]

(5.6)

\[
\frac{\partial^2 I_{OAF}(z_{2,i}|h_i)}{\partial z_{2,i}^2} = -\frac{a_i[2c_i z_{2,i} + d_i]}{2 \ln(2)[c_i z_{2,i}^2 + d_i z_{2,i} + e_i]^2},
\]

(5.7)

where

\[
a_i = q_1 \gamma_{1,i} \gamma_{2,i} (q_1 \gamma_{1,i} + 1),
\]

\[
c_i = \gamma_{2,i}^2 (q_1 \gamma_{0,i} + q_1 \gamma_{1,i} + 1),
\]

\[
d_i = \gamma_{2,i} (q_1 \gamma_{0,i} + 1) (2q_1 \gamma_{0,i} + q_1 \gamma_{1,i} + 2),
\]

\[
e_i = (q_1 \gamma_{0,i} + 1) (q_1 \gamma_{1,i} + 1)^2.
\]

(5.8)
Note that since $\gamma_0, i, \gamma_1, i, \gamma_2, i \geq 0$ and $q_1, z_2, i > 0$, the second derivative in (5.7) is non-positive and independent of $z_{2,k}, \forall k \neq i$. The Hessian of (5.5) is then diagonal with the non-positive elements in (5.7). Given that the constraints in (5.3) are linear, this is a concave optimization problem. The optimal value of $z_{2,i}$ can then be obtained using the KKT conditions by simply finding a stationary point in the Lagrangian [104]. In particular, the Lagrangian of (5.3) can be written as

$$L(z_2, \lambda_1) = -\frac{1}{N} \sum_{i=1}^{N} I_{OAF}(z_{2,i}|h_i) - \lambda_1 \left(z_2 - \frac{1}{N} \sum_{i=1}^{N} z_{2,i}\right), \quad (5.9)$$

where $z_2 = \{z_{2,i} | 1 \leq i \leq N\}$ and $\lambda_1 \geq 0$ is the Lagrange multiplier. The $i$-th component of the gradient of (5.9), $\nabla L(z_2, \lambda_1)$, is given by

$$\frac{\partial L(z_2, \lambda_1)}{\partial z_{2,i}} = -\frac{1}{N} \frac{\partial I_{OAF}(z_{2,i}|h_i)}{\partial z_{2,i}} + \frac{\lambda_1}{N},$$

with the derivative as in (5.6). By equating the gradient to zero, i.e., $\nabla L(z_2, \lambda_1) = 0$, and solving for $z_{2,i}$, the optimal instantaneous allocation for the OAF scheme can be expressed as

$$z_{OAF}^{2,i} = \begin{cases} \left[\frac{-d_i + \sqrt{d_i^2 - 4c_i(e_i - a_i \mu_1^*)}}{2c_i}\right]^+, & \gamma_{1,i}, \gamma_{2,i} > 0 \\ 0, & \text{o.w.}, \end{cases} \quad (5.10)$$

where $[x]^+ = \max\{0, x\}, \mu_1 = 1/[2 \ln(2)\lambda_1]$ and $d_i^2 - 4c_i(e_i - a_i \mu_1) = q_1 \gamma_{1,i} \gamma_{2,i}^2 (q_1 \gamma_{1,i} + 1) [q_1 \gamma_{1,i} (q_1 \gamma_{1,i} + 1) + 4\mu_1 \gamma_{2,i} (q_1 \gamma_{0,i} + q_1 \gamma_{1,i} + 1)] > 0$, so that the root in (5.10) is always real. Note that the second root has been ignored since it violates the non-negativity of $z_{2,i}$.

In (5.10), $\lambda_1^* > 0$ is a unique constant that must satisfy

$$g(\lambda_1^*) = \frac{1}{N} \sum_{i=1}^{N} z_{2,i}^* = z_2, \quad (5.11)$$

and can be easily found using bisection. This is because the root in (5.10) is strictly
decreasing with $\lambda_1 > 0$, is negative when $\lambda_1 \rightarrow +\infty$, and approaches $+\infty$ as $\lambda_1$ approaches zero from above $\lambda_1 \rightarrow 0^+$. As such, $\lim_{\lambda_1 \rightarrow 0^+} g(\lambda_1) = +\infty$, $\lim_{\lambda_1 \rightarrow +\infty} g(\lambda_1) = 0$ and $g(\lambda_1)$ is strictly decreasing with $\lambda_1 > 0$. Note also that the sum power constraint in (5.11) is tight as the objective function in (5.5) is strictly increasing with $z_{2,i}$ as long as $\exists i \in \{1, \ldots, N\}$ with $\gamma_{1,i}, \gamma_{2,i} > 0$, i.e., the first derivative in (5.6) is strictly positive.

Observe from (5.3) that the solution in (5.10) can be applied over any $N$ parallel channels as long as the relay has knowledge of $h_i$ for $1 \leq i \leq N$. However, as explained in Section 2.3, when $N$ large enough to capture the ergodicity of the channel $h_i$,

$$\frac{1}{N} \sum_{i=1}^{N} I|h_i| \rightarrow \mathbb{E}[I|h_i|] = I,$$

and

$$\frac{1}{N} \sum_{i=1}^{N} z_{2,i}^* \rightarrow \mathbb{E}[z_{2,i}^*] = z_2,$$

so that $(z_{2,i}^*, \lambda_1^*)$ become a causal function of $h_i$ and its distribution, rather than a non-causal function of future values of $h_i$. It can be seen from (5.10) that $\lambda_1^*$ also depends on the noise variances, and the power allocations at the source and relay.

Finally, as a consistency check, one can verify that the solution in (5.10) reduces to the relay adaptation scheme proposed in [101] for the special case of $\gamma_{0,i} = 0$, i.e., to maximize the achievable rate of the DHAF system.

### 5.2.2 NAF System

We now turn our attention to NAF system in which the source is allowed to transmit in both broadcasting and cooperative phases. As such, it can be assumed that $q_1$, $q_2$ and $z_2$ are strictly positive. Recall from Section 2.5 that the covariance matrix of the input vector to the NAF channel is given by $Q = \mathbb{E}[xx^\dagger]$ and has diagonal elements $q_1 > 0$ and $q_2 > 0$. As shown in [22], when the source uses a Gaussian codebook and has no channel knowledge, correlation in the input vector reduces the average mutual information of the NAF system. Hence, similar to Chapter 4, we consider a diagonal input covariance matrix $Q$. Substituting the RA coefficient in (5.1) into (2.20) with $q_{12} = 0$, the conditional mutual
information can be written as

\[ I_{\text{NAF}}|h_i = \frac{1}{2} \log[1 + f_{\text{NAF}}(z_{2,i}, h_i)], \]

where

\[ f_{\text{NAF}}(z_{2,i}, h_i) = f_{1,\text{NAF}}(z_{2,i}, h_i)/f_{2,\text{NAF}}(z_{2,i}, h_i) \]

with

\[ f_{1,\text{NAF}}(z_{2,i}, h_i) = q_1^2 q_2^2 \gamma_{0,i}^2 \gamma_{1,i} + q_1 (2q_2^2 \gamma_{0,i}^2 + q_1 \gamma_{0,i} \gamma_{1,i} + z_{2,i} \gamma_{0,i} \gamma_{2,i} + z_{2,i} \gamma_{1,i} \gamma_{2,i} + q_2 \gamma_{0,i} \gamma_{1,i}) \]

\[ + \gamma_{0,i}(q_1 + q_2), \]

\[ f_{2,\text{NAF}}(z_{2,i}, h_i) = q_1 \gamma_{1,i} + z_{2,i} \gamma_{2,i} + 1. \]

The optimal instantaneous power allocation at the relay can be obtained as in the previous subsection by solving (5.3) with \( I|h_i \) replaced by (5.12). It can be easily shown that

\[ \frac{\partial I_{\text{NAF}}(z_{2,i}|h_i)}{\partial z_{2,i}} = -A_i \]

\[ \frac{\partial^2 I_{\text{NAF}}(z_{2,i}|h_i)}{\partial z_{2,i}^2} = \frac{-A_i}{2 \ln(2)[C_i z_{2,i}^2 + D_i z_{2,i} + E_i]^2}, \]

where

\[ A_i = \gamma_{2,i} (q_1 \gamma_{1,i} + 1) (q_1 q_2^2 \gamma_{0,i}^2 + q_2 \gamma_{0,i} - q_1 \gamma_{1,i}), \]

\[ C_i = \gamma_{2,i} (q_1 \gamma_{0,i} + q_1 \gamma_{1,i} + 1), \]

\[ D_i = \gamma_{2,i} (q_1 \gamma_{1,i} + 1) (\gamma_{0,i} [2q_1 + q_2 + q_1 q_2 \gamma_{0,i}] + q_1 \gamma_{1,i} + 2), \]

\[ E_i = (q_1 \gamma_{0,i} + 1) (q_2 \gamma_{0,i} + 1) (q_1 \gamma_{1,i} + 1)^2. \]

Note that since \( \gamma_{0,i}, \gamma_{1,i}, \gamma_{2,i} \geq 0 \) and \( q_1, q_2, z_{2,i} > 0 \), the parameters \( E_i > 0 \) and \( C_i, D_i \geq 0 \). However, \( A_i \) might be positive, negative or zero depending on the noise variances, power allocations at \( S \), and instantaneous channel gains. Hence, different from the OAF case, the optimization problem in (5.3) for the NAF system is not concave. In the following, we first develop an achievable upper bound on the objective function (5.12) so that a concave-optimization problem can be established. The global maximizer of (5.3) is finally obtained in closed-form.

First, observe from (5.13) that \( I_{\text{NAF}}(z_{2,i}|h_i) \) can be either a strictly increasing function of \( z_{2,i} \) when \( A_i < 0 \), a strictly decreasing function when \( A_i > 0 \), or a constant with respect
to $z_{2,i}$ when $A_i = 0$. Define the sets $\text{AP} = \{i \mid A_i \geq 0\}$ and $\text{AN} = \{i \mid A_i < 0\}$. Furthermore, let $\mathcal{R} = \{z_2 \in \mathbb{R}^N \mid z_{2,i} \geq 0, \frac{1}{N} \sum_{i=1}^{N} z_{2,i} \leq z_2\}$ be the feasible region. It is then straightforward to show that for any $z_2 \in \mathcal{R}$, the objective function in (5.3) can be upper bounded as

$$
\sum_{i=1}^{N} I_{\text{NAF}}(z_{2,i}|h_i) \leq \sum_{i \in \text{AN}} I_{\text{NAF}}(z_{2,i}|h_i) \leq \sum_{i \in \text{AN}} I_{\text{NAF}}(z^*_{2,i}|h_i).
$$

In (5.16), the first upper bound is due to the fact that the objective function decreases in value with $z_{2,i} > 0$ for $i \in \text{AP}$ and is achieved with equality when $z_{2,i} = 0 \ \forall \ i \in \text{AP}$. The second inequality is the best feasible upper bound with $z_{2,i} = 0 \ \forall \ i \in \text{AP}$ and the values of $z^*_{2,i}$ for $i \in \text{AN}$ are the solution to:

$$
\max_{z_{2,i} \geq 0, i \in \text{AN}} \frac{1}{N} \sum_{i \in \text{AN}} I_{\text{NAF}}(z_{2,i}|h_i) \ \text{s.t.} \ \frac{1}{N} \sum_{i \in \text{AN}} z_{2,i} \leq z_2.
$$

Note that the above objective function is now concave since its Hessian is diagonal with the strictly negative elements in (5.14) for $A_i < 0$. Hence, due to the linearity of the constrains and the achievability of the upper bound in (5.16), any solution to (5.3) in the form of $z_{2,i} = 0 \ \forall \ i \in \text{AP}$ and $z_{2,i} = z^*_{2,i} \ \forall \ i \in \text{AN}$ is globally optimal. The problem in (5.17) can then be solved using the KKT conditions by finding a stationary point in the Lagrangian. The Lagrangian of (5.17) can be written as

$$
\mathcal{L}(z_2, \lambda_2) = -\frac{1}{N} \sum_{i \in \text{AN}} I_{\text{NAF}}(z_{2,i}|h_i) - \lambda_2 \left( z_2 - \frac{1}{N} \sum_{i \in \text{AN}} z_{2,i} \right).
$$

By equating the gradient of the above Lagrangian to zero and solving for $z_{2,i}$, the optimal instantaneous allocation $z^*_{2,i}$ for the NAF protocol and the solution to (5.3) is finally expressed as

$$
z^*_{2,i,\text{NAF}} = \begin{cases} 
\left[ -D_i + \sqrt{D_i^2 - 4C_i(E_i + A_i \mu^2)} \right]^{+} / 2C_i, & A_i < 0 \ (i \in \text{AN}) \\
0, & A_i \geq 0 \ (i \in \text{AP}),
\end{cases}
$$

where we have ignored the second root since it violates the non-negativity of $z_{2,i}$. In (5.18),
\[ \mu_2 = 1/[2 \ln(2) \lambda_2] \] and \[ D_i^2 - 4C_i(E_i + A_i \mu_i^*) > 0 \] for \( A_i < 0 \) so that \( z_{2,i}^* \) is always real. Furthermore, \( \lambda_*^2 > 0 \) is a unique constant which satisfies (5.11) and can be easily found using bisection as long as the set \( AN \neq \emptyset \). This is because, similar to (5.10), the root (5.18) is strictly decreasing with \( \lambda_2 > 0 \), is negative when \( \lambda_2 \to +\infty \), and approaches \( +\infty \) as \( \lambda_2 \to 0^+ \). The sum power constraint in (5.11) is again tight as the objective function in (5.17) is strictly increasing with \( z_{2,i}^* \) for \( i \in AN \).

### 5.2.3 TW System

Consider now the two-phase TW system in which both source nodes want to exchange information with \( q_1, q_2, z_2 > 0 \). For the system using the RA coefficient in (5.2), the conditional achievable rate in the \( S_k \to S_j \) direction (\( j \in \{1, 2\}, k = 3 - j \)) can be expressed from (2.29) as

\[
I_{j,\text{TW}}|h_i = \frac{1}{2} \log \left( 1 + \frac{q_kz_{2,i}^2 \gamma_1,i \gamma_2,i}{q_1 \gamma_1,i + q_2 \gamma_2,i + z_{2,i} \gamma_j,i + 1} \right), \tag{5.19}
\]

where for notational convenience we assume that \( P_{sj} = P_r \) and \( N_{dj} = N_r \) so that \( \gamma_{j,i} = (P_{sj} \alpha_j[i])/N_r = (P_r \alpha_j[i])/N_{dj} \). The conditional sum rate is then given by \( I_{\text{sum}}|h_i = I_1|h_i + I_2|h_i \). The optimal power allocation that maximizes the sum rate can be obtained by solving (5.4) using \( I_j|h_i \) in (5.19). The first and second derivatives of (5.19) are given by

\[
\frac{\partial I_{j,\text{TW}}(z_{2,i}|h_i)}{\partial z_{2,i}} = \frac{A_j'}{2 \ln(2) P_j(z_{2,i})},
\]

\[
\frac{\partial^2 I_{j,\text{TW}}(z_{2,i}|h_i)}{\partial^2 z_{2,i}} = -\frac{A_j'[2B_j'z_{2,i} + D_j']}{2 \ln(2) P_j^2(z_{2,i})}, \tag{5.20}
\]

where the second order polynomial

\[
P_j(z_{2,i}) = B_j'z_{2,i}^2 + D_j'z_{2,i} + E', \tag{5.21}
\]
with

\[ A'_j = q_k \gamma_{1,i} \gamma_{2,i} (q_1 \gamma_{1,i} + q_2 \gamma_{2,i} + 1), \]
\[ B'_j = \gamma_{j,i}^2 (q_k \gamma_{k,i} + 1), \]
\[ D'_j = \gamma_{j,i} (q_k \gamma_{k,i} + 2) (q_1 \gamma_{1,i} + q_2 \gamma_{2,i} + 1), \]
\[ E' = (q_1 \gamma_{1,i} + q_2 \gamma_{2,i} + 1)^2. \]

Given that \( q_1, q_2, z_2 > 0 \) and \( \gamma_{1,i}, \gamma_{2,i} \geq 0 \), the second derivative in (5.20) is non-positive. Similar to the OAF case, the problem in (5.4) is concave for the TW system and the optimal \( z_{2,i}^* \) can be obtained by finding a stationary point in the Lagrangian. The Lagrangian of (5.4) can be written as

\[ \mathcal{L}(z_2, \lambda_3) = -\frac{1}{N} \sum_{i=1}^{N} \left[ I_{1,TW}(z_{2,i} | h_i) + I_{2,TW}(z_{2,i} | h_i) \right] - \lambda_3 \left( z_2 - \frac{1}{N} \sum_{i=1}^{N} z_{2,i} \right), \]

and its gradient is given by

\[ \frac{\partial \mathcal{L}(z_2, \lambda_3)}{\partial z_{2,i}} = -\frac{1}{N} \left[ \frac{\partial I_{1,TW}(z_{2,i} | h_i)}{\partial z_{2,i}} + \frac{\partial I_{2,TW}(z_{2,i} | h_i)}{\partial z_{2,i}} \right] + \frac{\lambda_3}{N}, \]

with the derivatives as in (5.20). By equating the above derivative to zero, it can be shown that the stationary point of the Lagrangian must satisfy the following condition:

\[ P(z_{2,i}, \mu_3) = \frac{P_1(z_{2,i})P_2(z_{2,i}) - \mu_3 A'_1P_2(z_{2,i}) - \mu_3 A'_2P_1(z_{2,i})}{N} = 0, \]

for \( 1 \leq i \leq N \), where \( \mu_3 = 1/[2 \ln(2) \lambda_3] \) and \( P_j(\cdot) \) is given by (5.21). The above function is a quartic polynomial on \( z_{2,i} \) and has at most four real roots. Since (5.4) is a concave problem, at most one of these roots is real and positive.

Rather than writing the cumbersome equation for the root of a quartic polynomial, let \( r_i(\mu_3) \) be the largest real root of (5.24) at frame \( i \) for a given \( \mu_3 \), i.e., \( P(r_i(\mu_3), \mu_3) = 0 \). It is easy to show that when \( \gamma_{1,i} = 0 \) or \( \gamma_{2,i} = 0 \), \( P(z_{2,i}, \mu_3) \) becomes a quadratic with a single real negative root of multiplicity two. From (5.24) and the concavity of (5.4), the optimal
power allocation and the solution to (5.4) can then be expressed as

\[
    z_{2,i}^{TW} = \begin{cases} 
        [r_i(\mu^*_3)]^+, & \gamma_{1,i}, \gamma_{2,i} > 0 \\
        0, & \text{o.w.,} 
    \end{cases} \tag{5.25}
\]

As shown in Appendix C.1, \( \lim_{\lambda_3 \to +\infty} r_i(\mu_3) < 0 \), \( \lim_{\lambda_3 \to 0^+} r_i(\mu_3) = +\infty \) and \( r_i(\mu_3) \) is strictly decreasing with \( \lambda_3 > 0 \). Thus, as in the NAF and OAF cases, \( \lambda^*_3 > 0 \) in (5.25) can be found using bisection. The sum power constraint in (5.4) is again tight as the first derivative in (5.20) is strictly positive for \( \gamma_{1,i}, \gamma_{2,i} > 0 \). As before, when \( N \) is sufficiently large, we have that \( [1/N]\sum_{i=1}^{N} [I_1|h_i| + I_2|h_i|] \to \mathbb{E}[I_1|h_i| + I_2|h_i|] = I_1 + I_2 \). Furthermore, similar to the OAF system, we have verified that the solution in (5.25) reduces to the one in [101] when \( q_1 = 0 \) or \( q_2 = 0 \).

### 5.3 Discussions on Optimal Solutions

In this section, we provide further insights regarding the derived power adaptation schemes. It can be seen from the previous section that the results in (5.10), (5.18) and (5.25) are water-filling like schemes. Different from the water-filling formula for a single-link in (2.9) that simply allocates more power to strong channels, and less or no power to weak channels, the solutions in (5.10), (5.18) and (5.25) depend on multiple variables. These include the fading statistics of three time-varying channels, the noise variances, as well as the power allocations at the source. As such, the proposed solutions can be considered as a multi-D water-filling like scheme in time and space. In the following, we discuss some intuitive behavior of the derived solutions with respect to the channel gains.

#### 5.3.1 Zero and Non-zero Power Regions

First, we investigate the conditions under which a zero or a non-zero power is allocated to the relay at a given frame \( i \), i.e., whether the relay is active with \( z^*_{2,i} > 0 \) or inactive with \( z^*_{2,i} = 0 \) at any given frame. Consider the OAF protocol using the power adaptation scheme in (5.10). Observe from (5.10) that the relay for the OAF scheme remains silent at a given frame whenever \( e_i - a_i\mu^*_1 \geq 0 \), or equivalently,

\[
    \frac{q_1\gamma_{1,i}\gamma_{2,i}}{(q_1\gamma_{0,i} + 1)(q_1\gamma_{1,i} + 1)} \leq 1/\mu^*_1. \tag{5.26}
\]
It is easy to see from (5.26) that the relay is not used whenever $\gamma_{2,i}$ is too small, or when $\gamma_{1,i}$ is small and $\gamma_{2,i}$ is not large. The relay is also not used when $\gamma_{0,i}$ is sufficiently large compared to $\gamma_{2,i}$. Conversely, the relay is active when $\gamma_{2,i}$ is large compared to $\gamma_{0,i}$ as long as $\gamma_{1,i}$ is not small. Hence, the power adaptation for the OAF scheme in (5.10) can be intuitively interpreted as follows:

- The relay is inactive when either the $R-D$ link is weak; the $S-R$ is weak and the $R-D$ link is not strong; or the direct $S-D$ is strong compared to the $R-D$ link.
- The relay is active when the $R-D$ link is stronger than the direct $S-D$ link and the $S-R$ link is not weak.

To illustrate the above properties, Fig. 5.1 shows the active/inactive regions of the OAF protocol for different values of $\gamma_{0,i}$ ($\mu_1 = q_1 = 1$). The inactive or zero-power region is the lower-left area to the boundaries in Fig. 5.1, whereas the active or non-zero-power region is the upper-right area. As expected, it can be seen from Fig. 5.1 that no power is allocated to the relay when $\gamma_{1,i}$ and $\gamma_{2,i}$ are small compared to $\gamma_{0,i}$. On the other hand, the relay is active when $\gamma_{1,i}$ and $\gamma_{2,i}$ are sufficiently large. It can also be seen from Fig. 5.1 that the zero-power region expands as $\gamma_{0,i}$ increases.

For the NAF protocol, the behavior of the optimal power adaptation scheme in (5.18) is slightly different. In particular, it can be seen from (5.18) that the relay is off when either i) $A_i \geq 0$; or ii) $A_i < 0$ and $(E_i + A_i\mu_2^*) \leq 0$. When $A_i \geq 0$, it then follows from (5.15) that the relay shall keep silent whenever $(q_1 q_2 \gamma_{0,i}^2 + q_2 \gamma_{0,i} - q_1 \gamma_{1,i}) \geq 0$. For instance, when power is poured equally in two cooperative phases ($q_1 = q_2$), the relay remains off whenever $\gamma_{0,i} \geq \gamma_{1,i}$. On the other hand, when $A_i < 0$, it can be verified that the relay should not be used if

$$\frac{\gamma_{2,i} |q_1 \gamma_{1,i} - q_1 q_2 \gamma_{0,i}^2 - q_2 \gamma_{0,i}|}{(q_1 \gamma_{0,i} + 1) (q_2 \gamma_{0,i} + 1) (q_1 \gamma_{1,i} + 1)} \leq 1/\mu_2^*.$$  

(5.27)

It follows from (5.27) that the relay remains silent if $\gamma_{2,i}$ is small. On the contrary, the relay is active when $\gamma_{1,i}$ and $\gamma_{2,i}$ are large compared to $\gamma_{0,i}$. Overall, the multi-D water-filling formula for the NAF protocol in (5.18) can be interpreted as follows:

- The relay is inactive when either the $S-D$ link is stronger than the $S-R$ link; or when one has a weak $R-D$ channel.
Fig. 5.1 Zero and non-zero power regions for the OAF protocol ($\mu_1 = 1, q_1 = 1$).

- Power is allocated to the relay when $S-R$ and $R-D$ links are stronger than the $S-D$ one.

The above trends are illustrated in Fig. 5.2 for the NAF protocol with $\mu_2 = q_1 = q_2 = 1$. It can be seen from this figure that the relay is active/inactive when the $S-R$ and $R-D$ links are large/small compared to the direct one. Observe also that the shape of the boundaries in Figs. 5.1 and 5.2 differ greatly. Furthermore, compared to the behavior in Fig. 5.1, the zero-power region in Fig. 5.2 expands much faster as $\gamma_{0,i}$ increases.

Finally, for the TW system, the solution in (5.25) is more challenging to analyze due to the cumbersome closed-form expression for the root of a quartic polynomial. By looking into $P(z_{2,i}, \mu_3)$ in (5.24) in more detail, we show in Appendix C.2 that $r_i(\mu_{3,i}^* < 0$ when either $\gamma_{1,i}$ or $\gamma_{2,i}$ is small, and $r_i(\mu_{3,i}^* > 0$ when both $\gamma_{1,i}$ and $\gamma_{2,i}$ are large. Thus, the solution in (5.25) can be interpreted as follows:

- The relay is inactive when either the $S_1-R$ or the $S_2-R$ link is sufficiently small.
- The relay is active when both the $S_1-R$ and the $S_2-R$ gains are sufficiently large.
The above observations can be confirmed in Fig. 5.3, where (5.25) is illustrated for $\mu_3 = q_1 = q_2 = 1$. Different from the OAF and NAF systems, note that the boundary in Fig. 5.3 is symmetric along the line $\gamma_{1,i} = \gamma_{2,i}$. This is due to the symmetry of the mutual information in (5.19).

### 5.3.2 High and Low Power Regions

Assuming that the relay is in the active region, here we investigate the areas in which more or less power is allocated to the relay, i.e., the high and low power regions. Consider first the cooperative relay systems. From the previous discussion, we can guarantee that the relay is active for the OAF and NAF protocols when the overall S-R-D link is much stronger than the direct S-D link. Thus, when $\gamma_{1,i}, \gamma_{2,i} \gg \gamma_{0,i}$, the relay is active and the power adaptation schemes in (5.10) and (5.18) can both be approximated as

$$z^*_2,i \approx -\frac{q_1}{2} \left( \frac{\gamma_{1,i}}{\gamma_{2,i}} \right) + \sqrt{\frac{q_1^2}{4} \left( \frac{\gamma_{1,i}}{\gamma_{2,i}} \right)^2 + q_1 \mu^* \left( \frac{\gamma_{1,i}}{\gamma_{2,i}} \right)}.$$
Fig. 5.3 Zero and non-zero power regions for the TW protocol ($\mu_3 = 1, q_1 = q_2 = 1$).

Since the above power allocation is a strictly increasing function of $\gamma_{1,i}/\gamma_{2,i}$, more power is given to the relay when $\gamma_{1,i} > \gamma_{2,i}$, as long as the overall $S$-$R$-$D$ link remains stronger than the $S$-$D$ link. Conversely, less power is allocated when $\gamma_{2,i} > \gamma_{1,i}$. It is then intuitive that if the $S$-$R$ link is much stronger than the $R$-$D$ one, $R$ has a high confidence about the received symbol and hence allocates more power to guarantee correct reception at $D$. Equivalently, if the $S$-$R$ link is much weaker than the $R$-$D$ one, $R$ is not confident about the received symbol. As a result, it shall allocate less power for the re-transmission. To illustrate this behavior, Fig. 5.4 shows the contour of (5.18) as a function of the channel gains $\gamma_{1,i}$ and $\gamma_{2,i}$ for $\gamma_{0,i} = \mu_2 = q_1 = q_2 = 1$. When the relay is active, it can be seen from Fig. 5.4 that it is allocated more power if $\gamma_{1,i} > \gamma_{2,i}$, which can be understood as the high power region. Equivalently, less power is allocated when $\gamma_{2,i} > \gamma_{1,i}$ in the low power region. Although Fig. 5.4 only shows the contour for the NAF adaptation scheme, similar trends have been observed for the OAF scheme in (5.10).

Consider now the case of the TW protocol. As discussed before, the relay is active for this protocol when the $S_1$-$R$ and the $S_2$-$R$ links are sufficiently large. Assuming that $\gamma_{1,i} \gg 1$ and $\gamma_{2,i} \gg 1$, we show in Appendix C.2 that $r_i(\mu_3)$ increases as $\gamma_{2,i}/\gamma_{1,i}$ moves away
Fig. 5.4 Optimal adaptation scheme $z_{2,i}^{NAF}$ in (5.18) for $\gamma_{0,i} = \mu_2 = q_1 = q_2 = 1$.

from one. This means that when the relay is active, it allocates more power to imbalanced networks, i.e., when one of the two channel gains is larger than the other. Fig. 5.5 shows the contour of the optimal adaptation scheme in (5.25) for $\mu_3 = q_1 = q_2 = 1$. As expected, it can be seen from Fig. 5.5 that more power is allocated when $\gamma_{1,i} \gg \gamma_{2,i}$ or $\gamma_{2,i} \gg \gamma_{1,i}$ in the high power regions. As noted before, the allocation in Fig. 5.5 is symmetric along the line $\gamma_{1,i} = \gamma_{2,i}$.

5.4 Illustrative Examples

In this section, simulation results are presented to quantify the gain of the proposed adaptation schemes. The two conventional systems using the FG and CI amplification coefficients are used as a baseline for comparison. The channel gains are assumed to be independent with Rayleigh-distributed magnitude as $h_l[i] \sim CN(0, \phi_l)$. Similar to Chapters 3 and 4, we adopt the linear network model so that $\phi_0 = 1$, $\phi_1 = 1/d^\nu$ and $\phi_2 = 1/(1-d)^\nu$, where $0 \leq d \leq 1$ is the normalized $S_1$-$R$ distance and $\nu$ is the pathloss exponent. For all simulations, $\nu = 3$, $d = 0.5$ and unit noise power at the three nodes is considered. In addition,
equal power allocation is assumed with $P_s = P_r = P_{sj} = P_t$, $q_1 = q_2 = z_2 = 2/3$ for TW/NAF and $q_1 = z_2 = 1$ for OAF.

Fig. 5.6 shows the achievable rates of the OAF system in (5.5) against power $P_t$ for three different adaptation techniques. These include the FG coefficient in (2.19) with $\Xi = \phi$, the CI coefficient in (2.19) with $\Xi = \alpha$, and the RA coefficient in (5.1) with $z_2[i]$ in (5.10). In this and all subsequent figures, the rates are averaged over $10^6$ channel realizations and are normalized by the rate of the FG technique. First, observe from Fig. 5.6 that the CI system outperforms the FG one for high values of $P_t$, whereas the opposite holds for low values. As discussed in Chapter 4, note that the performance of the CI and FG techniques converge in low power regions. More importantly, it can be seen from Fig. 5.6 that the proposed RA technique provides the best performance for all values of $P_t$. The gains over the CI and FG schemes are particularly significant in low $P_t$ regions. For instance, a $1.4 \times$ increase in rate can be observed at $P_t = -20$ dB.

The normalized achievable rates of the NAF system in (5.12) are shown in Fig. 5.7 for different power adaptation schemes. Recall from Chapter 4 that the NAF FG system is better than the CI one in high $P_t$ regions, but presents a similar performance for low $P_t$
values. As expected, the proposed RA technique in (5.1) and (5.18) outperforms these two schemes for the entire range of $P_t$ in Fig. 5.7. Similar to the OAF case, the gains are more significant in low $P_t$ regions. For example, a $1.25 \times$ gain is achieved in Fig. 5.7 at $P_t = -20$ dB.

Fig. 5.8 shows the normalized sum rates of the TW system using the FG/CI coefficients in (2.28), and the RA coefficient in (5.2) and (5.25). Observe from Fig. 5.8 that, as discussed in Chapter 4, CI is superior to FG for high values of $P_t$, whereas FG is dominant for low values. Similar to the previous cases, it can be seen from Fig. 5.8 that the proposed RA technique provides impressive gains over the FG and CI techniques, specially in low power regions. Specifically, a 2-fold increase over the FG system can be seen at $-17$ dB. However, for $P_t > 10$ dB, the sum rates of the CI and RA systems converge and do not significantly outperform that of the FG system. The behavior of the TW system is thus similar to that in single-input single-output channels, where the rate is insensitive to power adaptation in high power regions, but very sensitive at low power regions [39].

Finally, we should note that although we only consider $d = 0.5$ in this section, similar trends have been observed for other values of $d$. Moreover, although here we consider the
Fig. 5.7 Normalized achievable rates for the NAF system using different power adaptation schemes.

Fig. 5.8 Normalized sum rates of the TW system using different power adaptation schemes.
achievability rate performance with Gaussian inputs, significant gains in terms of BER have also been observed in [78] by applying the proposed relay adaptation schemes to uncoded and coded systems with finite constellations.

5.5 Chapter Summary

This chapter considered the design of optimal relay adaptation schemes over OAF, NAF and TWAF channels. By assuming Gaussian codebooks at the source nodes and full CSI at the relay, optimal power amplifications were derived to maximize the achievable rate or sum rate of the considered systems. The obtained closed-form solutions indicate that the optimal schemes are a multi-D water-filling like formula in time and space. Important insights on the behavior of the proposed schemes with regards to the channel conditions were then presented. Simulation results were finally provided to demonstrate the advantage of the proposed relay adaptation system over conventional AF systems using either the FG or the CI coefficient.
Chapter 6

Multi-frame Precoding for Non-orthogonal AF Relaying over Fading Channels\(^1\)

In previous chapters, we addressed the capacity of several AF relay systems assuming Gaussian codebooks at the source nodes. Besides the rate performance with Gaussian inputs, it is also important to design practical coding schemes using finite constellations. In this regard, several techniques have been proposed in the literature to exploit the diversity of AF systems over fast fading channels. For instance, by applying the idea of signal space diversity \([41]\) to uncoded NAF systems via a precoding matrix, it was shown in \([109, 110]\) that full cooperative diversity can be achieved. By further incorporating such a precoding technique into a bit-interleaved coded modulation (BICM) system with iterative decoding (ID), references \([111, 112]\) showed that both time and cooperative diversities can be exploited simultaneously. In particular, it was demonstrated in \([111, 112]\) that the diversity gain function of the considered NAF-BICM-ID system is \(d_H\)-th power of that of the uncoded full-diversity cooperative system in \([109, 110]\). Since the precoders in \([111, 112]\) are restricted to a single cooperative frame, the maximum achievable diversity order is constrained by the minimum Hamming distance \(d_H\) of the outer code. Without this restriction, higher degrees

\(^1\)Parts of this chapter have been presented at the 2011 IEEE International Wireless Communications and Mobile Computing Conference in Istanbul, Turkey \([105]\), and the 2011 IEEE Vehicular Technology Conference in San Francisco, USA \([106]\); and have been published in the Wiley Journal of Wireless Communications and Mobile Computing \([107]\), and the IEEE Transactions on Vehicular Technology \([108]\).
of freedom, and consequently, larger time diversity order, can be attained and incorporated into the cooperative system for performance improvement.

Usually, diversity benefits offered by relaying can only be observed in the error-floor region when the system operates at a sufficiently high SNR. For an ergodic fading relay channel, designing a practical coding strategy that can approach closely to its fundamental limit in the turbo pinch-off or waterfall region is also particularly important. To the best of our knowledge, most of the designs available so far are made to DF relaying [113–115]. Despite its practicality, there is not much advancement in designing a good coding scheme for NAF systems.

Motivated by the above discussions, this chapter proposes a precoding scheme over multiple cooperative frames for a BICM system over a half-duplex NAF relay channel. We focus on both diversity improvement in the error-floor region and near-capacity performance in the turbo pinch-off region. Our detailed contributions can be summarized as follows:

- Focusing on the error-floor region, we show that the multiple-frame precoded system can provide a diversity gain function that is $N$-th power of that of the previously considered NAF-BICM-ID system in [111], where $N$ is the number of precoded cooperative frames. An optimal class of precoders is then derived based on the asymptotic coding gain. This class indicates that the source should transmit a superposition of all symbols in the broadcasting phases, while being silent in all cooperative phases to optimize the asymptotic performance. A design criterion is also developed to find optimal superposition angles for good convergence behavior. A pragmatic approach is then proposed to obtain good rotation angles. Analytical and simulation results indicate that the proposed precoders achieve higher diversity orders and provide significant coding gains over the single-frame precoding scheme.

- In the turbo pinch-off region, we demonstrate that a concatenation of multi-D mapping [116], multiple-frame precoding and a simple outer binary code can be used to achieve near-capacity performance. We first show that for a given unitary and full diversity precoder, the optimal multi-D labeling for NAF relaying shall maximize the average Euclidean distance between all pairs in the multi-D constellation whose labels differ in only one bit. The extrinsic information transfer (EXIT) charts are then used to match the outer code, the multi-D mapping, and the precoder for near-capacity performance. For various spectral efficiencies, we obtain a BER of $10^{-5}$ or lower at
a SNR that is 0.45 dB-1 dB below the achievable rate with finite constellations and within 1.55 dB-1.95 dB from the ergodic capacity with Gaussian inputs.

The rest of the chapter is organized as follows. Section 6.1 introduces the NAF-BICM system under consideration. The error performance of the considered system is then analyzed in Section 6.2. The diversity benefits in error-floor regions are investigated in Section 6.3 along with the design of asymptotically optimal precoders. Section 6.4 then addresses the design in turbo pinch-off areas to approach the capacity of the NAF channel. Simulation results are provided in Section 6.5 to verify the analysis. Finally, concluding remarks are drawn in Section 6.6.

6.1 System Model

The general block diagram of the proposed NAF-BICM-ID system using a $2N \times 2N$ precoder $G$ over $N$ cooperative frames is shown in Fig. 6.1. The information sequence $u$ is first encoded into a coded sequence $c$. This coded sequence is then interleaved by a bit-interleaver $\Pi$ to become the interleaved sequence $v$. Each group of $2Nm_c$ bits in $v$ is mapped to a signal $s = [s_1, s_2, \ldots, s_{2N}]^T$ in the complex $2N$-dimensional ($2N$-D) constellation $\Psi$ according to a mapping rule $\xi$. Each component $s_k$ is in a 1-D unit-energy constellation $\Omega$, such as QPSK or QAM, of size $2^{m_c}$. In general, the mapping rule $\xi$ can be implemented independently in $\Omega$ for each component $s_k$ or might span the entire $N$ cooperative frames, which is referred to as multi-D mapping [116]. The symbol $s \in \Psi$ is then rotated by a $2N \times 2N$ complex precoder $G$ with entries $\{g_{i,k}\}$ ($1 \leq i, k \leq 2N$). The rotated symbol $x = Gs$ now corresponds to a new rotated constellation $\Psi_r$. Each symbol $x \in \Psi_r$ is then transmitted over the NAF channel as will be described shortly. At the destination, as depicted in Fig. 6.1, the received vector $y$ is demodulated and decoded in an iterative manner. The receiver consists of a soft-output demodulator that follows the maximum a posteriori probability (MAP) algorithm similar to [50], and a soft-input soft-output (SISO) channel decoder which uses the MAP algorithm in [117].

The transmission from source to destination is carried using the NAF protocol. In particular, each symbol $x \in \Psi_r$ is sent via $N$ cooperative frames as described in (2.16). Hereafter, a group of $N$ cooperative frames shall be referred to as a super-frame. Let $x = Gs = [x_1^T, \ldots, x_N^T]^T$. Then, the $2 \times 1$ vector $x_i = [x_{1,i}, x_{2,i}]^T = G_i s$ is transmitted in the $i$-th cooperative frame and the sub-precoder $G_i$ is of size $2 \times 2N$. To illustrate this
idea, Table 6.1 compares the multiple-frame precoding technique over $N = 2$ consecutive cooperative frames to the conventional NAF transmission without precoder and the NAF scheme using a single-frame precoder. In Table 6.1, $A \xrightarrow{\leftarrow} B$ denotes that node $A$ transmits $x$ to node $B$, and $A \xrightarrow{\text{AF}} B$ means that $A$ amplifies and forwards the signal received in the previous phase to $B$.

**Table 6.1** NAF transmission in two consecutive cooperative frames.

<table>
<thead>
<tr>
<th>Precoder</th>
<th>Cooperative frame 1 Phase 1</th>
<th>Cooperative frame 2 Phase 1</th>
<th>Cooperative frame 1 Phase 2</th>
<th>Cooperative frame 2 Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without</td>
<td>$S \xrightarrow{s_1} R, D$</td>
<td>$S \xrightarrow{s_2} R$</td>
<td>$S \xrightarrow{s_3} R, D$</td>
<td>$S \xrightarrow{s_3} R, D$</td>
</tr>
<tr>
<td></td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$S \xrightarrow{s_3} R, D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
</tr>
<tr>
<td>Single</td>
<td>$S \xrightarrow{s_1 g_{1,1} + s_2 g_{1,2}} R, D$</td>
<td>$S \xrightarrow{s_1 g_{1,1} + s_2 g_{1,2}} D$</td>
<td>$S \xrightarrow{s_3 g_{1,1} + s_4 g_{1,2}} R, D$</td>
<td>$S \xrightarrow{s_3 g_{1,1} + s_4 g_{1,2}} R, D$</td>
</tr>
<tr>
<td></td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
</tr>
<tr>
<td>Multi</td>
<td>$S \xrightarrow{\sum_{k=1}^{4} g_{1,k} s_k} R, D$</td>
<td>$S \xrightarrow{\sum_{k=1}^{4} g_{1,k} s_k} D$</td>
<td>$S \xrightarrow{\sum_{k=1}^{4} g_{3,k} s_k} R, D$</td>
<td>$S \xrightarrow{\sum_{k=1}^{4} g_{3,k} s_k} R, D$</td>
</tr>
<tr>
<td></td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
<td>$R \xrightarrow{\text{AF}} D$</td>
</tr>
</tbody>
</table>

Let the channel gains at frame $i$ be given by $h_i = [h_{0,i}, h_{1,i}, h_{2,i}]$ according to Fig. 2.1. In this chapter, we again consider the block fast fading scenario in which the transmitted codeword spans several realizations. In particular, the channel gains remain constant during
one cooperative frame but change independently from one frame to another. Furthermore, we assume that the destination has perfect knowledge of these channel gains, while the relay only knows the second order statistics of the $S-R$ channel. To concentrate on the code design, unit-variance Rayleigh fading $\mathcal{CN}(0,1)$ and equal noise variances at relay and destination $N_d = N_r = N_0$ are considered. Denote the covariance matrix of the transmitted vector as $Q = \mathbb{E}[xx^\dagger] = GG^\dagger$. From (2.17), the power transmitted by the source is then $P_s \cdot \text{tr}(Q) = P_s \cdot \text{tr}(GG^\dagger)$ per cooperative frame or $P_s \cdot \text{tr}(Q)/2N$ per symbol period. Here, we assume that $\text{tr}(Q) = \text{tr}(GG^\dagger) = \sum_{i=1}^{2N} \sum_{k=1}^{2N} |g_{i,k}|^2 = 2N$ so that the transmitted power is $P_s$ per symbol period. Allocating a power of $P_r = P_s$ per cooperative frame to the relay, the FG amplification coefficient at frame $i$ can be written as $b_i = \sqrt{1/\left[\eta_i P_s + N_0\right]}$, where $\eta_i = \sum_{l=1}^{2N} |g_{2i-1,l}|^2$ for $1 \leq i \leq N$. Given these constraints, the source spends a total power of $2NP_s$, while the relay uses $NP_s$ in any given super-frame.

By stacking the received signals in (2.17) for $N$ frames and whitening the noise components, the $2N \times 1$ received vector at $D$ can be written as

$$y = \sqrt{P_s}Hx + n,$$

(6.1)

where $n \sim \mathcal{CN}(0,N_0I_{2N})$, and the $2N \times 2N$ equivalent channel matrix is expressed as

$$H = \text{bdiag}(H_1, H_2, \ldots, H_N).$$

(6.2)

In (6.2), $\text{bdiag}(\cdot)$ denotes the block-diagonal matrix function and the $2 \times 2$ channel submatrix $H_i$ is given by

$$H_i = \begin{pmatrix}
h_{0,i} & 0 \\
v_i \sqrt{P_s b_i} h_{2,i} & v_i^2 h_{0,i}
\end{pmatrix},$$

with the whitening factor $v_i = 1/\sqrt{1 + P_s b_i^2 |h_{2,i}|^2}$. For mathematical convenience, the signal component in (6.1) can then be rewritten as in [109]:

$$Hx = \Sigma X Th_{01},$$

(6.3)

where $\Sigma = \text{bdiag}(\Sigma_1, \Sigma_2, \ldots, \Sigma_N)$ with $\Sigma_i = \begin{pmatrix} 1 & 0 \\ 1 & v_i \end{pmatrix}$, $X = \text{bdiag}(X_1, X_2, \ldots, X_N)$
with $X_i = \begin{pmatrix} x_{1,i} & 0 \\ x_{2,i} & x_{1,i} \end{pmatrix}$, $T = \text{bdiag} (T_1, T_2, \ldots, T_N)$ with $T_i = \text{bdiag} (1, \sqrt{P_s} b_{i2}, h_{i1})$, and $h_{01}^\top = (h_{01,1}^\top, \ldots, h_{01,N}^\top)$ with $h_{01,i} = (h_{0,i}, h_{1,i})^\top$.

### 6.2 Performance Analysis

This section presents a union bound to the BER for the NAF-BICM-ID system precoding over multiple cooperative frames. This bound will be used in the next sections for the designs in both error-floor and turbo pinch-off areas. The derivation presented here can be considered as an extension of that given in [111] to multi-frame precoders.

In general, the union bound to the BER $P_b$ for the NAF-BICM-ID system using a rate-$k/n$ outer code is given by

$$P_b \leq \frac{1}{k_c} \sum_{d_H} c_d f(d, \Psi, \xi, G) , \quad (6.4)$$

where $c_d$ is the total information weight of all error events at Hamming distance $d$ and $d_H$ is the free Hamming distance of the code. The function $f(d, \Psi, \xi, G)$ is the average PEP between two codewords forming an error event and can be computed as follows.

First, let $e$ denote the super-frame index, i.e., super-frame $e$ contains $N$ cooperative frames or equivalently $2N$ symbol periods. Then, let $\mathbf{c}$ and $\tilde{\mathbf{c}}$ denote respectively the input and estimated sequences with Hamming distance $d$ between them. These binary sequences correspond to the sequences $\mathbf{s}$ and $\tilde{\mathbf{s}}$, whose elements are $2N$-D symbols in $\Psi$. Without loss of generality, assume $\mathbf{c}$ and $\tilde{\mathbf{c}}$ differ in the first $d$ consecutive bits. Hence, $\mathbf{s}$ and $\tilde{\mathbf{s}}$ can be redefined as sequences of $d$ complex $2N$-D symbols as $\mathbf{s} = [s_1, \ldots, s_d]$ and $\tilde{\mathbf{s}} = [\tilde{s}_1, \ldots, \tilde{s}_d]$. Also, let $\mathbf{H} = [H_1, \ldots, H_d]$, where $H_e$ represents the channel matrix in (6.2) at super-frame $e$ affecting the transmitted symbol $s_e$, $1 \leq e \leq d$. Then, the PEP conditioned on $\mathbf{H}$ can be computed from (6.1) as

$$P(\mathbf{s} \rightarrow \tilde{\mathbf{s}} | \mathbf{H}) = Q \left( \sqrt{\frac{1}{2N_0} \sum_{e=1}^d d^2(s_e, \tilde{s}_e | H_e)} \right) , \quad (6.5)$$

where $Q(\cdot)$ is the Gaussian probability integral in (2.3) and $d^2(s_e, \tilde{s}_e | H_e)$ is the squared
Euclidean distance between $s_{(e)}$ and $\tilde{s}_{(e)}$ conditioned on $H_{(e)}$ and in the absence of noise:

$$d^2(s_{(e)}, \tilde{s}_{(e)}|H_{(e)}) = P_s\|H_{(e)}G(s_{(e)} - \tilde{s}_{(e)})\|^2. \quad (6.6)$$

Using the alternate representation of $Q(x) = \frac{1}{\pi} \int_0^{\pi/2} \exp \left(-x^2/[2 \sin^2 \theta] \right) d\theta$ in (2.3) and by averaging over each $H_{(e)}$ in $\bar{H}$, the unconditional PEP is given by

$$\mathbb{P}(\tilde{g} \rightarrow \tilde{s}) = \frac{1}{\pi} \int_0^{\pi/2} \left[ \prod_{e=1}^d \Delta_{\theta} \left( s_{(e)}, \tilde{s}_{(e)} \right) \right] d\theta, \quad (6.7)$$

where

$$\Delta_{\theta} \left( s_{(e)}, \tilde{s}_{(e)} \right) = \mathbb{E}_{H_{(e)}} \left[ \exp \left(-c_{\theta}\|H_{(e)}G(s_{(e)} - \tilde{s}_{(e)})\|^2 \right) \right], \quad (6.8)$$

with $c_{\theta} = P_s/[4N_0 \sin^2 \theta]$. Then, from the alternate matrix model in (6.3), the distance in (6.6) can be re-written as

$$d^2(s_{(e)}, \tilde{s}_{(e)}|H_{(e)}) = P_s h_{01}^{(e)^T} T_{(e)} U_{(e)}^T \Sigma_{(e)} \Sigma_{(e)} U_{(e)} T_{(e)} h_{01}^{(e)},$$

where $U_{(e)} = X_{(e)} - \bar{X}_{(e)} = \text{bdiag} \left( U_1^{(e)}, \ldots, U_N^{(e)} \right)$ with $U_i^{(e)} = \begin{pmatrix} u_{1,i}^{(e)} & 0 \\ u_{2,i}^{(e)} & u_{1,i}^{(e)} \end{pmatrix}$. Using the fact that for a given $a \sim \mathcal{CN}(0, \Phi)$ and a Hermitian matrix $A$, $\mathbb{E}[\exp(-a^H Aa)] = 1/\det(I + \Phi A)$ to average over $h_{01}^{(e)}$, (6.8) can be simplified as in [109] to

$$\Delta_{\theta} \left( s_{(e)}, \tilde{s}_{(e)} \right) = \mathbb{E}_{h_{2,1,\ldots,2,N}^{(e)}} \left[ \frac{1}{\det(A_{(e)})} \right],$$

where $A_{(e)} = I_{2N} + c_{\theta} T_{(e)} U_{(e)}^T \Sigma_{(e)} \Sigma_{(e)} U_{(e)} T_{(e)}$. It is easy to verify that $A_{(e)}$ is a block diagonal matrix and therefore can be written as $A_{(e)} = \text{bdiag} \left( A_1^{(e)}, \ldots, A_N^{(e)} \right)$, where

$$A_i^{(e)} = I_2 + c_{\theta} \begin{pmatrix} |u_{1,i}^{(e)}|^2 + v_i^{(e)^2} |u_{2,i}^{(e)}|^2 & v_i^{(e)^2} \sqrt{P_s} b_{2,i}^{(e)} u_{2,i}^{(e)} u_{1,i}^{(e)} H \v_i^{(e)^2} \sqrt{P_s} b_{1,i}^{(e)} u_{1,i}^{(e)} u_{2,i}^{(e)} \v_i^{(e)^2} P_s b_i^{(e)^2} |h_{2,i}^{(e)}|^2 |u_{1,i}^{(e)}|^2 \end{pmatrix}.$$
\( \Delta_{\theta}(s_e, \hat{s}_e) \) can be simplified to

\[
\Delta_{\theta}(s_e, \hat{s}_e) = \prod_{i=1}^{N} \Delta_{\theta,i}(s_e, \hat{s}_e), \tag{6.9}
\]

where \( \Delta_{\theta,i}(s_e, \hat{s}_e) = E_{h_{2,i}}[\det(A_{i}^{(e)})^{-1}] \) has already been computed in [111] and can be expressed in a closed-form as

\[
\Delta_{\theta,i}(s_e, \hat{s}_e) = \frac{1}{(c_{\theta}\|u_{i}^{(e)}\|^2 + 1)^2} \left\{ 1 + \left( \frac{1}{P_s b_i^2} - a \right) J(a) \right\}. \tag{6.10}
\]

In (6.10), \( a = \frac{(c_{\theta}\|u_{i}^{(e)}\|^2 + 1)}{[P_s b_i^2(c_{\theta}\|u_{i}^{(e)}\|^2 + 1)^2]} \); \( J(x) = \exp(x)E_1(x) \) with \( E_1(x) \) the exponential integral as in (4.4); and \( \|u_{i}^{(e)}\|^2 = |u_{1,i}^{(e)}|^2 + |u_{2,i}^{(e)}|^2 \) with \( u_{i}^{(e)} = [u_{1,i}^{(e)}, u_{2,i}^{(e)}]^T = G_i(s_e - \hat{s}_e) = G_i\epsilon_e \).

The PEP can finally be obtained by substituting (6.9) into (6.7). \( f(d, \Psi, \xi, G) \) can then be upper bounded by averaging over all possible cases of PEPs in (6.7). Using the error-free feedback bound [50], one has perfect a priori information of the coded bits fed back to the demodulator. As such, the other coded bits carried by the transmitted symbol can be assumed to be known perfectly and thus the error happens only when the labels of \( s_e \) and \( \hat{s}_e \) differ by only one bit. By further using the fact that the channel changes independently with \( e \), the upper bound on \( f(d, \Psi, \xi, G) \) can be simply obtained by averaging over the constellation \( \Psi \) as

\[
f(d, \Psi, \xi, G) \leq \frac{1}{\pi} \int_{0}^{\pi/2} \left[ \mathbb{E}_{s,p}[\Delta_{\theta}(s, p)] \right]^{d} d\theta \\
= \frac{1}{\pi} \int_{0}^{\pi/2} [\gamma_e(\Psi, \xi, G)]^d d\theta. \tag{6.11}
\]

In (6.11), the expectation is taken over all 1-bit neighbors in the constellation \( \Psi \), i.e., over all pairs of symbols \( s, p \in \Psi \) whose labels differ in only one bit. Furthermore, \( \Delta_{\theta}(s, p) \) is calculated as in (6.9), with \( s \) and \( p \) substituted for \( s_e \) and \( \hat{s}_e \), respectively. The
expectation in (6.11) is then expressed as
\[
\gamma_\theta(\Psi, \xi, G) = \frac{1}{2Nm_c} \sum_{s \in \Psi} \sum_{k=1}^{2Nm_c} \left[ \prod_{i=1}^{N} \Delta_{\theta,i}(s, p) \right],
\]
(6.12)
where \( u_i \) in \( \Delta_{\theta,i}(s, p) \) is given as
\[
u_i = \left[ u_{1,i}^\top, u_{2,i}^\top \right]^\top = G_i(s - p) = G_i \epsilon. \]
The perfect feedback bound can then be obtained from (6.4) by substituting \( f(d, \Psi, \xi, G) \) as in (6.11), \( \gamma_\theta(\Psi, \xi, G) \) as in (6.12) and \( \Delta_{\theta,i}(s, p) \) as in (6.10). As shall be shown in Section 6.5, this bound is tight at practical BER levels and is therefore useful to predict the error performance of the considered NAF-BICM-ID system.

6.3 Diversity Benefits in Error-floor Regions

Although the error bound derived earlier is helpful to predict the error performance, it does not provide an insight about the effect of the proposed precoding scheme to the performance. In this section, by further simplifying the bound, we examine the diversity benefits offered by the multiple-frame precoding technique in the error-floor region. The optimal precoding scheme is then developed to minimize the error performance in the error-floor area.

6.3.1 Diversity and Coding Gain Functions

By applying the Chernoff bound \( Q(\sqrt{2x}) < \frac{1}{2} \exp(-x) \) for \( x > 0 \), \( f(d, \Psi, \xi, G) \) in (6.11) can further be upper bounded as
\[
f(d, \Psi, \xi, G) \leq \frac{1}{\pi} \int_0^{\pi/2} \left[ \gamma_\theta(\Psi, \xi, G) \right]^d d\theta < \frac{1}{2} \left[ \gamma_{\pi/2}(\Psi, \xi, G) \right]^d,
\]
(6.13)
where
\[
\gamma_{\pi/2}(\Psi, \xi, G) = \frac{1}{2Nm_c} \sum_{s \in \Psi} \sum_{k=1}^{2Nm_c} \left[ \prod_{i=1}^{N} \Delta_{\pi/2,i}(s, p) \right].
\]
(6.14)
In (6.14), \( \Delta_{\pi/2,i}(s, p) \) is calculated as \( \Delta_{\theta,i}(s, p) \) by substituting \( \theta = \pi/2 \).

Let \( \epsilon = s - p \), \( G^\top = [G_1^\top, \ldots, G_N^\top] \), and the \( 2 \times 2N \) sub-matrix \( G_i^\top = [g_{1,i}, \ldots, g_{2N,i}] \). Hence, \( G_i \) spans over the \( i \)-th frame \( (1 \leq i \leq N) \) and \( g_{l,i} = [g_{l,1,i}, \ldots, g_{l,2N,i}]^\top \) is of size
2N \times 1 \ (l = 1, 2). Following a similar approach as in [109], \( \Delta_{\pi/2,i}(s, p) \) in (6.14) can be approximated at high transmitted powers as

\[
\Delta_{\pi/2,i}(s, p) = \frac{16N_0^2\eta_i}{|u_{1,i}|^4} \cdot P_s^{-2} \ln(P_s) + O(P_s^{-2})
\]

\[
\approx \frac{16N_0^2\|g_{1,i}\|^2}{|g_{1,i}^\top \epsilon|^4} \cdot P_s^{-2} \ln(P_s),
\]

where recall that \( u_{1,i} = g_{1,i}^\top \epsilon \) and \( \eta_i = \|g_{1,i}\|^2 \). Consequently, \( \gamma_{\pi/2}(\Psi, \xi, G) \) in (6.14) can be also approximated as

\[
\gamma_{\pi/2}(\Psi, \xi, G) \approx 16N_0^2 \frac{\mathcal{F}(G, \epsilon)}{2N_m c^{2N_m c}} \cdot P_s^{-2N} \ln^N(P_s),
\]

(6.15)

with \( \mathcal{F}(G, \epsilon) \) given by

\[
\mathcal{F}(G, \epsilon) = \sum_{s \in \Psi} \prod_{k=1}^{2N_m c} \prod_{i=1}^{N} \frac{\|g_{1,i}\|^2}{|g_{1,i}^\top \epsilon|^4}.
\]

(6.16)

The parameter \( \mathcal{F}(G, \epsilon)^{-1} \) is defined as the coding gain function of the system. Now, by considering only the first term in (6.4) which is dominant at high powers, \( P_b \) can be approximated from (6.13) and (6.15) as

\[
P_b \approx c_{dh} 2^{k_c} \left( \frac{16N_0^2 \mathcal{F}(G, \epsilon)}{2N_m c^{2N_m c}} \right)^{d_H} \times [P_s^{-2} \ln(P_s)]^{N-d_H}.
\]

(6.17)

The term \( [P_s^{-2} \ln(P_s)]^{N-d_H} \) is designated as the diversity gain function of the system. From (6.17), it can be seen that as long as the coding gain function \( \mathcal{F}(G, \epsilon)^{-1} \neq 0 \), a diversity gain function of \( [P_s^{-2} \ln(P_s)]^{N-d_H} \) can be fully achieved. The diversity function of the proposed system is simply \( (N \cdot d_H) \)-th and \( N \)-th power of the diversity gain functions of uncoded cooperative systems in [109,110] and the NAF-BICM-ID systems in [111], respectively. It is important to note that the asymptotic performance of the system is independent of the second row of \( G_i \), i.e., it does not depend on \( g_{2,i} \).

For the special case in which the mapping \( \xi \) is implemented independently and identically for each component \( s_k \in \Omega \), only one element in \( \epsilon \) has a non-zero value. Hence, for every term in the \( \gamma_{\pi/2}(\Psi, \xi, G) \) summation, \( u_{1,i} = \epsilon(k)g_{1,i,k} \) and \( u_{2,i} = \epsilon(k)g_{2,i,k} \) for some \( 1 \leq k \leq 2N \).
Therefore, \( \gamma_{\pi/2}(\Psi, \xi, G) \) can be simplified by averaging over the 1-bit neighbors \( s, p \in \Omega \) rather than \( s, p \in \Psi \) as

\[
\gamma_{\pi/2}(\Psi, \xi, G) = \frac{1}{2Nm_c^2m_c} \sum_{s \in \Omega} \sum_{k=1}^{m_c} 2N \left[ \prod_{l=1}^{N} \Delta_{\pi/2,l}(s, p) \right],
\]

(6.18)

where now

\[
\Delta_{\pi/2,l}(s, p) \approx \frac{16N_0^2 \|g_{1,i}\|^2}{|g_{1,i}|^4|\epsilon|^4} \cdot P_s^{-2} \ln(P_s),
\]

and \( \epsilon = s - p \). Note from (6.18) that the performance of the system with independent mapping depends only on the magnitude of each component of \( G \) and not on the actual \( G \). The probability of error \( P_b \) can now be approximated as

\[
P_b \approx \frac{c_{dl}}{2k_c} \left( \frac{16N_0N^{2N}}{2Nm_c^2m_c} \cdot F_{\text{ind}}(G) \right) \sum_{s \in \Omega} \sum_{k=1}^{m_c} \frac{1}{|\epsilon|^4N} \left[ P_s^{-2} \ln(P_s) \right]^{d_{hl}},
\]

(6.19)

where \( F_{\text{ind}}(G) \) becomes

\[
F_{\text{ind}}(G) = \left( \prod_{i=1}^{N} \left[ \sum_{m=1}^{2N} |g_{1,m}^{(i)}|^2 \right] \right) \left( \sum_{l=1}^{2N} \left[ \prod_{i=1}^{N} \frac{1}{|g_{1,i}^{(l)}|^4} \right] \right).
\]

(6.20)

In this case, a full diversity gain function can be achieved as long as \( F_{\text{ind}}(G)^{-1} \neq 0 \).

### 6.3.2 Optimal Class of Precoders

When both diversity and coding gain functions are optimized, the probability of error can be minimized. For the general system using multi-D mapping \( \xi \), this can be achieved by solving the following optimization problem:

\[
\min_G \mathcal{F}(G, \epsilon) \quad \text{s.t.} \quad \|G\|^2 \leq 2N.
\]

(6.21)

It can be seen from (6.16) that a full diversity gain function of \( [P_s^{-2} \ln(P_s)]^{N\cdot d_{hl}} \) is guaranteed as long as \( g_{1,i}^\top \epsilon \neq 0 \) for all 1-bit neighbors \( s, p \in \Psi \). Hence, the rows of any optimal precoder \( g_{1,i} \) (\( 1 \leq i \leq N \)) must satisfy this condition. Given that this property is independent of the magnitude of \( g_{1,i} \), the problem in (6.21) can be carried in two steps.
First, to find the optimal power distribution $\eta$, let $g_{1,i} = \sqrt{\eta_i} \tilde{g}_i$ where $\tilde{g}_i$ is a unit vector ($\|\tilde{g}_i\| = 1$). The coding gain function $F(G, \epsilon)$ in (6.16) can then be re-written as

$$F(G, \epsilon) = \sum_{s \in \Psi} \sum_{k=1}^{2Nm_c} \frac{1}{\eta_i |\tilde{g}_{i,\epsilon}|^4}. \quad (6.16)$$

From the inequality of the geometric and arithmetic means and from the power constraint of the precoder, it can be shown that

$$F(G, \epsilon) \geq \frac{1}{2N} \sum_{s \in \Psi} \sum_{k=1}^{2Nm_c} \frac{1}{\|\tilde{g}_{i,\epsilon}\|^4}, \quad (6.17)$$

where the equality is achieved when $\eta_i = 2 \forall i$. The objective function in (6.21) is then minimized when the $2N$ power of $G$ is equally shared to the $N$ rows $g_{1,i}$, and no power is given to $g_{2,i}$, i.e., $g_{2,i} = 0$. The asymptotic optimal class of precoders $G^*$ for any mapping can then be written in general form as

$$G^* = \begin{pmatrix} G^*_{1} \\ \vdots \\ G^*_{N} \end{pmatrix}, \quad \text{where} \quad G^*_{i} = \begin{pmatrix} g_{1,i}^\top \\ 0 \end{pmatrix}, \quad (6.22)$$

and $\|g_{1,i}\|^2 = 2$. Similar to [111], (6.22) indicates that the source and relay must transmit orthogonally to achieve the best asymptotic performance. That is, the source should equally distribute all its power in $N$ broadcasting phases to transmit a superposition of $2N$ symbols, while being silent in the respective $N$ cooperative phases.

The second step is to maximize the coding gain over the optimal class of precoders in (6.22). From this optimal class, the optimization problem in (6.21) simplifies to

$$\min_{\hat{g}_i} \sum_{s \in \Psi} \sum_{k=1}^{2Nm_c} \frac{1}{\|\tilde{g}_{i,\epsilon}\|^4} \quad \text{s.t.} \quad \|\tilde{g}_i\| = 1. \quad (6.23)$$

Intuitively, we must find the $N$ vectors $\tilde{g}_i$ lying on the surface of a $2N$-D unit-radius hypersphere that minimize the objective function in (6.23). For full cooperative diversity, these vectors must not be perpendicular to any of the 1-bit error vectors $\epsilon$, i.e., $g_{1,i}^\top \epsilon \neq 0$. Hence, the problem in (6.23) is constellation and mapping dependent, and does not
have a general solution. Furthermore, due to the number of variables involved, (6.23) is numerically and analytically very complex to solve for multi-D mappings with \( N > 1 \). For the system using independent mapping \( \xi \), the problem in (6.21) simplifies to

\[
\min_{\mathbf{G}} \mathcal{F}_{\text{ind}}(\mathbf{G}) \quad \text{s.t.} \quad \|\mathbf{G}\|^2 \leq 2N, \tag{6.24}
\]

and full diversity is guaranteed as long as \( g_{1,i} \) has no zero components. By the inequality of the geometric and arithmetic means and from the power constraint of the precoder \( \mathbf{G} \), it can be shown that

\[
\mathcal{F}_{\text{ind}}(\mathbf{G}) \geq \left( 2N^N \prod_{i=1}^{N} \prod_{m=1}^{2N} |g_{1,m}^{(i)}|^\frac{1}{N} \right) \left( 2N^N \prod_{l=1}^{2N} \prod_{i=1}^{N} |g_{l,i}^{(i)}| - \frac{2}{N} \right) \geq 2^{N+1} \frac{N^{2N+1}}{N^2},
\]

where the equalities are achieved when \( |g_{1,1}^{(i)}|^2 = |g_{1,2}^{(i)}|^2 = \cdots = |g_{1,2N}^{(i)}|^2 = 1/N \) for all \( 1 \leq i \leq N \). Hence, the optimization problem in (6.24) reaches its minimum when \( |g_{1,k}^{(i)}|^2 = 1/N \) and \( |g_{2,k}^{(i)}|^2 = 0, \forall i, k \). The asymptotically optimal precoder for independent mappings can then be written as

\[
\mathbf{G}(\varphi) = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{G}_{1}(\varphi_1) \\ \vdots \\ \mathbf{G}_{N}(\varphi_N) \end{pmatrix}, \quad \text{where} \quad \mathbf{G}_{i}(\varphi_i) = \begin{pmatrix} g_{\varphi_i}^\top \\ 0 \end{pmatrix}, \tag{6.25}
\]

with \( \varphi^\top = [\varphi_1^\top, \ldots, \varphi_N^\top] \), \( \varphi_i = [\varphi_1^{(i)}, \ldots, \varphi_{2N-1}^{(i)}] \) and \( g_{\varphi_i}^\top = [1, e^{j\varphi_1^{(i)}}, \ldots, e^{j\varphi_{2N-1}^{(i)}}] \). In the next subsection, we shall further address the design of the angles \( \varphi \) to achieve good convergence properties for independent mappings.

### 6.3.3 Design of Superposition Angles For Independent Mappings

Among a wide range of optimal \( \mathbf{G}(\varphi) \) in (6.25), the one that yields a faster convergence (i.e., using a smaller number of iterations) to the asymptotic performance is preferred. Therefore, one also needs to take into account the performance at the first iteration. A reasonable approach is to consider the worst-case of PEP in (6.7) and to find \( \varphi \) that minimizes this PEP.
Applying the Chernoff bound to the Gaussian integral in (6.7), the PEP can be approximated at high powers as

\[ P(s \to \tilde{s}) \approx \frac{1}{2} \prod_{e=1}^{d} \prod_{i=1}^{N} \Delta_{\pi/2,i} \left( s(e), \tilde{s}(e) \right). \]  
(6.26)

Using the $G(\phi)$ precoder and following a similar approach as in the previous section, the PEP can be further approximated as

\[ P(s \to \tilde{s}) \approx \frac{32dN N^{2dN} P_s^{-2dN} \ln^{dN}(P_s)}{2} \prod_{e=1}^{d} \prod_{i=1}^{N} \frac{1}{|u_{1,i}^{(e)}|^4}, \]  
(6.27)

where $u_{1,i}^{(e)} = g_{\phi_i}^\top \epsilon = \epsilon_1 + \sum_{k=1}^{2N-1} \epsilon_{k+1} \exp(j\varphi_k^{(i)})$. Note from (6.27) that the worst-case PEP is independent of $e$ and hence the super-frame index can be dropped. The worst-case PEP can then be minimized by solving the following:

\[
\max_{\varphi} \min_{s, \tilde{s} \in \Psi, s \neq \tilde{s}} \prod_{i=1}^{N} |u_{1,i}^{(e)}|^4, \quad (6.28)
\]

where $u_{1,i} = g_{\varphi_i}^\top \epsilon = \epsilon_1 + \sum_{k=1}^{2N-1} \epsilon_{k+1} \exp(j\varphi_k^{(i)})$, with $\epsilon = s - \tilde{s} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_{2N})^\top \neq 0$.

The optimization problem in (6.28) is constellation dependent and does not have a general solution. For the simplest case of $N = 1$, this problem has been analytically solved in [111] for several $M$-QAM modulation schemes. However, for multiple-frame precoding with $N > 1$, the problem in (6.28) is mathematically and numerically involved. As an alternative, in the following, we propose a pragmatic approach based on (6.28) to find good rotation angles $\varphi$.

First, the optimization problem in (6.28) can be relaxed to find the optimal $\varphi_1$ as

\[
\max_{\varphi_1} \min_{s, \tilde{s} \in \Psi, s \neq \tilde{s}} |g_{\varphi_1}^\top \epsilon|^4. \quad (6.29)
\]

This guarantees that the first term of the product in (6.28) is reasonably large. The solution to (6.29) is similar to that of multi-user systems in [118] solved numerically for various values of $N$ and several square $M$-QAM modulation schemes. Then, to find $\varphi_i$ for $2 \leq i \leq N$, we use the vector $g_{\varphi_i}$ obtained from (6.29) for all $i$, except that the signs
of \( N \) values in this vector are flipped to make sure that all vectors in the set \( \{ g_{\varphi_i} \} \) are orthogonal. The product in (6.28) is then reasonably large. For instance, by applying this approach to the case of \( N = 2 \), one has

\[
G(\varphi) = \frac{1}{\sqrt{2}} \begin{pmatrix}
1 & e^{j\varphi_1} & e^{j\varphi_2} & e^{j\varphi_3} \\
0 & 0 & 0 & 0 \\
1 & -e^{j\varphi_1} & e^{j\varphi_2} & -e^{j\varphi_3} \\
0 & 0 & 0 & 0
\end{pmatrix}.
\] (6.30)

In (6.30), the angles \( \{ \varphi_1, \varphi_2, \varphi_3 \} \) are the solution to (6.29) and are provided in [118]. As will be shown in the Section 6.5, the system using the optimal precoder in (6.30) presents significant coding gains as compared to other precoders. Furthermore, the angles obtained from the pragmatic approach indeed result in a good convergence property. The results for the case of \( N > 2 \) can also be obtained in similar manner.

### 6.4 Near-Capacity Design in Turbo Pinch-off Areas

The diversity analysis presented earlier is useful in the BER floor region. In this region, a reasonably low BER can be achieved at a sufficiently high transmitted powers. Also of importance is the so-called turbo pinch-off or waterfall region, where a significant BER reduction is observed over iterations. The turbo pinch-off region can be used to examine whether the coded system under consideration can achieve near-capacity performance.

In order to approach near-capacity performance, state-of-the-art codes such as turbo and turbo-like codes (e.g., LDPC) are usually used. In fact, such powerful codes have been successfully applied to many channels, including the relay channels in [113–115]. Lately, over Rayleigh fading channels, reference [119] proposes a powerful coding scheme in which a multi-D mapping is employed in a rotated multi-D constellation. Interestingly, it is demonstrated in [119] that by only using a simple outer convolutional code, the proposed technique outperforms any coded system using turbo-like codes with traditional modulation schemes.

In this section, using EXIT chart analysis [53], we show that by employing a suitable multi-D mapping \( \xi \) in multiple precoded cooperative frames, the considered BICM-ID system can also offer a capacity-approaching performance. For simplicity, particular attention
is paid to the case of \( N = 2 \). Extension to larger \( N \) can be done in a similar manner.

### 6.4.1 Multi-D Mapping in Precoded Multiple Cooperative Frames

Here, we propose the use of multi-D mapping \( \xi \) in \( N \) precoded cooperative frames along with a full-diversity rotation \( \mathbf{G} \) [41]. Full diversity rotations have the property that \( \mathbf{x} - \hat{x} = \mathbf{G}(\mathbf{s} - \hat{s}) \) contains all nonzero components as long as \( \mathbf{s} \neq \hat{s} \). Note from (6.17) that this type of rotation achieves the full cooperative diversity gain \( (P_s^{-2} \ln(P_s))^N \cdot d_H \) of the precoded NAF-BICM system for any labeling technique. Furthermore, full-diversity rotations can increase the achievable rate over ergodic fading channels [119]. For example, Fig. 6.2 shows the achievable information rates over a NAF channel achieved by three different inputs, namely Gaussian input, QPSK constellation, and 4-D precoded QPSK constellation using the \( 4 \times 4 \) full-diversity rotation \( \mathbf{G}_{st} \) given in [41]:

\[
\mathbf{G}_{st} = \frac{1}{\sqrt{4}} \begin{pmatrix}
1 & \omega_1 & \omega_1^2 & \omega_1^3 \\
1 & \omega_2 & \omega_2^2 & \omega_2^3 \\
1 & \omega_3 & \omega_3^2 & \omega_3^3 \\
1 & \omega_4 & \omega_4^2 & \omega_4^3
\end{pmatrix},
\]

(6.31)

with \( \omega_1 = \omega = \exp(j\pi/8) \) and \( \omega_i = \omega \exp(j\pi(i - 1)/2) \). The rates in Fig. 6.2 are plotted against information bit SNR \( E_b/N_0 \). As can be seen from this figure, the achievable rate can be greatly improved by using the rotated multi-D constellation.

For a good mapping rule \( \xi \), one must minimize \( \gamma_{\pi/2}(\Psi, \xi, \mathbf{G}) \) in (6.14). To obtain more insights about the mapping rule, we first approximate \( \Delta_{\pi/2,i}(\mathbf{s}, \mathbf{p}) \) in (6.14) as

\[
\Delta_{\pi/2,i}(\mathbf{s}, \mathbf{p}) \approx \frac{1}{(c_{\pi/2}|u_{1,i}|^2 + 1)^2} \left\{ 1 + \left( \frac{1}{P_s b^2} - a \right) \ln \left( 1 + \frac{1}{a} \right) \right\}.
\]

(6.32)

Observe that \( \Delta_{\pi/2,i}(\mathbf{s}, \mathbf{p}) \) is only a function of \( (|u_{1,i}|^2, |u_{2,i}|^2) \), where recall that \( \mathbf{u}_i = [u_{1,i}^\top, u_{2,i}^\top]^\top = \mathbf{G}_i(\mathbf{s} - \mathbf{p}) \). Hence, \( \Delta_{\pi/2,i}(\mathbf{s}, \mathbf{p}) \) can be rewritten as \( \Delta_{\pi/2,i}(|u_{1,i}|^2, |u_{2,i}|^2) \). One then has the following proposition:

**Proposition 6.1.** \( \Delta_{\pi/2,i}(\mathbf{s}, \mathbf{p}) \) is a strictly decreasing function of \( |u_{1,i}|^2 \) and \( |u_{2,i}|^2 \).

**Proof.** See Appendix D.
Using Proposition 6.1 and given the fact that $|u_{1,i}|^2$ and $|u_{2,i}|^2$ are both positive for a full diversity $G$, it can be seen that the BER can be minimized by maximizing $|u_{1,i}|^2$ and $|u_{2,i}|^2$ for each term in the summation (6.14). Since $G$ is also unitary, it is equivalent to maximizing $\|u\|^2 = \|G(s - p)\|^2 = \|G\epsilon\|^2$ for all 1-bit neighbors $s$ and $p$ $(s, p \in \Psi)$. Using (6.14), it turns out that the design criterion is to maximize the average Euclidean distance between all 1-bit neighbors in the multi-D constellation $\Psi$. Interestingly, such a design criterion for multi-D mapping is the same as that over a single-antenna AWGN channel considered in [116]. Therefore, the optimal multi-D mapping for NAF channels can be obtained in a similar manner as that over an AWGN channel. As shall be seen shortly, by using such a multi-D mapping, one can select a relatively simple outer code for capacity-approaching performance.

### 6.4.2 EXIT Chart Analysis

To demonstrate the advantage of the proposed coding scheme, we first apply the EXIT chart technique [53] to examine the demapper’s EXIT curves. Then a combination of the demapper and a simple convolutional decoder, with which close-capacity performance can
be achieved, is studied. Due to space limitation, we only consider the QPSK modulation scheme and adopt the optimal 4-D mapping proposed in [116]. The results can be straightforwardly extended to other constellations.

Similar to [53], let $I_{A_1}$ and $I_{E_1}$ denote the mutual information between the a priori log-likelihood ratio (LLR) and the transmitted coded bit, and between the extrinsic LLR and the transmitted coded bit at the input and output of the demodulator, respectively. The demodulator EXIT characteristic is then given by $I_{E_1} = T_1 (I_{A_1}, E_b/N_0)$ and can be determined by Monte Carlo experiments [53]. Similarly, let $I_{A_2}$ and $I_{E_2}$ be the mutual information representing the a priori knowledge and the extrinsic information of the coded bits at the input and output of the SISO decoder. The decoder EXIT characteristic is then defined as $I_{E_2} = T_2 (I_{A_2})$. Note that due to the presence of the interleaver, this value does not depend on SNR. After being deinterleaved, the extrinsic output of the detector is used as the a priori input to the decoder, i.e., $I_{A_2} = I_{E_1}$. Furthermore, after being interleaved, the extrinsic information of the decoder becomes the a priori information to be provided to the demodulator, i.e., $I_{A_1} = I_{E_2}$.

To demonstrate the benefits obtained by multi-D mapping employed in two precoded cooperative frames, Fig. 6.3 shows the demmapper’s EXIT curves for three cases: i) 1-D Gray mapping without rotation; ii) optimal 4-D mapping without rotation; and iii) optimal 4-D mapping using $G_{st}$. It can be seen that by using the optimal multi-D mapping, one obtains two EXIT curves with a very steep slope, which make them suitable for a class of codes having also a decayed EXIT curve, such as convolutional codes. Furthermore, with the same optimal 4-D mapping, the full diversity rotation provides larger $I_{E_1}$ in a wide range of $I_{A_1}$. This advantage comes from the capacity improvement discussed earlier.

Now, given the superiority of multi-D mapping and precoding, we can apply the EXIT chart technique [53] to select the most suitable convolutional code for the proposed system. By using EXIT charts, both EXIT curves of the multi-D demodulator and the decoder are plotted in the same graph, but the axes of the decoder curve are swapped [53]. The convergence behavior can therefore be visualized. Note that for a given rate-$r_c$ code, its EXIT curve does not depend on SNR and always crosses the middle point $(0.5, r_c)$.

We have examined different code rates and observed that the multi-D demapper EXIT curve can be matched with very simple codes. For example, Fig. 6.4 plots the demapper EXIT curves at $E_b/N_0 = 5.1$ dB, along with that of the simple rate-2/3, 4-state punctured convolutional code obtained from the rate-1/2 convolutional code with generator matrix
Fig. 6.3  Demapper EXIT curves for three different systems.

The above SNR is chosen to make sure that an open tunnel exists between the two curves in Fig. 6.4. Note that at this rate, the ergodic capacity of the NAF channel calculated as in [22] and the constrained capacity with QPSK are at $E_b/N_0 = 3.6$ dB and $E_b/N_0 = 5.6$ dB, respectively. It can be seen from Fig. 6.4 that the two EXIT curves match very well at the chosen SNR. Specifically, the two curves only intersect at an ending point that is very close to $I_{E_1}$ (1). Apparently, our proposed system can operate below the constrained capacity with QPSK. This fact is later confirmed by simulations.

The proposed scheme can also work well with other code rates. Though not explicitly shown here, we observe that one can achieve a good match for the rate-3/4 system at $E_b/N_0 = 5.9$ dB using the rate-3/4, 4-state punctured convolutional code obtained from the rate-1/2 convolutional code with generator matrix $\{5; 7\}$ [120]. It should be mentioned that the ergodic capacity with Gaussian inputs and the constrained capacity with QPSK for this rate are at $E_b/N_0 = 4.05$ dB and $E_b/N_0 = 7$ dB, respectively.
6 Multi-frame Precoding for NAF Relaying over Fading Channels

6.5 Illustrative Examples

In this section, analytical and simulation results are provided to confirm the tightness of the derived bound and the superiority of the multiple-frame precoding technique in both error-floor and turbo pinch-off regions. Unless otherwise stated, a random interleaver of length 80000 bits is used in all simulations. In the computation of the bound to $P_b$ in (6.4), the first twenty terms of the summation are retained to provide an accurate bound. Furthermore, for simplicity, we only consider the NAF-BICM-ID system using the QPSK modulation scheme. Similar results can be obtained for different modulation schemes.

6.5.1 Tightness of the Derived Bound

To demonstrate the tightness of the bound when precoding over three frames, Fig. 6.5 shows the BER performance after 1, 2 and 3 iterations of the NAF-BICM-ID system using the $6 \times 6$ $G_{cyclo6}$ precoder from [121]. Gray mapping and the rate-1/2 4-state convolutional code with generator matrix \{5; 7\} are considered in all figures in this and the next subsection. Similar results can also be obtained for other codes and mapping rules. Note from Fig. 6.5 that
the BER performance of the system converges to the analytical bound after 2 iterations at sufficiently high powers. Specifically, the analytical and simulation results converge around the BER level of $10^{-3}$.

![BER performance graph](image)

**Fig. 6.5** BER performance after 1, 2 and 3 iterations of the NAF-BICM-ID system with $N = 3$ using the $G_{\text{cyclo6}}$ precoder.

To verify the tightness of the derived bound when precoding over two cooperative frames, Fig. 6.6 presents the BER performances after 5 iterations of the NAF-BICM-ID system employing three different $4 \times 4$ precoders. The considered precoders include $I_4$, i.e., no precoder, and the $4 \times 4 G_{\text{mixed}}$ and $G_{\text{krus}}$ precoders given in [121]. Observe from Fig. 6.6 that the error performances of all systems converge to their respective bounds at practical BER levels. In particular, the analytical and simulation results converge around the BER level of $10^{-3}$ for all systems under consideration. This makes the bound an effective tool to predict the error performance of systems precoding over several cooperative frames.

6.5.2 Superiority of Multiple-frame Precoding in Error-floor Regions

To confirm the optimality of the proposed multiple-frame precoding scheme in error-floor regions, Fig. 6.7 shows the BER performances after 5 iterations of the systems precoding
Fig. 6.6 BER performances after 5 iterations of the NAF-BICM-ID system with $N = 2$ and different $4 \times 4$ precoders.

over two frames with the following $4 \times 4$ matrices: i) $I_4$, no precoder; ii) $G_{st}$ in (6.31), which is the optimal precoder for uncoded and coded systems over single-antenna fading channels [41, 50]; iii) $G_{\pi/6 Bk} = \text{bdiag} \left( G_{\pi/6}, G_{\pi/6} \right)$ with $G_{\pi/6} = \begin{pmatrix} 1 & e^{j\pi/6} \\ 0 & 0 \end{pmatrix}$, which is equivalent to the optimal single-frame precoder for the system using QPSK [111]; and iv) the proposed $G(\varphi)$ in (6.30), which belongs to the optimal class in (6.25) with the angles for good convergence behavior $\varphi_1 = 0.15$, $\varphi_2 = 0.30$ and $\varphi_3 = 0.62$ as derived in [118]. First, it is observed that the analytical bounds in Fig. 6.7 are again tight for all systems under consideration. Furthermore, it can be seen from Fig. 6.7 that precoding over two frames as in the case of $G_{st}$ and $G(\varphi)$ provides a higher diversity order than precoding over a single frame (i.e., $G_{\pi/6 Bk}$) or no precoding at all. It can be easily verified from (6.19) that both $G_{st}$ and $G(\varphi)$ achieve full cooperative diversity. More importantly, the system using $G(\varphi)$ greatly outperforms the systems using the other precoders. In particular, the proposed optimal precoder presents coding gains of $4.7$ dB, $1.7$ dB, and $1.2$ dB over $I_4$, $G_{\pi/6 Bk}$, and $G_{st}$, respectively, at the BER level of $10^{-5}$. These observations are in agreement with the analysis in Section 6.3.
Fig. 6.7 BER performances after 5 iterations of the NAF-BICM-ID system with $N = 2$ using the $I_4$, $G_{st}$, $G_{\pi/6 Bk}$ and $G(\phi)$ precoders.

The effect of the rotation angles on the system performance can be observed in Fig. 6.8, where the BERs after 1 and 5 iterations are shown for the systems with $N = 2$ using $G(\varphi)$ and $G_d = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & e^{j\varphi_1} & e^{j\varphi_2} & e^{j\varphi_3} \\ 0 & 0 & 0 & 0 \\ 1 & e^{j\varphi_1} & e^{j\varphi_2} & e^{j\varphi_3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$. The same angles from [118] are used for both precoders. Although the two precoders belong to the optimal class in (6.25) and use the same angles, it can be seen from Fig. 6.8 that the performance of the system with $G_d$ cannot converge to the bound after 5 iterations. This is because the value of (6.28) is greatly reduced when the two precoding vectors are in parallel. Fig. 6.8 highlights the importance of the rotation angles for good convergence behavior, and verifies that the proposed pragmatic approach indeed provides a reasonably good convergence performance.
Finally, to demonstrate the advantage of multiple-frame precoding in near-capacity regions, Fig. 6.9 shows the BER performance after 50 decoding iterations of the systems using the $G_{st}$ precoder, the multi-D mapping from [116], and the rate-2/3 and rate-3/4 punctured convolutional codes mentioned in Section 6.4. A 160000-length bit-interleaver is used for the simulations in Fig. 6.9. For comparison, we also plot the ergodic capacity with Gaussian inputs and the constrained capacity with QPSK. It can be seen from Fig. 6.9 that the proposed system can achieve a BER of $10^{-5}$ or lower at $E_b/N_0 = 5.15$ dB and $E_b/N_0 = 6$ dB for rate-2/3 and rate-3/4 systems, respectively, which are only 1.55 dB-1.95 dB away from the capacity with Gaussian inputs. Compared with the constrained capacities with QPSK, which are at $E_b/N_0 = 5.6$ dB and $E_b/N_0 = 7$ dB, we achieve significant gains of 0.45 dB and 1 dB for rate-2/3 and rate-3/4 systems, respectively. These results clearly match well to the analysis made by EXIT chart. Note that for all systems, the BER performance converges to the asymptotic error bound. However, it only happens in the error-floor region at high powers as discussed earlier.
Fig. 6.9  BER performances with 50 iterations of the NAF-BICM-ID system using the $G_{st}$ precoder, the multi-D mapping from [116], and the rate-2/3 and rate-3/4 outer codes.
6.6 Chapter Summary

This chapter considered a precoding scheme over multiple cooperative frames for NAF-BICM-ID systems. A tight union bound on the BER was first derived for the proposed system. It was then shown that multiple-frame precoding can provide significant benefits with respect to both diversity function in the error-floor region and capacity-approaching performance in the turbo pinch-off region. Specifically, in the error-floor region, an optimal class of precoders with respect to both diversity and coding gains was developed for independent and multi-D mapping. This class indicated that the source and relay should transmit orthogonally for best asymptotic diversity and coding gains. A pragmatic approach was then proposed to find good rotation angles in the case of independent mapping.

In the turbo pinch-off area, by concatenating the proposed system with a simple outer code and applying multi-D mapping, we demonstrated that one can operate below the traditional capacity with QPSK and approach close to the ergodic capacity with Gaussian inputs. The multiple-frame precoding techniques can therefore be utilized to exploit the benefits offered by NAF relaying.
Chapter 7

Capacity of Full-Duplex AF Relaying under Residual Self-interference\(^1\)

In all previous chapters, we addressed the capacity and code designs for several HD relaying protocols. We now turn our attention to FD relaying schemes. As explained in Chapter 1, pioneering works on FD relaying concentrated on the ideal scenario in which the relay is able to transmit and receive simultaneously without any self-interference. Since self-interference cannot be completely mitigated in practice, this ideal assumption overestimated the benefits of FD relaying. As such, current studies on FD relaying take the effect of residual self-interference into account [58, 123–134]. In particular, by considering some imperfection in the cancellation process (e.g., imperfect channel state information, imperfect knowledge of the transmitted signal), most of these works assume that the variance of the residual self-interference is proportional to the average transmitted power [58, 123–132, 134]. Under this assumption, the performance of FD relaying has been shown to be severely degraded in terms of rate and reliability, specially in high power regions. For instance, the achievable rate of a FD DHAF system was investigated in [123, 125] along with break-even boundaries between HD and FD modes. From the boundaries in [123, 125], HD is preferred in high relay power regions when full power allocation is applied. To control the self-interference, optimal power allocation strategies were then derived in closed-form in [123, 125], where it was observed that full power at the FD relay is not necessarily optimal.

Recently, experimental results in [30] suggest that the previous assumption on the vari-

\(^1\)Parts of this chapter have been published in the IEEE Wireless Communications Letters [122].
ance of the interference might only correspond to the worst-case scenario. Specifically, it was observed in [30] that the variance of the residual self-interference after analog and/or digital cancellation is proportional to the $\lambda$-th power of the average transmitted power ($0 \leq \lambda \leq 1$), where $\lambda$ indicates the quality of the self-interference cancellation technique in use. Similar trends for other mitigation prototypes can also be observed in [35]. Different from previous theoretical works that focused on $\lambda = 1$, the experimental results in [30, 35] indicate that $\lambda$ is often less than one. The main reason for this behavior is that it is easier to estimate the self-interference as its power increases, which results in more accurate subtraction and an increase in the average amount of cancellation. While the empirically-based model in [30] might not represent reality perfectly, it provides a better match to the residual self-interference and allows more flexibility to evaluate different scenarios. Therefore, more benefits of FD relaying can be expected under this model. In addition, it is clear that the optimal allocations derived earlier in [123, 125] no longer hold true.

This chapter investigates the capacity and respective optimal power allocation strategy for a FD DH system without direct link and under the residual self-interference model in [30]. The focus is on the static scenario similar to Chapter 3. Both individual and global power constraints are considered. First, the maximization of the achievable rates is shown to be a quasiconcave optimization problem [73, Ch.3.4] for $0 \leq \lambda \leq 1$. Given that the derivatives of the mutual information are non-linear when $0 < \lambda < 1$, closed-form solutions to the power allocation as in the case of $\lambda = 1$ in [123, 125] cannot be obtained. As such, bisection is proposed to obtain the optimal power allocation schemes. The capacity and optimal power strategies are then investigated in different high power regions. Specifically, when the relay has a large power constraint, we apply the method of dominant balance [75, Ch.3.4] to show that using full power at the relay is detrimental to the achievable rate. Following a similar approach, the multiplexing gain of the FD system using the optimal power allocation is shown to be $1/[1 + \lambda]$ when both the source and relay have a large power constraint. The FD DH scheme is finally compared to the HD system. Analytical and simulation results reveal that FD is advantageous when the source has a large power constraint and the relay does not, or when both the source and relay have a large constraint and $\lambda < 1$.

The rest of the chapter is organized as follows. Section 7.1 first introduces the FD system under consideration. The optimal power allocation schemes at the source and relay are derived in Section 7.2 under individual and global power constraints. Asymptotic
7 Capacity of FD AF Relaying under Residual Self-interference

analysis in different high power regions is carried in Section 7.3. Simulation examples are then provided in Section 7.4 along with comparisons to HD systems. Concluding remarks are finally drawn in Section 7.5.

7.1 Problem Formulation

In this chapter, we consider the static scenario similar to Chapter 3. In this scenario, the channel gains $\mathbf{h} = [h_1, h_2]$ are not time-varying so that full channel knowledge can be acquired at all nodes. Since full CSI is available, the CI coefficient $b = \sqrt{z_2/[\alpha_1 q_1 P_s + N_r + V]}$ in (2.40) is used. The FD DHAF system with input-output relation in (2.39) is adopted. In particular, we assume that the direct source-destination link is under deep shadowing so that $h_0 = 0$ in (2.39).

7.1.1 Residual Self-interference Model

The distribution of the self-interference after several stages of cancellation is unknown in practice. Here, similar to previous works, we assume that the residual self-interference is zero-mean, additive and Gaussian as $\mathcal{CN}(0, V)$. The Gaussian assumption might hold in reality due to the various sources of imperfections in the cancellation process (i.e., due to the central limit theorem). If the interference is not Gaussian, this assumption can be considered as the worst-case scenario in terms of achievable rate [135]. In addition, based on the experimental results in [30], the variance of the residual self-interference is modeled as

$$V = \beta (z_2 P_r)^\lambda,$$

(7.1)

where $z_2 P_r$ is the average power transmitted by the relay; and $\beta$ (given in units of [Watts]$^{1-\lambda}$) and $\lambda$ ($0 \leq \lambda \leq 1$) are constants that reflect the quality of the selected cancellation technique and can be found in [30, 35]. Smaller values of $\beta$ and $\lambda$ represent a better self-interference cancellation quality\(^2\). Note that the model in (7.1) is general in that it includes two important scenarios: the optimistic case in which the interference variance is simply a constant and not a function of transmitted power ($\lambda = 0$) [34, 35], and the pessimistic one in which the variance increases linearly with transmitted power.

\(^2\)It should be noted from the experimental results in [30] that $\beta$ is in fact a function of $\lambda$, i.e., $\beta = \beta(\lambda)$.\(^2\)
(λ = 1) [58, 123–126, 128–132, 134]. As we will show in this and the following chapter, the value of λ plays a crucial role in the performance of FD relay systems.

7.1.2 Achievable Rate and Capacity Formulation

Substituting the amplification coefficient \( b \) and the self-interference \( V \), the achievable rate in (2.41) can be written when \( \alpha_0 = 0 \) as

\[
I_{FD|h} = \log \left( 1 + \frac{q_1 z_2 \gamma_1 \gamma_2}{q_1 \gamma_1 + z_2 \gamma_2 + z_2^2 \gamma_3 + z_2^{1+\lambda} \gamma_3 + 1} \right) = \log[1 + f(q_1,z_2)],
\]

(7.2)

where \( \gamma_1 = [P_s \alpha_1]/N_r \), \( \gamma_2 = [P_r \alpha_2]/N_d \), and \( \gamma_3 = [\beta P_t^\lambda]/N_r \). The objective of this chapter is to maximize the above rate. First, in the individual power constraint scenario, we assume that \( q_1 \leq q_s \) and \( z_2 \leq z_r \). From (2.39), the power constraints at \( S \) and \( R \) are \( q_s P_s \) and \( z_r P_r \), respectively. The constants \( q_s \) and \( z_r \) might be set to one to have constraints of \( P_s \) and \( P_r \). Similar to [61, 123], we also consider the global constraint scenario with \( P_s = P_r = P_t \) and \( q_1 + z_2 \leq q_t \), where \( q_t P_t \) is the joint power constraint. The capacity of the FD DH system can be obtained by solving the following optimization problem:

\[
C_{FD} = \max_{q_1, z_2 \geq 0} I_{FD|h}, \quad \text{s.t.} \quad \begin{cases} 
q_1 \leq q_s, z_2 \leq z_r & \text{(indiv.)} \\
q_1 + z_2 \leq q_t & \text{(joint)}.
\end{cases}
\]

(7.3)

The rate in (7.2) is in general not a concave function of \( q = [q_1,z_2] \) for any \( 0 \leq \lambda \leq 1 \). However, as shown in the next section, the problem in (7.3) is quasiconcave [73, Ch.3.4] and has a unique global maximizer \( q^* = [q_1^*, z_2^*] \).

7.2 Capacity and Optimal Power Allocation

In this section, we derive the capacity in (7.3) along with the optimal power allocation schemes. Since \( \log(\cdot) \) is monotonically increasing, (7.3) can be equivalently solved by maximizing \( f(\cdot,\cdot) \) in (7.2). In the following, we first consider per-node constraints before extending the results to joint constraints.

**Proposition 7.1.** The optimal power allocation scheme under individual constraints is
given by
\begin{align*}
q_1^* &= q_s, \\
 z_2^* &= \begin{cases} 
 z_r, & \lambda = 0 \\
 \min \{ p_1, z_r \}, & 0 < \lambda < 1 \\
 \min \left\{ \sqrt{\frac{a_1z_2 + 1}{\gamma_2\gamma_3}}, z_r \right\}, & \lambda = 1.
\end{cases}
\end{align*}
(7.4)

Proof. Under individual constraints, it can be first shown that \( \partial f(q_1, z_2)/\partial q_1 > 0 \). Full power at \( S \) is hence optimal and \( q_1^* = q_s \). To find the optimal power at \( R \), we take the derivative of \( f(q_s, z_2) \) as

\[ \frac{\partial f(q_s, z_2)}{\partial z_2} = \frac{q_s\gamma_1\gamma_2 P_1(z_2, \lambda)}{\left( z_2^\lambda \gamma_3 + q_s\gamma_1 + z_2 \gamma_2 + z_2^{1+\lambda} \gamma_2 \gamma_3 + 1 \right)^2}, \]
(7.5)

where

\[ P_1(z_2, \lambda) = a_1 z_2^{1+\lambda} + b_1 z_2^\lambda + c_1, \]

with \( a_1 = -\lambda \gamma_2 \gamma_3 \), \( b_1 = \gamma_3 (1 - \lambda) \) and \( c_1 = q_s \gamma_1 + 1 \). Consider the following subcases according to \( \lambda \): i) When \( \lambda = 0 \), \( P_1(z_2, 0) > 0 \) and full power at \( R \) is optimal \( z_2^* = z_r \). ii) When \( 0 < \lambda < 1 \), it can be easily shown that \( P_1(0, \lambda) > 0 \) and \( \lim_{z_2 \to +\infty} P_1(z_2, \lambda) < 0 \).

By equating the derivative of \( P_1(z_2, \lambda) \) to zero, it has a single stationary point at \( z_2 = [1 - \lambda]/[\gamma_2(1 + \lambda)] > 0 \). \( P_1(z_2, \lambda) \) then has a single strictly positive root. Denote this root as \( p_1 \) if it is feasible, i.e., \( z_2^* = \min\{p_1, z_r\} \). Given that \( P_1(z_2, \lambda) \) is highly non-linear when \( 0 < \lambda < 1 \), \( p_1 \) cannot be obtained in closed-form and bisection is needed. Note that the convergence of bisection is guaranteed due to the quasiconcavity of the problem. iii) When \( \lambda = 1 \), \( P_1(z_2, 1) = -z_2^2 \gamma_2 \gamma_3 + c_1 \) is a quadratic polynomial. In this case, \( f(q_s, z_2) \) is again quasiconcave and the maximizer has been derived in [125] as \( z_2^* = \min\{\sqrt{c_1/[\gamma_2 \gamma_3]}, z_r\} \).

From these subcases, the optimal scheme is given by (7.4).

Proposition 7.2. The optimal power allocation scheme under global constraints is given
Consider the following subcases: i) When \( \lambda = 0, q_1 + z_2 = q_t \) since \( \partial f(q_1, z_2) / \partial q_1 > 0 \). Taking the derivative along the line \( q_1 + z_2 = q_t \),

\[
\frac{\partial f(q_1, q_t - q_1)}{\partial q_1} = \frac{\gamma_1 \gamma_2 P_2(q_1, \lambda)}{[(q_t - q_1) \lambda \gamma_3 + q_1 \gamma_1 + (q_t - q_1) \gamma_2 + (q_t - q_1) \lambda + 1] \gamma_2 \gamma_3 + 1}.
\]

where

\[
P_2(q_1, \lambda) = q_t - 2q_1 - q_t^2 \gamma_1 + \gamma_2 (q_1 - q_t)^2 + \gamma_3 (q_t - q_1)^3 [q_t - 2q_1 + \gamma_2 (q_1 - q_t)^2 + \lambda q_1 (q_t \gamma_2 - q_1 \gamma_2 + 1)].
\]

Consider the following subcases: i) When \( \lambda = 0, P_2(q_1, 0) = A_0 q_1^2 + B_0 q_1 + C_0 \) is a quadratic polynomial with \( A_0 = \gamma_2 - \gamma_1 + \gamma_2 \gamma_3, B_0 = -2(q_t \gamma_2 + 1)(\gamma_3 + 1) \), and \( C_0 = q_t (\gamma_3 + 1)(q_t \gamma_2 + 1) \). Given that the self-interference is not a function of \( z_2 \) when \( \lambda = 0 \), (7.3) is equivalent to the rate maximization of a HD system with a raised noise floor. This was solved in [61] and the solution is \( q_1^* = [-B_0 - \sqrt{B_0^2 - 4A_0 C_0}]/[2A_0] \) when \( A_0 \neq 0 \), or \( q_1^* = -C_0/B_0 = q_t/2 \) when \( A_0 = 0 \). ii) When \( 0 < \lambda < 1 \), it is easy to show that \( P_2(0, \lambda) > 0 \) and \( P_2(q_t, \lambda) < 0 \). Similar to the previous subsection, \( P_2(q_1, \lambda) \) has a single positive root in the range \([0, q_t] \) and the problem is again quasiconcave. Let \( p_2 \) be this root. As discussed before, a closed-form solution for \( q_1^* = p_2 \) cannot be obtained and bisection is required. iii) When \( \lambda = 1, P_2(q_1, 1) = A_1 q_1^2 + B_1 q_1 + C_1 \) is a quadratic polynomial with \( A_1 = \gamma_2 - \gamma_1 + \gamma_3 + q_t \gamma_2 \gamma_3, B_1 = -2(q_t \gamma_3 + 1)(q_t \gamma_2 + 1) \), and \( C_1 = q_t (q_t \gamma_3 + 1)(q_t \gamma_2 + 1) \). The problem is also quasiconcave with the solution \( q_1^* = [-B_1 - \sqrt{B_1^2 - 4A_1 C_1}]/[2A_1] \) when \( A_1 \neq 0 \), or \( q_1^* = -C_1/B_1 = q_t/2 \) when \( A_1 = 0 \) as derived in [123]. Combining all these cases, we have (7.6).
7.3 Asymptotic Analysis

In the previous section, we derived optimal power allocation strategies that maximize the achievable rate in (7.2) under individual and global power constraints. In this section, we provide further insights on the behavior of the capacity in (7.3) and the optimal schemes in (7.4) and (7.6) in different asymptotically high power regions. Specifically, we will consider four different asymptotic cases as follows:

Large Source Power, Fixed Relay Power

Consider first the case under individual constraints where \( P_s \rightarrow \infty \) while \( P_r \) remains fixed. When \( 0 < \lambda < 1 \), applying the method of dominant balance [75, Ch.3.4] to \( P_1(z_2, \lambda) \), it can be shown that \( O(z_2^{1+\lambda}) = O(q_s\gamma_1) = O(P_s) \). Therefore, \( p_1 = O(P_s^{1/[1+\lambda]}) \) and \( z_2^* = \min\{p_1, z_r\} = \min\{O(P_s^{1/[1+\lambda]}), z_r\} = z_r \). When \( \lambda = 1 \), \( z_2^* = \min\{\sqrt{(q_s\gamma_1 + 1)/(\gamma_2\gamma_3)}, z_r\} = \min\{O(P_s^{1/2}), z_r\} = z_r \). Thus, full power at both nodes, \( q_1^* = q_s \) and \( z_2^* = z_r \), is asymptotically optimal for \( 0 \leq \lambda \leq 1 \) and the capacity \( C_{FD} \rightarrow \log(1 + z_r\gamma_2) \) approaches a constant.

Large Relay Power, Fixed Source Power

Consider now the per-node constrained case where \( P_r \rightarrow \infty \) while \( P_s \) remains fixed. When \( \lambda = 0 \), \( z_2^* = z_r \) and \( C_{FD} \rightarrow \log (1 + [q_s\gamma_1]/[1 + \gamma_3]) \). When \( 0 < \lambda < 1 \), applying the method of dominant balance to \( P_1(z_2, \lambda) \), \( O(z_2^{1+\lambda}\gamma_2\gamma_3) = O(z_2^2\gamma_3) = O(q_s\gamma_1 + 1) = O(1) \). Therefore, \( z_2^* = p_1 = O(P_r^{-1}) \) and the capacity also approaches a constant. When \( \lambda = 1 \), \( z_2^* = \sqrt{(q_s\gamma_1 + 1)/(\gamma_2\gamma_3)} = O(P_r^{-1}) \) and the capacity approaches a constant as well. Note that when \( \lambda > 0 \), \( z_2^*P_r = O(1) \). Hence, as the power constraint at \( R \) approaches infinity, the power transmitted by the relay saturates to a certain value to control the self-interference.

Large Source and Relay Power

Consider the individual constraint case with \( P_s = P_r \rightarrow \infty \). When \( \lambda = 0 \), \( q_1^* = q_s \), \( z_2^* = z_r \) and

\[
C_{FD} = \log \left( \frac{q_1^*z_2^*\gamma_1\gamma_2 + O(P_s)}{q_1^*\gamma_1 + z_2^*\gamma_2 + z_2^*\gamma_2\gamma_3 + O(1)} \right) = \log(P_s) + O(1).
\]

The capacity can then be written as \( C_{FD} = m \cdot \log(P_s) + O(1) \) where \( m = 1 \) is the multiplexing gain. When \( 0 < \lambda < 1 \), applying the dominant balance method, \( O(z_2^{1+\lambda}\gamma_2\gamma_3) = \)
$O(q_s \gamma_1) = O(P_s)$. The root $p_1 = z_2^* = O(P_s^{-\lambda/[1+\lambda]})$ and

$$C_{FD} = \log \left( \frac{q_1^* z_2^* \gamma_1 \gamma_2 + O(P_s)}{q_1^* \gamma_1 + (z_2^*)^{1+\lambda} \gamma_2 \gamma_3 + O(P_s^{1/1+\lambda})} \right) = \frac{1}{1 + \lambda} \log(P_s) + O(1),$$

with the multiplexing gain $m = 1/[1+\lambda]$. When $\lambda = 1$, $q_1^* = q_s$, $z_2^* = \sqrt{(q_s \gamma_1 + 1)/(\gamma_2 \gamma_3)} = O(P_s^{-1/2})$ and

$$C_{FD} = \log \left( \frac{q_1^* z_2^* \gamma_1 \gamma_2 + O(P_s)}{q_1^* \gamma_1 + (z_2^*)^{2} \gamma_2 \gamma_3 + O(P_s^{1/2})} \right) = \frac{1}{2} \log(P_s) + O(1),$$

with $m = 1/2$. Note that the power transmitted by $R$, $z_2^* P_s$, only grows as $O(P_s^{1/[1+\lambda]})$ when $\lambda > 0$, i.e., full relay power is not optimal.

**Large Global Power**

Finally, consider the joint constraint with $P_t \to \infty$. Following a similar approach as above, when $\lambda = 0$, $q_1^* = O(1)$, $z_2^* = O(1)$, and $C_{FD} = \log(P_t) + O(1)$ with a multiplexing gain $m = 1$. When $0 < \lambda < 1$, $q_1^* = p_2 = q_t - O(P_t^{-\lambda/[1+\lambda]})$ and $z_2^* = O(P_t^{-\lambda/[1+\lambda]})$. The capacity is then given by $C_{FD} = [1/(1+\lambda)] \log(P_t) + O(1)$ with $m = 1/[1+\lambda]$. When $\lambda = 1$, $A_1 > 0$, $q_1^* = q_t - O(P_t^{-1/2})$ and $z_2^* = O(P_t^{-1/2})$. Hence, $C_{FD} = [1/2] \log(P_t) + O(1)$ and $m = 1/2$.

The behavior of the capacity and optimal power allocation schemes for the above asymptotic cases is summarized in Table 7.1.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$q_1^* P_s$</th>
<th>$z_2^* P_r$</th>
<th>$m$</th>
<th>Preferred protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>High $P_s$, fixed $P_r$</td>
<td>$O(P_s)$</td>
<td>$O(1)$</td>
<td>$-$</td>
<td>FD</td>
</tr>
<tr>
<td>High $P_r$, fixed $P_s$</td>
<td>$O(1)$</td>
<td>${O(P_r) \ (\lambda = 0)}$</td>
<td>$-$</td>
<td>HD $(\lambda &gt; 0)$</td>
</tr>
<tr>
<td>High $P_s = P_t$</td>
<td>$O(P_s)$</td>
<td>$O(P_s^{1/[1+\lambda]})$</td>
<td>$1/[1+\lambda]$</td>
<td>FD $(\lambda &lt; 1)$</td>
</tr>
<tr>
<td>High $P_t$</td>
<td>$O(P_t)$</td>
<td>$O(P_t^{1/[1+\lambda]})$</td>
<td>$1/[1+\lambda]$</td>
<td>FD $(\lambda &lt; 1)$</td>
</tr>
</tbody>
</table>
7.4 Illustrative Examples and Comparisons to HD Schemes

In this section, we provide simulation results to verify our theoretical analysis. Comparisons between the FD and HD DH schemes are also carried. For all simulations, unit noise power is considered $N_r = N_d = 1$. To concentrate on the effect of $\lambda$, we set $\beta = 1$ in (7.1). In addition, to keep the same average power constraints between HD and FD, $q_s = z_r = q_t = 1$ for the FD system and $q_{s, \text{HD}} = z_{r, \text{HD}} = q_{t, \text{HD}} = 2$ for the HD one. Besides the optimal power allocations derived in (7.4) and (7.6), the full ($q_1 = q_s$, $z_2 = z_r$) and uniform ($q_1 = z_2 = q_t/2$) schemes are also considered. Note that full power is optimal for HD with per-node constraints, whereas uniform allocation is optimal for HD under joint constraints with $\gamma_1 = \gamma_2$. Also note that the rate of the HD system can be obtained by setting $\gamma_3 = 0$ in (7.2) and pre-multiplying by a factor of 1/2, i.e., the mutual information in (2.26).

Fig. 7.1 shows the achievable rates of the DH system against $P_s$ for a fixed $P_r = 5$ dB and different values of $\lambda$. The performance of the ideal FD system without self-interference ($V = 0$) and using the optimal power scheme is also plotted as a benchmark. First, observe that full power is asymptotically optimal for the FD systems, as expected from Section 7.3. When compared to the HD system with full power, it can be seen that the FD systems perform better in high source power regions. This is because in this asymptotic case, it is easy to show that $C_{\text{FD}} \rightarrow \log(1 + z_r \gamma_2) > C_{\text{HD}} \rightarrow [1/2] \log(1 + 2z_r \gamma_2)$. Specifically, the FD systems with $\lambda = [0, 0.5, 1]$ outperform the HD one for $P_s$ values greater than 6.5, 9 and 11.5 dB, respectively, and provide an asymptotic 1.4× gain.

Fig. 7.2 shows the achievable rates against $P_r$ for a fixed $P_s = 5$ dB. First, note that the proposed allocation schemes in (7.4) outperform full power for all values of $\lambda$. This is in agreement with Section 7.2. As discussed in Section 7.3, Fig. 7.2 indicates that using full power at the relay hurts the rate in high relay power regions when $\lambda > 0$. This is because full power leads to $I_{\text{FD}} \rightarrow \log(1 + O(P_r^{-\lambda})) \rightarrow 0$. In addition, it can be seen that the HD system outperforms all FD schemes. This is because when $\lambda = 0$, it can be shown that $C_{\text{HD}} \rightarrow [1/2] \log(1 + 2q_s \gamma_1) > C_{\text{FD}} \rightarrow \log(1 + [q_s \gamma_1]/[1 + \gamma_3])$ when $2\gamma_3(1 + \gamma_3) > q_s \gamma_1$, which is the case for the parameters in Fig. 7.2.

The rates of the DH system are plotted in Fig. 7.3 against $P_s = P_r$. To study the multiplexing gains and for clarity of the figure, the rates are normalized by $\log(1 + P_s)$. As before, note that the normalized rates of the FD systems using the allocations from Sec-
Fig. 7.1 Achievable rates of DH systems against $P_s$ ($\alpha_1 = \alpha_2 = 1$, $\lambda = [0, 0.5, 1]$, $P_r = 5$ dB).

Fig. 7.2 Achievable rates of DH systems against $P_r$ ($\alpha_1 = \alpha_2 = 1$, $\lambda = [0, 0.3, 0.5, 0.7, 1]$, $P_s = 5$ dB).
tion 7.2 outperform those with full power. Moreover, the normalized rates with the optimal allocation approach the multiplexing gain values of $1/[1+\lambda]$, which agrees with Section 7.3. Following a similar analysis, it can be shown that $m = [1 - \lambda]$ for the FD systems with full power. As $C_{HD} = [1/2] \log(P_s) + O(1)$, the FD systems with the optimal and full allocation present a higher multiplexing gain than HD when $\lambda < 1$ and $\lambda < 0.5$, respectively. In particular, the FD systems using the optimal strategy with $\lambda = [0, 0.3, 0.5, 0.7]$ outperform the HD scheme for $P_s$ greater than 3.5, 5, 8 and 17.5 dB, respectively, and provide gains of 1.6, 1.3, 1.2 and 1.0× at $P_s = 20$ dB. However, the capacity of the FD system with $\lambda = 1$ is inferior for the parameters in Fig. 7.3.

![Fig. 7.3](image.png)

**Fig. 7.3** Normalized rates of DH systems against $P_s = P_r$ ($\alpha_1 = \alpha_2 = 2$, $\lambda = [0, 0.3, 0.5, 0.7, 1]$).

Similar results can be observed for the joint constraint scenario in Fig. 7.4, where the normalized rates are shown against $P_t$. As in the previous figures, note that the allocation in (7.6) outperforms the uniform scheme for all FD systems. In this scenario, the FD systems with optimal allocation and $\lambda < 1$ are advantageous over HD when $P_t > 8.5$ dB, while the FD system with $\lambda = 1$ remains inferior. The dominant protocols (HD or FD) for the scenarios considered in this section are summarized in Table 7.1.
7 Capacity of FD AF Relaying under Residual Self-interference

Fig. 7.4 Normalized rates of DH systems against $P_t$ ($\alpha_1 = \alpha_2 = 2$, $\lambda = [0, 0.3, 0.5, 0.7, 1]$).

7.5 Chapter Summary

In this chapter, we derived the capacity and optimal allocation for a FD DHAF system under residual self-interference. The optimization problems were first shown to be quasiconcave and bisection was proposed to find the root of the non-linear derivative. The capacity and power allocation were then investigated in high power regions. Specifically, we showed that full power at the relay is not necessarily optimal and that the multiplexing gain is $1/[1 + \lambda]$. Our analysis also revealed that the FD system is able to outperform its HD counterpart when the source uses a large power, and the relay has either a fixed or a large power constraint.
Chapter 8

Error Performance of Full-Duplex AF Relaying under Residual Self-interference

In Chapter 7, we studied the capacity of FD relaying assuming Gaussian codebooks at the source. We now consider the error performance of FD systems using finite constellations. As explained in Chapter 7, most works on FD relaying with residual self-interference assume that the variance of the interference is proportional to the average transmitted power. Moreover, studies in the literature have only considered the non-cooperative dual-hop approach, in which the direct link from source to destination is either assumed to be a source of interference [125, 127–129, 132], or completely ignored [123, 125, 126, 131, 133, 134]. While these non-cooperative FD schemes allow the source to transmit continuously, their error performance has been shown to be significantly affected in the above residual self-interference model. For example, it has been shown via BER and outage analysis that the FD systems present a zero diversity order [129, 131, 132, 134], i.e., the outage or BER curves present an error floor. Such floor persists even with relay [131] or antenna [134] selection.

As noted in Chapter 7, recent experimental results in [30, 35] suggest that the variance of the residual self-interference is proportional to the $\lambda$-th power of the average transmitted power, where $\lambda$ is less than one. This means that previous works have only considered the

---

1 Parts of this chapter have been presented at the 2014 IEEE International Conference on Communications in Sydney, Australia [136]; and have been published in the IEEE Journal on Selected Areas in Communications [137].
the worst-case scenario of $\lambda = 1$. In addition, we know from the studies on cooperative HD protocols that the direct link is important in terms of both rate and diversity advantages. Therefore, it is expected that similar benefits can be exploited in FD systems by making use of the direct link. However, to our knowledge, the idea of cooperative relaying has not been explored in FD systems under self-interference.

Inspired by the above observations, this chapter investigates the error and diversity performance of FD AF relay systems under the residual self-interference model in [30]. We consider the cooperative FD system that makes use of the direct link for diversity improvement, and the DH system without direct link. For the system with direct link, we focus on the cooperative FD LR protocol introduced in Section 2.5, where transmission is carried over $L$ consecutive periods. This protocol has only been studied in terms of achievable rate under no self-interference [56, 57] and includes the HD NAF scheme as a special case when $L = 2$. We shall examine the error and diversity performances of both the uncoded system, and the coded system under the framework of BICM similar to Chapter 6.

As a first step, we derive closed-form expressions of the uncoded PEP for the LR and DH protocols. The derived expressions can be used not only to analyze the uncoded systems, but also to obtain tight bounds to the BER of the coded ones. Asymptotic PEP and BER expressions in high transmission power regions are then obtained. Based on these expressions, it is shown that FD LR systems can achieve the same diversity function as their HD counterparts [109, 111, 138] for any $0 \leq \lambda \leq 1$ as long as a suitable precoder at the source is applied. Different from previous works where the direct link is treated as interference, a non-zero diversity order is attained and no error floor is observed despite the self-interference in FD. By further analyzing the coding gain of the LR system when $\lambda > 0$, it is demonstrated that the source only needs to transmit a superposition of $L$ signals in the first time slot to maximize this gain. Equivalently, it turns out that using HD orthogonal relaying together with a superposition constellation is asymptotically optimal. This behavior is similar to that in HD relaying shown in [111, 138]. For DH systems, the asymptotic analysis indicates that their diversity functions are dependent on $\lambda$. In particular, the diversity order is a decreasing function of $\lambda$ for $0 < \lambda \leq 1$ and an error floor is observed only when $\lambda = 1$. The FD DH system must then sacrifice its diversity when $\lambda > 0$ to achieve a full rate of one transmitted symbol per period. Although HD relaying is asymptotically optimal for both relaying protocols, illustrative results show that FD relaying is advantageous at practical BER levels in good self-interference cancellation.
scenarios, i.e., when $\lambda$ is sufficiently small.

The rest of the chapter is organized as follows. The considered systems are introduced in Section 8.1. The uncoded PEP analysis for both relaying protocols is carried in Section 8.2 along with the derivation of tight BER bounds on the coded systems. Asymptotic PEP and BER expressions are derived in Section 8.3 followed by the diversity analysis. Illustrative examples are presented in Section 8.4. Finally, Section 8.5 concludes the chapter.

### 8.1 System Model

The general block diagram of the uncoded relay systems is shown in Fig. 8.1. First, for the uncoded FD LR protocol, the information sequence $\mathbf{u}$ is divided into groups of $Lm_c$ bits. Each group is mapped to a signal $\mathbf{s} = [s_1, \ldots, s_L]^\top$ in the complex $L$-D constellation $\Psi$. Each component $s_i$ is assumed to be in a 1-D unit-energy constellation $\Omega$, such as QPSK or QAM, of size $2^{m_c}$. We assume that the mapping $\xi$ from binary bits to symbols is implemented independently in $\Omega$ for each component $s_i$. The symbol $\mathbf{s}$ is then rotated by a $L \times L$ single-frame complex precoder $\mathbf{G}$ as $\mathbf{x} = \mathbf{Gs}$. As we demonstrate shortly and similar to Chapter 6, the use of such precoder provides diversity benefits for the LR system. Each precoded symbol $\mathbf{x}$ is then sent over $L$ periods according to the LR protocol described in (2.33) with $\alpha_l > 0$. At the destination, a maximum likelihood (ML) detector is applied on the received vector $\mathbf{y}$ to obtain the estimated information sequence $\mathbf{\hat{u}}$.

![Fig. 8.1](Uncoded LR and DH systems.)

For the uncoded FD DH protocol, $\mathbf{u}$ is divided into groups of $m_c$ bits which are mapped directly to $x_i = s_i \in \Omega$ using the mapping rule $\xi$. Given that space diversity is not exploited in DH relaying, no precoder is considered in this case. The signals $x_i$ are then transmitted using the DH protocol in (2.39), where the direct link is assumed to be heavily shadowed.
with $\alpha_0 = 0$ similar to Chapter 7. At the destination, ML decoding is applied on the received signal $y_i$ as in the LR case. Note that demodulation is carried symbol-by-symbol over $\Omega$ for the DH protocol, whereas LR requires demodulation over the $L$-D constellation $\Psi$. Note also that both the DH and LR protocols achieve a full rate of one transmitted symbol per period.

In the coded framework, the general block diagram of the LR and DH BICM-ID systems is similar to the one described in Fig. 6.1. Specifically, the information sequence $u$ in Fig. 8.1 is first encoded and then interleaved prior to modulation. At the destination, as illustrated in Fig. 6.1, the received sequence is decoded in an iterative manner using a MAP demodulator and a SISO channel decoder.

In this chapter, we consider the fading environment in which the transmitted codeword spans numerous channel realizations of $h = [h_0, h_1, h_2]$. Moreover, the destination possesses full CSI, whereas the relay has only statistical knowledge of the incoming $S$-$R$ link. Assuming that the relay only amplifies the signal received in the previous period, the FG amplification coefficients for the LR and DH protocols are given by (2.37) and (2.40) with $\Xi = \phi$, respectively. Recall that the covariance matrices of the signals transmitted by $S$ and $R$ in the LR protocol are given by $Q = \mathbb{E}[xx^\dagger] = GG^\dagger$ and $Z = \mathbb{E}[tt^\dagger]$. We assume that $\text{tr}(Q) = L$ and $\text{tr}(Z) = L - 1$ so that the source and relay have respectively an average power constraint of $P_s$ and $P_r$ per symbol period when active. The entries of $G$, $\{g_{i,k}\}$ ($1 \leq i, k \leq L$), must then satisfy the power constraint $\text{tr}(Q) = \text{tr}(GG^\dagger) = \sum_{i=1}^{L} \sum_{k=1}^{L} |g_{i,k}|^2 = L$. In a similar manner, we assume that $\mathbb{E}[|s_i|^2] = 1$ and $z_2 = 1$ for the DH protocol.

Finally, similar to Chapter 7, we adopt the residual self-interference model in [30]. Specifically, the residual self-interference is modeled as a complex Gaussian random variable with variance given by

$$V = \beta P_r^\lambda,$$

where $P_r$ is the average power transmitted by the relay when active, and $\beta$ and $\lambda$ ($0 \leq \lambda \leq 1$) are constants that can be found in [30, 35].
8.2 PEP Analysis and Tight BER Bounds

In this section, we first derive closed-form PEP expressions for the uncoded relay systems. Such derivation is also useful in developing tight BER bounds for the coded systems.

8.2.1 PEP for the LR System

The PEP is defined as the probability of deciding in favor of \( \bar{s} = [\bar{s}_1, \ldots, \bar{s}_L]^\top \) given that \( s \) was transmitted, \( \bar{s}, s \in \Psi \) and \( s \neq \bar{s} \). These two signal points correspond to the rotated symbols \( x = Gs \) and \( \bar{x} = [\bar{x}_1, \ldots, \bar{x}_L]^\top = G\bar{s} \). Assuming perfect CSI at \( D \), the PEP conditioned on \( h \) can be written from (2.33) as

\[
P(s \to \bar{s}|h) = Q\left( \sqrt{\frac{d^2(s, \bar{s}|h)}{2}} \right),
\]

(8.2)

where \( Q(\cdot) \) is the Gaussian \( Q \)-function in (2.3) and \( d^2(s, \bar{s}|h) \) is the squared Euclidean distance conditioned on \( h \):

\[
d^2(s, \bar{s}|h) = P_s \| K^{-1/2} H_{LR}(x - \bar{x}) \|^2 = P_s \| K^{-1/2} H_{LR} G(s - \bar{s}) \|^2.
\]

(8.3)

Following a similar approach as in Chapter 6 and assuming that only the previous symbol is forwarded by the relay, the signal component in (2.33) can be alternatively written as

\[
H_{LR}Gs = H_{LR}x = XTh_{01},
\]

(8.4)

where the \( 2 \times 1 \) \( h_{01} = [h_0, h_1]^\top \), the \( L \times 2 \)

\[
T = \begin{pmatrix}
1 & 0 \\
0 & \sqrt{P_r} h_2 b_{2,1} \\
\vdots & \vdots \\
0 & \sqrt{P_r} h_2 b_{L,L-1}
\end{pmatrix},
\]

(8.5)
and the $L \times L$

$$
X = \begin{pmatrix}
x_1 & 0 & \cdots & \cdots & 0 \\
x_2 & x_1 & 0 & \cdots & 0 \\
x_3 & 0 & x_2 & 0 & \cdots \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
x_L & 0 & \cdots & \cdots & x_{L-1}
\end{pmatrix}.
$$

(8.6)

The distance in (8.3) can then be expressed from (8.4) as

$$
d^2(s, \tilde{s}|h) = P_s h_{01}^\dagger T^\dagger U^\dagger K^{-1} U T h_{01},
$$

(8.7)

where $U = X - \tilde{X}$ and $\tilde{X}$ is given as in (8.6) but replacing $x_i$ by $\tilde{x}_i$. Applying the alternate representation of the Gaussian probability integral $Q(x) = \frac{1}{\sqrt{\pi}} \int_{0}^{\pi/2} \exp \left( -\frac{x^2}{[2 \sin^2 \theta]} \right) d\theta$ in (2.3) and from (8.7), the conditional PEP becomes

$$
\mathbb{P}(s \rightarrow \tilde{s}|h) = \frac{1}{\pi} \int_{0}^{\pi/2} \exp \left( -c_\theta h_{01}^\dagger T^\dagger U^\dagger K^{-1} U T h_{01} \right) d\theta,
$$

where $c_\theta = P_s / [4 \sin^2 \theta]$. Averaging over $h$, the unconditional PEP can be written as

$$
\mathbb{P}(s \rightarrow \tilde{s}) = \frac{1}{\pi} \int_{0}^{\pi/2} \Delta_\theta(s, \tilde{s}) d\theta,
$$

(8.8)

where

$$
\Delta_\theta(s, \tilde{s}) = \mathbb{E}_h \left[ \exp \left( -c_\theta h_{01}^\dagger T^\dagger U^\dagger K^{-1} U T h_{01} \right) \right].
$$

(8.9)

To solve the expectation in (8.9), we can first use the fact that given $a \sim \mathcal{CN}(0, \Phi)$ and a Hermitian matrix $A$, $\mathbb{E}[\exp(-a^H A a)] = 1 / \det(\mathbf{I} + \Phi A)$ [109]. Thus, we have

$$
\Delta_\theta(s, \tilde{s}) = \mathbb{E}_{h_2} \left[ \frac{1}{\det(I_2 + c_\theta \Phi_{01} T^\dagger U^\dagger K^{-1} U T)} \right],
$$

(8.10)
where \( \Phi_{01} = \begin{pmatrix} \phi_0 & 0 \\ 0 & \phi_1 \end{pmatrix} \). Given that the relay only forwards the previous symbol to the source, the covariance in (2.35) is written as \( K = \text{diag}(k) \) with the \( L \times 1 \) vector \( k \):

\[
k = [k_1, \ldots, k_L]^T
\]

\[
= [N_d, N_d + N_r \alpha_2 |b_{2,1}|^2, N_d + N_r \alpha_2 |b_{3,2}|^2 (V + N_r), \ldots, N_d + N_r \alpha_2 |b_{L,L-1}|^2 (V + N_r)]^T.
\]

Let the second column of \( UT \) be given by

\[
\tau = [\tau_1, \ldots, \tau_L]^T = [0, \sqrt{P_r h_2 b_{2,1} u_1}, \ldots, \sqrt{P_r h_2 b_{L,L-1} u_{L-1}}]^T,
\]

with \( u = [u_1, \ldots, u_L] = x - \bar{x} \). The determinant in (8.10) can then be written as

\[
\det(I_2 + c \Phi_{01} T^T U^T K^{-1} U T) = \det \left( 1 + c \phi_0 \sum_{l=1}^{L} |u_l|^2 k_l \begin{pmatrix} c \phi_0 \sum_{l=2}^{L} |\tau_l|^2 k_l & \begin{pmatrix} c \phi_1 \sum_{l=2}^{L} |\tau_l|^2 k_l \\ 1 + c \phi_1 \sum_{l=2}^{L} |\tau_l|^2 k_l \end{pmatrix} \end{pmatrix} \right)
\]

\[
= P_1(\alpha_2)/P_2(\alpha_2),
\]

where

\[
P_1(\alpha_2) = \left( \prod_{l=1}^{L} k_l \right) + c \phi_0 \left( \sum_{l=1}^{L} |u_l|^2 \prod_{m=1}^{L} k_m \right) + c \phi_1 \left[ k_1 + c \phi_0 |u_1|^2 \right] \left( \sum_{l=2}^{L} |\tau_l|^2 \prod_{m=2}^{L} k_m \right)
\]

\[
+ c^2 \phi_0 \phi_1 k_1 \left( \sum_{l=2}^{L} L_{m=2}^{L} m \neq l |u_l|^2 |\tau_m|^2 - \tau_l^* u_l \tau_m u_l^* \prod_{n=2}^{L} k_n \right),
\]

\[
P_2(\alpha_2) = \prod_{l=1}^{L} k_l.
\]

It can be shown from the definition of \( \tau \) and \( k \) that the functions \( P_1(\alpha_2) \) and \( P_2(\alpha_2) \) are polynomials in \( \alpha_2 \) of order \( L - 1 \). Let \( p_l \) be the roots of \( P_1(\alpha_2) \) \((1 \leq l \leq L - 1)\). Using partial fraction expansion,

\[
\frac{P_2(\alpha_2)}{P_1(\alpha_2)} = \omega_0 + \sum_{l=1}^{L-1} \frac{\omega_l}{\alpha_2 - p_l},
\]
where \( \omega_l \) are constants that depend on \( \tau \) and \( k \), and can be obtained from the partial fraction procedure for a given \( L \). The expectation in (8.9) can finally be solved as

\[
\Delta_\theta(s, \bar{s}) = \mathbb{E}_{\alpha_2} \left[ \frac{P_2(\alpha_2)}{P_1(\alpha_2)} \right] = \omega_0 + \sum_{l=1}^{L-1} \mathbb{E}_{\alpha_2} \left[ \frac{\omega_l}{\alpha_2 - p_l} \right] = \omega_0 + \frac{1}{\phi_2} \sum_{l=1}^{L-1} \omega_l J \left( \frac{-p_l}{\phi_2} \right), \tag{8.12}
\]

with \( J(x) = \exp(x)E_1(x) \) and the exponential integral \( E_1(x) \) as in (4.4). The PEP for the LR system can then be obtained from (8.8) by substituting \( \Delta_\theta(\cdot, \cdot) \) as in (8.12), where the integral over \( \theta \) can be easily calculated via numerical integrations. Note that the above derivation is applicable to any LR system in which the relay amplifies the previous symbol with an arbitrary constant-gain coefficient, i.e., the elements in the subdiagonal of \( B \) are not necessarily equal as in (2.37).

### 8.2.2 PEP for the DH System

We now consider the PEP for the DH protocol with \( h_0 = 0 \). In the DH case, the conditional PEP \( \mathbb{P}(s_i \rightarrow \bar{s}_i|h_1, h_2) \) can be written as in (8.2) with

\[
d^2(s_i, \bar{s}_i|h_1, h_2) = \frac{P_sP_r\alpha_1\alpha_2b^2|\epsilon_i|^2}{P_r\alpha_2b^2(N_r + V) + N_d},
\]

and \( \epsilon_i = s_i - \bar{s}_i \neq 0 \). Following a similar approach as in (8.8), the unconditional PEP is given by

\[
\mathbb{P}(s_i \rightarrow \bar{s}_i) = \frac{1}{\pi} \int_0^{\pi/2} \Delta_\theta(s_i, \bar{s}_i) \, d\theta, \tag{8.13}
\]

where

\[
\Delta_\theta(s_i, \bar{s}_i) = \mathbb{E}_{h_1, h_2} \left[ \exp \left( \frac{-P_sP_r\alpha_1\alpha_2b^2|\epsilon_i|^2}{4\sin^2 \theta \cdot [P_r\alpha_2b^2(N_r + V) + N_d]} \right) \right]. \tag{8.14}
\]

The expectation in (8.14) can then be solved similar to (8.10) as:

\[
\Delta_\theta(s_i, \bar{s}_i) = \mathbb{E}_{h_2} \left[ 1 + \frac{P_sP_r\phi_1\alpha_2b^2|\epsilon_i|^2}{4\sin^2 \theta \cdot [P_r\alpha_2b^2(N_r + V) + N_d]} \right]^{-1} = \mathbb{E}_{h_2} \left[ \frac{A_2\alpha_2 + B_2}{C_2\alpha_2 + B_2} \right],
\]
where $A_2 = 4 \sin^2 \theta \cdot P_r b^2 (N_r + V)$, $B_2 = 4 \sin^2 \theta \cdot N_d$ and $C_2 = P_r b^2 [4 \sin^2 \theta \cdot (N_r + V) + P_s |\epsilon_i|^2 \phi_1]$. Using partial fraction expansion,

$$\Delta_\theta (s_i, \bar{s}_i) = \frac{A_2 C_2 - B_2 A_2}{C_2} \mathbb{E}_{h_2} \left[ \frac{1}{\alpha_2 + (B_2 / C_2)} \right]$$

$$= \frac{A_2}{C_2} + \frac{B_2 C_2 - B_2 A_2}{\phi_2 C_2^2} J \left( \frac{B_2}{\phi_2 C_2} \right). \quad (8.15)$$

The PEP for the DH system is then obtained by substituting (8.15) in (8.13). Similar to the LR protocol, the above analysis is applicable to any constant-gain coefficient $b$.

### 8.2.3 Tight BER Bounds of BICM-ID Systems

The PEP expressions in (8.12) and (8.15) can be used not only to analyze uncoded systems, but also to provide tight bounds to the BER of the coded systems as was shown in Chapter 6. In particular, by applying the error-free feedback approach, the union bound on the BER of a LR BICM-ID system that employs a rate-$k_c/n_c$ outer code can be written as in (6.4) with $f(d, \Psi, \xi, G)$ upper bounded as in (6.11) but now with

$$\gamma_\theta (\Psi, \xi, G) = \frac{1}{L m_c^2 L m_c} \sum_{s \in \Psi} \sum_{k=1}^{L m_c} \Delta_\theta (s, p). \quad (8.16)$$

The bound to the BER of the coded LR system is then obtained by substituting (8.12), (6.11) and (8.16) into (6.4). A similar approach as in (6.4) can be followed for the coded DH BICM-ID system using (8.15). As will be shown in Section 8.4, the bound in (6.4) is tight after a number of iterations at sufficiently high transmit powers for both LR and DH protocols. Hence, this bound is useful to predict the BER performance of the coded relay systems without resorting to lengthy Monte Carlo simulations.

### 8.3 Diversity and Coding Gain Analysis

In the previous section, closed-form expressions of the uncoded PEP were derived for the LR and DH protocols, which were also useful for the development of tight bounds to the error performance of coded systems. However, the derived expressions do not provide insights about the diversity performance of the considered systems. In this section, focusing on
the high transmission power region, we develop simplified PEP expressions that are useful to carry out the diversity and coding gain analysis under uncoded and coded frameworks. For simplicity, hereafter, it is assumed that $P_s = P_r$. We shall start with the LR protocol, followed by the DH one.

### 8.3.1 LR Analysis

For the LR protocol, we first consider the diversity analysis for the uncoded system before extending the results to the coded one. Insights into the precoder design shall also be discussed. Our analysis is applicable to any $L \geq 3$ and can thus be considered as a generalization of the HD NAF studies in [109, 138] for uncoded and in [111] for coded scenarios.

#### Uncoded System

For the uncoded LR system, the PEP in (8.8) can first be upper bounded using the Chernoff bound $Q(\sqrt{2x}) < [1/2] \exp(-x)$ as

$$P(s \to \tilde{s}) = \frac{1}{\pi} \int_0^{\pi/2} \Delta_\theta(s, \tilde{s}) d\theta < \frac{1}{2} \Delta_\theta(s, \tilde{s})\bigg|_{\theta = \pi/2}, \quad (8.17)$$

where $\Delta_\theta(\cdot, \cdot)$ is given in (8.12). Let $\epsilon = s - \tilde{s} = [\epsilon_1, \ldots, \epsilon_L] \neq 0$ and $g_i^T$ be the $i$-th row of $G$. As $\Delta_\theta(\cdot, \cdot)$ is still very involved even when $\theta = \pi/2$, we have the following proposition with regards to this function in high transmit power regions.

**Proposition 8.1.** When $P_s$ is sufficiently large, $\Delta_{\pi/2}(\cdot, \cdot)$ in (8.12) can be simplified as

$$\Delta_{\pi/2}(s, \tilde{s}) = \begin{cases} \frac{16N^2_\eta}{\phi_0 \phi_2 (L-1) \lambda^2} \cdot P_s^{-2} \ln(P_s) + O(P_s^{-2}), & \lambda = 0 \\ \frac{16N^2_\eta \lambda}{\phi_0 \phi_2 (L-1)^2 |\epsilon|^2} \cdot P_s^{-2} \ln(P_s) + O(P_s^{-2}), & 0 < \lambda < 1 \\ \frac{16N^2_\eta |\phi_1 + (L-2) \beta|}{\phi_0 \phi_2 |u_1|^2} \cdot P_s^{-2} \ln(P_s) + O(P_s^{-2}), & \lambda = 1. \end{cases} \quad (8.18)$$

In (8.18), $U = (\sum_{i=1}^L |u_i|^2)(\sum_{i=1}^{L-1} |u_i|^2) - |\sum_{i=1}^{L-1} u_i u_i^\dagger|^2$, $u_i = g_i^T \epsilon$, and $\eta = \sum_{i=1}^{L-1} q_i = \sum_{i=1}^{L-1} \sum_{k=1}^L |g_{i,k}|^2$.

**Proof.** See Appendix E.1. \qed
Note that when $L = 2$, FD LR reduces to HD NAF and it is not hard to verify that
\[
\Delta_{\pi/2}(s, \tilde{s}) = \frac{16N_d^2 \eta}{\phi_0 \phi_2 |u_1|^4} \cdot P_s^{-2} \ln(P_s) + O(P_s^{-2}).
\]  
(8.19)

Substituting (8.18) into (8.17) and ignoring the lower order terms, the PEP for the uncoded LR system can be approximated for high transmitted powers as
\[
\Pr(s \rightarrow \tilde{s}) \approx \frac{8N_d^2 \cdot F_{\text{unc}}^{\text{LR}}(G, \epsilon)}{\phi_0 \phi_2 (L - 1)} \cdot P_s^{-2} \ln(P_s),
\]  
(8.20)

where
\[
F_{\text{unc}}^{\text{LR}}(G, \epsilon) = \begin{cases} 
\eta/\mathcal{U}, & \lambda = 0 \\
n_\eta \left(\frac{L}{(L-1)|u_1|^T}\right), & 0 < \lambda < 1 \\
n_\phi + L - 1 \beta \phi_1 (L-1)|u_1|^T, & \lambda = 1.
\end{cases}
\]  
(8.21)

First, observe from (8.20) that a full-diversity gain function of $P_s^{-2} \ln(P_s)$ can be achieved for any $0 \leq \lambda \leq 1$ as long as the coding gain function $F_{\text{unc}}^{\text{LR}}(G, \epsilon)^{-1} \neq 0$. This diversity function is the same as the one obtained for the uncoded HD NAF system studied in [109, 138]. Thus, any precoder with $|u_1| = \|g_1^\top \epsilon\| \neq 0$ for $0 < \lambda \leq 1$ or $\mathcal{U} \neq 0$ for $\lambda = 0$ offers a full-diversity performance for the uncoded LR system. This class of precoders will be referred hereafter as full-diversity precoders. More importantly, thanks to the use of the direct $S-D$ link, the FD LR system does not suffer from an error floor as long as a full-diversity precoder is used.

Among the class of full-diversity $G$, it is then of interest to find the optimal precoder that maximizes the coding gain using the worst-case PEP analysis. Let $\eta_i = \|g_i\|^2$. From (8.20), this is equivalent to solving the following optimization problem:
\[
\min_{G} \max_{\epsilon \neq 0} F_{\text{unc}}^{\text{LR}}(G, \epsilon) \quad \text{s.t.} \quad \sum_{i=1}^L \eta_i = L.
\]  
(8.22)

Consider first the case of $0 < \lambda \leq 1$. As an initial step, we find the optimal value for the magnitude of each row of $G$, $\eta_i^*$. This problem can be separated from (8.22) as follows. Let $g_i = \sqrt{\eta_i} \tilde{g}_i = \sqrt{\eta_i} \cdot [\tilde{g}_{i,1}, \ldots, \tilde{g}_{i,L}]^\top$, where $\|\tilde{g}_i\|^2 = 1$. Then, $\eta/|u_1|^4$ in (8.21) for $0 < \lambda < 1$
can be written as
\[ \eta \left| u_1 \right|^4 = \frac{\sum_{l=1}^{L-1} \eta_l}{\eta_1^2 \left| g_1^\top \epsilon \right|^4} = \frac{1}{\eta_1} \left( \sum_{l=2}^{L-1} \eta_l \right) \left( \frac{1}{\eta_1} + \sum_{l=2}^{L-1} \eta_l \right). \]

The optimal values of \( \eta_l \) for \( 0 < \lambda < 1 \) can then be found by minimizing \( \frac{1}{\eta_1} + \sum_{l=2}^{L-1} \frac{\eta_l}{\eta_1^2} \). It is easy to see that this term is decreasing with \( \eta_1 \) and increasing with \( \eta_l \) for \( 2 \leq l \leq L-1 \). Since \( \sum_{l=1}^{L} \eta_l = L \), at optimality, \( \eta_1^* = L \) and \( \eta_l^* = 0 \) for \( 2 \leq l \leq L \). For \( \lambda = 1 \), \( \frac{1}{\eta_1^2 \left| g_1^\top \epsilon \right|^4} \) is also decreasing with \( \eta_1 \). Therefore, \( \eta_1^* = L \) and \( \eta_l^* = 0 \) for \( 2 \leq l \leq L \) are optimal when \( 0 < \lambda \leq 1 \). The problem in (8.22) simplifies to
\[
\max \min_{g_1} \frac{\left| u_1 \right|^2}{\| g_1 \|^2} \quad \text{s.t.} \quad \| g_1 \|^2 = L. \tag{8.23}
\]

From the above derivation, the optimal strategy for \( 0 < \lambda \leq 1 \) is to transmit from the source a superposition of all symbols in the first time slot \( (i = 1) \) using all the power and to remain silent in the following \( L-1 \) time slots. The relay in turn must amplify and forward in the second slot \( i = 2 \) the information it received in the previous time. Hence, at sufficiently large \( P_s \), the relay is in fact trying to transmit in HD mode to maximize the coding gain and the full-duplexity of the system is not exploited. This is due to the fact that among all transmitted symbols \( \{x_1, \ldots, x_L\} \) in Table 2.2, \( x_1 \) is the only one that: 1) passes through both the direct \( S-D \) and the relay \( S-R-D \) link (unlike \( x_L \) that only goes through the direct), and 2) is not affected by the FD self-interference (unlike \( \{x_2, \ldots, x_{L-1}\} \)). Consequently, given that the source and relay transmit in an orthogonal manner, HD orthogonal relaying using a superposition constellation can be considered asymptotically optimal. A similar behavior was observed in [138] for the uncoded HD NAF system.

Note that the above observations hold as long as \( V = O(P_s^\lambda) \) with \( \lambda > 0 \). When \( \lambda = 0 \), \( V = O(1) \) and the asymptotic performance depends not only on \( u_1 \) and \( \eta \), but also on \( u_i \) for \( 2 \leq i \leq L \) through the parameter \( U \). Therefore, an explicit construction of the optimal precoder appears to be much more challenging and is left for future work.

**Coded System**

Similar to the uncoded system, it is straightforward to show that the BER of the coded LR scheme can be approximated by considering the most significant term in (6.4) and applying
the Chernoff bound to (6.11) as

\[ P_b \approx \frac{c_{d_H}}{2k_c} [\gamma_0(\Psi, \xi, G)]^{d_H} \bigg|_{\theta=\pi/2}. \quad (8.24) \]

Then, substituting the simplified expression of (8.18) in (8.16) and ignoring the lower order terms, \( \gamma_{\pi/2}(\Psi, \xi, G) \) can be approximated as

\[ \gamma_{\pi/2}(\Psi, \xi, G) \approx \frac{1}{Lm_c2^{Lm_c}} \sum_{s \in \Psi} \sum_{k=1}^{Lm_c} \Delta_{\pi/2}(s, p) = \frac{16N_d^2P_s^{-2}\ln(P_s)}{Lm_c2^{Lm_c}\phi_0\phi_2(L-1)} \sum_{s \in \Psi} \sum_{k=1}^{Lm_c} F_{\text{LR}}^{\text{code}}(G, \epsilon). \]

When the labeling \( \xi \) is implemented independently and identically for each component \( s_i \in \Omega, \epsilon = s - p \) has only one non-zero component. The above function can then be simplified as in Chapter 6 by averaging over the 1-bit neighbors \( s, p \in \Omega \) with \( \epsilon = s - p \neq 0 \) (rather than \( s, p \in \Psi \)) as

\[ \gamma_{\pi/2}(\Psi, \xi, G) = \frac{16N_d^2P_s^{-2}\ln(P_s)}{Lm_c2^{Lm_c}\phi_0\phi_2(L-1)} \cdot F_{\text{LR}}^{\text{code}}(G) \cdot \sum_{s \in \Omega} \sum_{k=1}^{m_c} \frac{1}{|\epsilon|^1} \]

with

\[ F_{\text{code}}^{\text{LR}}(G) = \begin{cases} \lambda -1 \sum_{l=1}^{L} \left[ \left( \sum_{i=1}^{L} |g_{i,l}|^2 \right) \left( \sum_{i=1}^{L-1} |g_{i,l}|^2 \right) - \left| \sum_{i=1}^{L-1} g_{i,l}g_{i+1,l}^+ \right|^2 \right], & \lambda = 0 \\ \frac{\lambda}{L-1} \left( \sum_{l=1}^{L-1} \frac{m_l}{\eta_l^2} \right) \left( \sum_{l=1}^{L} \frac{1}{|g_{i,l}|^4} \right), & 0 < \lambda < 1 \\ \frac{1}{L-1} \left( \frac{(L-2)^2}{\phi\eta_1^3} + \sum_{l=1}^{L-1} \frac{m_l}{\eta_l^4} \right) \left( \sum_{l=1}^{L} \frac{1}{|g_{i,l}|^4} \right), & \lambda = 1. \end{cases} \quad (8.25) \]

Then, the BER expression in (8.24) simplifies to

\[ P_b \approx \frac{c_{d_H}}{2k_c} \left[ \frac{16N_d^2 \cdot F_{\text{code}}^{\text{LR}}(G)}{Lm_c2^{Lm_c}\phi_0\phi_2(L-1)} \cdot \sum_{s \in \Omega} \sum_{k=1}^{m_c} \frac{1}{|\epsilon|^4} \right]^{d_H} \times [P_s^{-2}\ln(P_s)]^{d_H}. \quad (8.26) \]

Observe from (8.26) that a full-diversity gain function of \( [P_s^{-2}\ln(P_s)]^{d_H} \) can be achieved for any \( 0 \leq \lambda \leq 1 \) as long as the coding gain function \( F_{\text{code}}^{\text{LR}}(G)^{-1} \neq 0 \).
case, this is the same diversity gain function as the one observed for the coded HD NAF system in [111], and any precoder with $\mathcal{F}^\text{LR}_{\text{code}}(\mathbf{G})^{-1} \neq 0$ belongs to the class of full-diversity precoders for the coded LR system. Similarly, the coded system does not present an error floor as long as a full-diversity precoder is used.

The optimal precoder $\mathbf{G}^*$ for the coded system can then be obtained by minimizing (8.26), or equivalently (8.25), given the constraint $\sum_{i=1}^{L} \eta_i = L$. First, note that (8.25) depends only on the magnitude of each component of $\mathbf{G}$, $|g_{i,k}|^2$, and not on its actual value. As before, by focusing on $0 < \lambda \leq 1$, it can be easily shown that $\mathcal{F}^\text{LR}_{\text{code}}(\mathbf{G})$ is a decreasing function of $\eta_1$ and an increasing function of $\eta_l$ for $2 \leq l \leq L - 1$. Thus, $\eta_1^* = L$ and $\eta_l^* = 0$ for $2 \leq l \leq L$. Furthermore, by setting $\eta_i$ to the above values, it can be shown from the arithmetic/geometric mean inequality that $|g_{1,k}^*|^2 = 1$ for $1 \leq k \leq L$ minimizes $\mathcal{F}^\text{LR}_{\text{code}}(\mathbf{G})$ when $\lambda > 0$. Therefore, as in the uncoded scenario, transmitting a superposition constellation using HD orthogonal relaying is asymptotically optimal for $0 < \lambda \leq 1$. This behavior has also been observed for the coded HD NAF system in [111].

### 8.3.2 DH Analysis

We now turn our attention to the analysis of the uncoded and coded DH systems.

#### Uncoded System

First, the PEP of the uncoded DH system can be upper bounded similar to (8.17) by replacing $\Delta_\phi(\mathbf{s}, \tilde{\mathbf{s}})$ with $\Delta_\phi(s_i, \tilde{s}_i)$ in (8.15). As in the LR analysis, we then have the following proposition with regards to the asymptotic behavior of $\Delta_\phi(\cdot, \cdot)$ in (8.15).

**Proposition 8.2.** When $P_s$ is sufficiently large, $\Delta_{\pi/2}(\cdot, \cdot)$ in (8.15) can be simplified depending on the value of $\lambda$ as

$$\Delta_{\pi/2}(s_i, \tilde{s}_i) = \begin{cases} 
\frac{4N_d}{\phi_2|c_1|^2} \cdot P_s^{-1} \ln(P_s) + O(P_s^{-1}), & \lambda = 0 \\
\frac{4^\beta}{\phi_1|c_1|^2} \cdot P_s^{-(1-\lambda)} + O(P_s^{-1} \ln(P_s)), & 0 < \lambda < 1 \\
\frac{4^\beta}{4^\beta + \phi_1|c_1|^2} + O(P_s^{-1} \ln(P_s)), & \lambda = 1.
\end{cases} \quad (8.27)
$$

**Proof.** See Appendix E.2.

Based on the above proposition and (8.17), the PEP for the uncoded DH system can...
be approximated as
\[ P(s_i \rightarrow \bar{s}_i) \approx 2 \cdot F_{\text{unc}}^{\text{DH}}(\epsilon_i) \cdot D_{\text{unc}}^{\text{DH}}(P_s), \quad (8.28) \]

where
\[ F_{\text{unc}}^{\text{DH}}(\epsilon_i) = \begin{cases} N_d / [\phi_2|\epsilon_i|^2], & \lambda = 0 \\ \beta / [\phi_1|\epsilon_i|^2], & 0 < \lambda < 1 \\ \beta / [4\beta + \phi_1|\epsilon_i|^2], & \lambda = 1, \end{cases} \quad (8.29) \]

and the diversity gain function
\[ D_{\text{unc}}^{\text{DH}}(P_s) = \begin{cases} P_s^{-1}\ln(P_s), & \lambda = 0 \\ P_s^{-(1-\lambda)}, & 0 < \lambda < 1 \\ 1, & \lambda = 1. \end{cases} \quad (8.30) \]

First, observe from (8.28) and (8.30) that different from the LR system, the diversity gain function of the DH scheme depends on the value of \( \lambda \). In particular, a diversity gain function of \( P_s^{-(1-\lambda)} \) is attained when \( 0 < \lambda < 1 \), whereas the gain function is \( P_s^{-1}\ln(P_s) \) when \( \lambda = 0 \). Similar to previous works, an error floor is observed when \( \lambda = 1 \). Following the same approach as in Appendix E.2 for \( \lambda = 0 \), it can be easily shown that the diversity function of a HD DH system is \( P_s^{-1}\ln(P_s) \). Thus, in contrast to the HD scheme, the FD DH system can achieve a full rate of one transmitted symbol per period but must sacrifice its diversity to \( P_s^{-(1-\lambda)} \) when \( \lambda > 0 \).

**Coded System**

For the coded DH system, the BER can be approximated following the approach in (8.24) as
\[ P_b \approx \frac{c_{\text{di}}}{2k_c} \left[ \frac{4}{m_c2^{m_c}} \sum_{s \in \Omega} \sum_{k=1}^{m_c} F_{\text{unc}}^{\text{DH}}(\epsilon) \right]^{d_{\text{di}}} \times [D_{\text{unc}}^{\text{DH}}(P_s)]^{d_{\text{di}}}. \quad (8.31) \]

Similar to the uncoded case, observe from (8.31) that the diversity gain function of the system depends on the value of \( \lambda \). The diversity function is \( P_s^{-(1-\lambda)d_{\text{di}}} \) for \( 0 < \lambda < 1 \) and \( [P_s^{-1}\ln(P_s)]^{d_{\text{di}}} \) for \( \lambda = 0 \). A zero-diversity order is still attained when \( \lambda = 1 \).
8.4 Illustrative Examples

In this section, we provide simulation results to verify the theoretical analysis carried out in previous sections. In all simulations, the Gray-labeled QPSK constellation is employed unless otherwise stated. Similar results can be obtained for other modulation schemes and mapping rules. The channel gains are assumed to be Rayleigh-distributed with unit variance. The powers at $S$ and $R$ are also assumed to be equal $P_s = P_r$ and unit noise power at both nodes is considered $N_r = N_d = N_0 = 1$. Furthermore, to concentrate on the effect on $\lambda$, we set $\beta = 1$ in (8.1). For the coded systems, the rate-1/2 convolutional code with generator matrix $\{5; 7\}$ is considered along with an 80000-length bit interleaver.

8.4.1 LR Examples

Fig. 8.2 shows the BER performance of the uncoded LR system for $L = 3$ against bit SNR $E_b/N_0$ (in dB). The $3 \times 3$ precoders considered in Fig. 8.2 are the identity $I_3$ precoder (i.e., no precoder) and the $G_{cyclo3}$ precoder from [121]. Furthermore, three values of $\lambda$ in (8.1) are considered: $\lambda = 0, 0.5, \text{and } 1$. First, it can be checked from (8.21) that $G_{cyclo3}$ is a full-diversity precoder for the uncoded system, i.e., $F_{\text{LR unc}}(G_{cyclo3}, \epsilon)^{-1} \neq 0$. As a consequence, the slopes of the BER curves with $G_{cyclo3}$ remain unchanged for any $\lambda$ and match with the diversity order (Div.) $P_s^{-2} \log(P_s)$, which is also plotted in Fig. 8.2 as a reference. On the other hand, $I_3$ is not a full-diversity precoder and thus it can be seen from Fig. 8.2 that its diversity performance depends on the value of $\lambda$. Specifically, an error floor is observed when $\lambda = 1$.

The diversity behavior in the coded scenario is illustrated in Fig. 8.3 for $L = 3$ and $\lambda = \{0, 0.5, 1\}$. In particular, Fig. 8.3 shows the BER performances of the LR BICM-ID system after 3 decoding iterations along with the BER bound in (6.4)$^2$. First, observe from Fig. 8.3 that the BER performance of all systems converges to the bound in (6.4) at a sufficiently high transmitted power. Specifically, the analytical and simulation results converge around the practical BER levels of $10^{-3}$-$10^{-4}$. This confirms the accuracy of the PEP expressions derived in (8.8) and (8.12). Moreover, it can be checked from (8.25) that $G_{cyclo3}$ is a full-diversity precoder for the coded LR system, whereas $I_3$ is not since $F_{\text{code}}(I_3)^{-1} = 0$. Therefore, the BER curves with $G_{cyclo3}$ present the same slope, while the

$^2$The first 20 terms of the summation in (6.4) are retained in this and all subsequent figures.
slopes of the identity systems change with $\lambda$. An error floor can still be observed in Fig. 8.3 when $\lambda = 1$. This is in agreement with the analysis in Section 8.3.1.

To demonstrate the asymptotic optimality of HD relaying when $\lambda > 0$, Fig. 8.4 shows the BER performance of the uncoded LR system for $L = 3$ and $\lambda = \{0.5, 0.7, 1\}$. Besides $G_{cyclo3}$, we consider the optimal $3 \times 3$ precoder for the uncoded system $G_{ex3}$ where $g_{1,ex3}^\top = [0.755, 0.387 e^{j0.236}, 1.510]$ and $g_{i,ex3}^\top = 0$ for $i = 2, 3$. The vector $g_{1,ex3}$ is the solution to (8.23) for the QPSK constellation and was obtained using exhaustive search. Note that since the optimal precoder transforms the FD system into a HD one, its performance is independent of $\lambda$. Observe from Fig. 8.4 that at the BER level of $10^{-5}$, the $G_{ex3}$ precoder outperforms $G_{cyclo3}$ when $\lambda \geq 0.7$. However, at this BER level, it is yet not able to outperform $G_{cyclo3}$ with $\lambda = 0.5$ and the crossover happens only at the BER of $10^{-6}$. Therefore, although HD transmission is asymptotically optimal, FD might be better at practical BER levels when $\lambda$ is sufficiently small, i.e., $\lambda \leq 0.5$.

Fig. 8.5 shows the BER performance after 10 iterations of the coded LR system for $L = 3$ and $\lambda = \{0, 0.5, 1\}$. The considered $3 \times 3$ precoders are $G_{cyclo3}$, and $G_{damen3}$ with $g_{1,damen3}^\top = [1, e^{j\varphi}, e^{j2\varphi}]$ and $g_{i,damen3}^\top = 0$ for $i = 2, 3$. The $G_{damen3}$ precoder is asymptotically optimal.
Fig. 8.3 BER performances after 3 iterations of the LR BICM-ID system using different $3 \times 3$ precoders ($\lambda = 0, 0.5, 1$).

Fig. 8.4 BER performances of the uncoded LR system for $L = 3$ and $\lambda = 0.5, 0.7, 1$. 
for the coded system as it minimizes (8.25) when $\lambda > 0$. Furthermore, the angle is set to $\varphi = 0.312$, which is the solution to (8.23) for the structure $g_1^T = [1, e^{j\varphi}, e^{j2\varphi}]$ [118]. Such angle guarantees a good first iteration performance and a faster convergence to the bound [111]. It can be seen from Fig. 8.5 that the asymptotic performance of $G_{\text{damen3}}$ outperforms those of $G_{\text{cyclo3}}$ for any value of $\lambda$. In particular, the LR system with $G_{\text{damen3}}$ presents a gain of more than 2 dB over $G_{\text{cyclo3}}$ at the BER level of $10^{-5}$. Thus, different from the uncoded case, the HD system takes advantage of the iterative gain in the BICM-ID structure and it is preferred over FD at this practical BER level.

![Fig. 8.5 BER performances after 10 iterations of the LR BICM-ID system for $L = 3$ and $\lambda = 0, 0.5, 1$.](image)

### 8.4.2 DH Examples

The BER performances of the uncoded DH system are shown in Fig. 8.6 for different values of $\lambda$. The diversity orders in (8.30) are also illustrated in Fig. 8.6 as a reference. First, note that the slopes of the BER curves are in agreement with the diversity analysis in Section 8.3.2. As expected, it can then be seen from Fig. 8.6 that the diversity behavior is highly dependent on the value of $\lambda$. In particular, the diversity order is a decreasing
function of $\lambda > 0$ and an error floor is observed in Fig. 8.6 when $\lambda = 1$.

![Graph showing BER performances of the uncoded DH system for different values of $\lambda$.](image)

**Fig. 8.6** BER performances of the uncoded DH system for different values of $\lambda$.

Similar trends for the coded scenario can be seen in Fig. 8.7. For the DH BICM system, it was observed that iterations do not significantly improve the performance. As such, all curves in Fig. 8.7 are plotted after 1 iteration. First, note that the BER bounds are tight for all considered systems. This confirms the correctness of the expressions derived in (8.13) and (8.15). As in the uncoded scenario, it can be seen from Fig. 8.7 that $\lambda$ has a great impact on the diversity order of the coded DH system. Specifically, a zero diversity order is still attained when $\lambda = 1$.

Finally, to compare the FD and HD schemes, the BER of the uncoded/coded HD DH systems is also plotted in Figs. 8.6 and 8.7. To maintain the same throughput, we assume that the HD systems employ the Gray-labeled 16-QAM modulation scheme. Note from these figures that at the BER level of $10^{-5}$, the HD systems outperform all FD schemes with $\lambda \geq 0.3$. However, the FD systems are advantageous when $\lambda \leq 0.2$. As such, for practical purposes, $\lambda < 0.3$ for FD to be worth using in this scenario.
Fig. 8.7  BER performances of the DH BICM system for different values of $\lambda$.

8.5 Chapter Summary

This chapter investigated the error performance and diversity behavior of the FD LR and DH systems under the effect of residual self-interference. Closed-form expressions of the PEP were first derived for the uncoded systems. These expressions were then used to provide tight bounds to the BER of the BICM-ID systems. Simplified PEP and BER expressions were also presented assuming high transmission powers. It was then shown that the FD LR systems achieve the same diversity function as their HD counterparts as long as a suitable precoder is applied. In the case of FD DH, it was demonstrated that the diversity order is a decreasing function of $\lambda$ and is equal to zero only when $\lambda = 1$. Although the HD LR and DH systems were shown to be asymptotically optimal, simulations results revealed that their FD counterparts might be advantageous at practical BER levels when a high cancellation quality is used, i.e., when $\lambda$ is small.
Chapter 9

Conclusion

This thesis was devoted to the design and analysis of reliable and spectrally-efficient HD and FD AF relay systems. In particular, we investigated achievable rates, power allocation schemes, and code designs for several AF relay protocols. The main contributions of this thesis can be summarized as follows.

In Chapter 3, we derived the capacity and respective optimal transmission scheme for a HD NAF system in a static environment. Specifically, by deriving and comparing all local solutions of the mutual information, the optimal covariance matrix at the source and power allocation at relay were obtained. First, for the individual power constraint scenario, we showed that the capacity is achieved by either a DT, a NAF-FR, or a NAF-BF scheme. We then showed that only a DT or a NAF-BF protocol can achieve the capacity in the global constraint scenario. As discussed in Chapter 3, the OAF protocol is thus suboptimal under both power constraints. The capacity and optimal transmission schemes were also analyzed for some concrete examples such as in high and low transmission powers, and over the symmetric and linear network models.

In Chapter 4, we investigated achievable rates and power allocation schemes for several OW and TW HD-AF protocols over Rayleigh fading channels. First, a general approach was proposed to obtain tight and simple approximations to the achievable rates in high and low transmission power regions. As observed in Chapter 4, the proposed approach is applicable to different AF relaying schemes, and to both CI and FG coefficients. The derived approximations were then used not only to compare the CI and FG techniques, but also to obtain asymptotically optimal power allocation schemes to maximize the rate or
sum rate for OW and TW relaying, respectively. For the case of the DH scheme using the CI coefficient, a closed-form expression of the rate was derived and bisection was proposed to obtain the optimal power allocation for any power region. Comparisons among different protocols were also presented.

In Chapter 5, we proposed power adaptation policies at the relay for cooperative OW and TW HD-AF systems. Specifically, by assuming Gaussian codebooks at the source nodes and full CSI at the relay, optimal power adaptation schemes were derived to maximize the achievable rate or sum rate under a long-term average power constraint. As noted in Chapter 5, the derived solutions can be interpreted as water-filling like schemes over time and space. Insights on the optimal amplification coefficients with respect to the channel gains were then provided. It was finally observed that the proposed schemes provide significant rate benefits over the systems using the conventional FG and CI techniques.

In Chapter 6, we introduced the idea of precoding over multiple cooperative frames to improve the diversity and approach the capacity of a HD NAF system over Rayleigh fading channels. Considering a BICM-ID framework, a tight union bound on the BER was first derived. The NAF-BICM-ID system was then analyzed in both error-floor and turbo pinch-off regions. In particular, in the error-floor area, an optimal class of precoders in terms of diversity and coding gain was developed along with a pragmatic approach to obtain good rotation angles for such class. In the turbo pinch-off region, it was demonstrated that a combination of multi-frame precoding and multi-D mapping can be used to approach the capacity of the relay channel.

In Chapter 7, we turned our attention to the capacity and optimal power allocation scheme for a static FD DH system under residual self-interference and without direct link. By modeling the variance of the residual self-interference as proportional to the $\lambda$-th power of the transmitted power, the optimization problems were first shown to be quasiconcave under per-node and sum power constraints. Bisection was then proposed to find the single root of the non-linear derivative. The capacity and optimal schemes were also analyzed in different high power regions. Specifically, we showed that full power at the relay is not necessarily optimal and that the optimal multiplexing gain is $1/[1 + \lambda]$. The conditions under which FD outperforms HD were finally discussed.

In Chapter 8, we analyzed the error and diversity performance of the FD DH and LR protocols over Rayleigh fading channels. Applying the same interference model as in Chapter 7, we first derived closed-form expressions of the uncoded PEP for the considered
systems. Such expressions allowed us to then derive tight BER bounds for the coded BICM systems. A diversity analysis for both uncoded and coded scenarios was also carried. In particular, we showed that the FD LR systems with a suitable precoder can attain the same diversity function as their HD counterparts. We then demonstrated that the diversity order of the FD DH systems is a decreasing function of $\lambda$ and is equal to zero when $\lambda = 1$. Although HD relaying was proved to be asymptotically optimal, simulations results revealed that FD is advantageous at practical BER levels when $\lambda$ is sufficiently small.

9.1 Future Work

The proposed future work is explained in detail below.

9.1.1 Extension Problems

We first present some problems which are extensions of the ones encountered in this thesis.

- In Chapter 4, asymptotically optimal power allocation schemes for several AF protocols were derived under the assumption of a sum power constraint. Similar to Chapter 3, it is also important to carry out this analysis for the individual constraint scenario. In addition, although the MRC approach proposed in Chapter 4 could potentially be applied to other AF schemes, a different method must be devised for protocols in which the asymptotic rate expressions still present significant cross terms of different channel gains, e.g., the three-phase TW scheme or the OAF protocol in (4.19). This presents another interesting research direction.

- In Chapter 5, power adaptation policies were derived assuming full CSI at the relay. However, the average power allocation schemes at the source nodes and relay were left as parameters (i.e., vector $q$). The achievable rates of the considered systems could be further improved by optimizing these average values similar to Chapter 4. Thus, investigating optimal average power allocations is one promising research direction. Furthermore, it can be seen from the simulation results in Chapter 5 that, under some scenarios, there is a significant performance gap between the proposed relay adaptation methods and the conventional CI and FG techniques. This provides an incentive to study relay adaptation schemes based on partial CSI at the relay.
• In Chapter 7, the capacity of the static FD DH system with no direct link was investigated. Given the importance of the direct link, this study should be extended to the FD LR protocol. Such study would be more challenging than the one for HD NAF relaying presented in Chapter 3 as \( L \) variables at the source and \( L - 1 \) at the relay have to be optimized. The capacity of both LR and DH protocols over fading channels has also not been addressed in the literature. As such, carrying an analysis similar to Chapter 4 for FD relaying is important as well.

• In Chapters 6 and 8, the error performance of different AF systems was investigated assuming a fixed average power allocation at the source and relay. As observed in the other chapters, further optimizing the average power allocation can yield significant gains. Hence, extending the analyzes in these chapters to a general power allocation scheme and finding the optimal allocation should be addressed in future works. This is particularly important for the FD DH system as full relay power has been shown to be not necessarily optimal (see Chapter 7). We conjecture that a better diversity order can then be achieved by optimizing the power of the FD DH scheme.

• In Chapter 8, the precoder design of the LR system with \( \lambda = 0 \) was not addressed. Given that experimental results have shown that \( \lambda \) can be quite close to zero [35], such design needs to be considered in the future.

• In the simulation results in Chapters 7 and 8, the \( \beta \) parameter of the considered interference model was set to one to focus on the effect of \( \lambda \). However, as noted in Section 7.1, \( \beta \) is in fact a function of \( \lambda \), i.e., \( \beta = \beta(\lambda) \). Thus, further research is required to analyze the relation between \( \beta \) and \( \lambda \), and study their joint effect on the system performance.

• Finally, applying the idea of multi-frame and capacity-approaching designs to the FD systems similar to Chapter 6 is another interesting future work.

9.1.2 More Advanced Problems

We now present more advanced research problems which we believe are worthwhile to investigate.
9 Conclusion

- Chapters 3 and 5 studied power allocation schemes among the nodes for a static system, and relay adaptation policies for a faded one, respectively. These two problems are subsets of a general formulation in which both the source and relay adapt their power over time to optimize the achievable rate under an average power constraint. The latter problem is significantly more challenging than the ones addressed in these chapters as the power adaptation at the relay will become a function of that at the source. This results in a non-concave optimization problem with multiple variables, which is not analytically tractable. Thus, proposing suboptimal source/relay adaptation solutions and finding a rate upper bound to assess these solutions is an interesting research direction.

- In this thesis, the problem of relay location was not considered as we mainly focused on finding optimal power allocation schemes. Therefore, jointly optimizing the relay location and power allocation is another promising research area.

- As this thesis focused on single-relay single-antenna AF systems, a natural extension is to look into multi-antenna and multi-relay scenarios. This will certainly provide more degrees of freedom and a significant performance improvement. Also, given the importance of DF in the literature, another interesting topic is to extend the analysis in this thesis to DF relaying.

- In Chapters 7 and 8, we adopted the empirical residual self-interference model from [30] in which the interference variance is proportional to the \( \lambda \)-th power of the transmitted power. Moreover, the self-interference was assumed to be Gaussian distributed. However, we should note that the relation between transmitted signal and residual self-interference has just started to be investigated, and the relation adopted in this thesis is one of the pioneering models for full-duplex radios. Therefore, as more studies on modeling the residual self-interference for different FD prototypes are undertaken, the FD results in this thesis might need to be revisited.

- Finally, many chapters in this thesis concentrated on the achievable rate or sum rate as the performance criteria. Due to the broadcast nature of the wireless channel, there has been an increasing research interest on how to transmit confidential information securely in the presence of eavesdroppers. In this context, physical layer security based on information theory has been gaining considerable attention, e.g., [139, 140].
9 Conclusion

The rate at which a pair of nodes can communicate in perfect secrecy, i.e., the secrecy rate, then becomes the performance metric of interest. Recently, the use of helper nodes has been proposed to improve the secrecy rate of a source-destination pair (see for example [141–143]). However, many issues regarding AF relaying techniques for secrecy remain unanswered. For instance, optimal power allocation schemes to achieve the secrecy capacity of HD AF wire-tap channels are scarce in the literature. The secrecy capacity of non-orthogonal or FD relaying systems has also not been investigated. Consequently, AF relaying techniques for secrecy is a fruitful area for future research.
Appendix A

Proofs for Chapter 3

A.1 Proof of Lemma 3.1

We analyze $f_C(q, q_{12}^*)$ in $P_1$ and $P_2$:

i) When $q \in P_1$, $f_C(q, q_{12}^*) = f_{P_1}([q_1, q_s - q_1, z_2])$. By taking the derivatives of $f_{P_1}([q_1, q_s - q_1, z_2])$ with respect to $q_1$ and $z_2$, it can be shown that there is a single stationary point inside this region at $q_1 = [q_s \gamma^2_0 - \gamma_1]/\gamma^2_0$ and $z_2 = [(q_s \gamma^2_0 \gamma_1 + \gamma^2_0 - \gamma^2_1)(q_s \gamma^2_0 - 2\gamma_1)]/[\gamma^2_0 \gamma_2 (\gamma_0 + \gamma_1)]$, which is feasible when $q_s \gamma^2_0 > 2\gamma_1$ and $z_2 < z_r$. However, the Hessian of $f_{P_1}([q_1, q_s - q_1, z_2])$ at the above point is given as $\nabla^2 f_{P_1}([q_1, q_s - q_1, z_2]) = \begin{pmatrix} A_1 & B_1 \\ B_1 & 0 \end{pmatrix}$ with $P(q_1) - z_2 > 0$, $A_1 < 0$ and $B_1 > 0$ when this point is feasible. The Hessian is then indefinite and this point cannot be a local maximizer. Note from the Hessian that $f_C(q, q_{12}^*)$ is not quasiconcave in $R_1$ as there is a stationary point that is saddle rather than a local maximizer [73].

ii) When $q \in P_2$, $f_C(q, q_{12}^*) = f_{P_2}([q_1, q_s - q_1, z_2])$. It can then be shown that $\partial f_{P_2}([q_1, q_s - q_1, z_2])/\partial z_2 = 0$ when $z_2 = [q_1 \gamma_1(q_1 \gamma_1 + 1)]/[\gamma_0 \gamma_2 (q_s - q_1)]$, which is feasible as long as $z_2 < z_r$. However, at this point, $\partial f_{P_2}([q_1, q_s - q_1, z_2])/\partial q_1 > 0$. Thus, no stationary points exist in $P_2$.

iii) When $z_2 = P(q_1)$, we are right at the boundary between $P_1$ and $P_2$. It is easy to show that $f_C(q, q_{12}^*)$ is continuously differentiable in $R_1$. Hence, by continuity and from i) and ii), there are no local maximizers when $z_2 = P(q_1)$.

Given that the Hessian at the only stationary point in $P_1$ is indefinite and that there are no stationary points in $P_2$, $f_C(q, q_{12}^*)$ is not quasiconcave in $R_1$ and has no local solutions.
in the interior of $\mathcal{R}_1$.

### A.2 Solutions to Subcases in Lemma 3.3

Consider the following subcases according to $P(q_1)$.

i) When $P(q_1) \leq z_r$ for the entire range $0 \leq q_1 \leq q_s$, $\ell_{IV} = \ell^{P_2}_{IV}(z_r)$. The function $f_C([q_1, q_s - q_1, z_r], q_s)$ over this line segment can then be written as

$$f_{P_2}([q_1, q_s - q_1, z_r]) = g_1(q_1) + \sqrt{g_2(q_1)},$$

(A.1)

where

$$g_1(q_1) = 1 + \frac{q_s \gamma_0 + q_1 z_r \gamma_2 + q_1 q_s \gamma_0 \gamma_1 + q_1 z_r \gamma_1 \gamma_2}{q_1 \gamma_1 + z_r \gamma_2 + 1},$$

$$g_2(q_1) = \frac{4q_1 z_r \gamma_1 \gamma_2 (q_s - q_1)(q_1 \gamma_1 + 1)}{(q_1 \gamma_1 + z_r \gamma_2 + 1)^2}.$$

The derivatives of the functions in (A.1) are given as

$$\Delta_1(q_1) = \frac{\partial g_1(q_1)}{\partial q_1} = \frac{z_r \gamma_2 (\gamma_0 + \gamma_1 + z_r \gamma_0 \gamma_2 + q_s \gamma_0 \gamma_1 + z_r \gamma_1 \gamma_2)}{(q_1 \gamma_1 + z_r \gamma_2 + 1)^2},$$

$$\Delta_2(q_1) = \frac{\partial \sqrt{g_2(q_1)}}{\partial q_1} = -\frac{\sqrt{\gamma_0 \gamma_1 \gamma_2 z_r} P_5(q_1)}{\sqrt{q_1 (q_s - q_1)(q_1 \gamma_1 + 1)(q_1 \gamma_1 + z_r \gamma_2 + 1)^2}}.$$

(A.2)

where the cubic function

$$P_5(q_1) = A_5 q_1^3 + B_5 q_1^2 + C_5 q_1 + D_5,$$

(A.3)

with $A_5 = -\gamma_1^2, B_5 = -3\gamma_1(z_r \gamma_2 + 1), C_5 = q_s \gamma_1 - 2z_r \gamma_2 + 2z_r q_s \gamma_1 - 2$, and $D_5 = q_s(z_r \gamma_2 + 1)$. Note that $A_5, B_5 < 0$ and $D_5 > 0$. By using Descartes’ rule of signs [144], it can be seen that $P_5(\cdot)$ always has a real positive root regardless of the sign of $C_5$. Since $P_5(0) = D_5 > 0$ and $P_5(q_s) < 0$, this root lies between 0 and $q_s$. Let $r_5$ be the positive real root of $P_5(\cdot)$. $P_5(\cdot)$ is then positive for $0 \leq q_1 < r_5$, negative for $r_5 < q_1 \leq q_s$, and zero when $q_1 = r_5$. From this and (A.2), the limit of $\Delta_2(q_1)$ as $q_1$ approaches zero from above is $\lim_{q_1 \to 0^+} \Delta_2(q_1) = +\infty$ and as it approaches $q_s$ from below is $\lim_{q_1 \to q_s^-} \Delta_2(q_1) = -\infty$. $\Delta_2(q_1)$ is then also positive for $0 \leq q_1 < r_5$, negative for $r_5 < q_1 \leq q_s$, and zero when $q_1 = r_5$, i.e., $\sqrt{g_2(q_1)}$ is
quasiconcave in $0 \leq q_1 \leq q_s$ with $r_5$ as its maximizer. From (A.2) and by taking the second derivative, $g_1(q_1)$ can be shown to be concave and increasing. Hence, $f_{P_2}([q_1, q_s - q_1, z_r])$ is an increasing function in $0 \leq q_1 \leq r_5$ and the local solution cannot lie in this range. For $r_5 \leq q_1 \leq q_s$, $f_{P_2}([q_1, q_s - q_1, z_r])$ can be shown to be concave as follows. By taking the second derivative of (A.1), it can be shown that $g_2(q_1)$ is concave for $q_1 \geq r_5$. Since the square root function is concave and increasing, $\sqrt{g_2(q_1)}$ is also concave for $q_1 \geq r_5$. Given that $g_1(q_1)$ is concave, $f_{P_2}([q_1, q_s - q_1, z_r])$ in (A.1) is a sum of two concave functions for $q_1 \geq r_5$ and thus as well concave over this range. Hence, $f_{P_2}([q_1, q_s - q_1, z_r])$ is concave for $r_5 \leq q_1 \leq q_s$ and a single maximizer must lie in this region. Since $\Delta_1(r_5) + \Delta_2(r_5) = \Delta_1(r_5) > 0$ and $\Delta_1(q_s) + \Delta_2(q_s) = -\infty$, the maximizer must be a stationary point. Thus, we have to find the point in $r_5 < q_1 < q_s$ such that $\Delta_1(q_1) + \Delta_2(q_1) = 0$. By equating $\Delta_2(q_1) = -\Delta_1(q_1)$ in (A.2) and squaring both sides, it can be shown that the maximizer is the root in $r_5 < q_1 < q_s$ of the following 6th order polynomial

$$P_6(q_1) = \gamma_0 \gamma_1 P_5(q_1)^2 - z_r \gamma_2 A_8^2 q_1 (q_s - q_1)(q_1 \gamma_1 + 1), \quad (A.4)$$

where $A_8 = -\gamma_1 (3z_r^2 \gamma_2^2 + 4z_r \gamma_2 + q_s \gamma_1 + 2z_r q_s \gamma_1 \gamma_2 + 1)$. Note from the analysis that this root always exists and it is unique. Let $r_6$ be the root of (A.4) such that $r_5 < r_6 < q_s$. The maximizer of $\ell_{IV} = \ell_{IV}^2(z_r)$ is then $q_1 = r_6$. It is important to point out that $f_{P_2}([r_6, q_s - r_6, z_r]) > f_{P_2}([q_1, 0, z_r])$ as $\lim_{q_1 \rightarrow q_s} -\Delta_1(q_1) + \Delta_2(q_1) = -\infty$. Hence, the maximizer in $\ell_{II}$ cannot be the solution to (3.7). Note also that although $f_{P_2}([q_1, q_s - q_1, z_r])$ is only concave for $r_5 \leq q_1 \leq q_s$, it is a quasiconcave function in $0 \leq q_1 \leq q_s$ with $r_6$ as its maximizer since $f_{P_2}([q_1, q_s - q_1, z_r])$ is increasing for $0 \leq q_1 \leq r_5$.

**ii)** When $P(q_1) > z_r$ for some range in $0 < q_1 < q_s$, the extremes of $\ell_{IV}$ are in $\ell_{IV}^2(z_r)$, while the mid-section is in $\ell_{IV}^1(z_r)$. Let the positive roots of $P(q_1) - z_r = 0$ be denoted as $r$ and $r'$ with $0 < r' < r < q_s$. Note that for $P(q_1) > z_r$, $r' \neq r$. Then, $\ell_{IV}^1(z_r) = \{q \in \ell_{IV} \mid r' \leq q_1 \leq r\}$ and $\ell_{IV}^2(z_r) = \{q \in \ell_{IV} \mid 0 \leq q_1 \leq r', \ r \leq q_1 \leq q_s\}$. In $\ell_{IV}^1(z_r)$, the derivative of $f_{C}([q_1, q_s - q_1, z_r], q_r') = f_{P_1}([q_1, q_s - q_1, z_r])$ is given as

$$\frac{\partial f_{P_1}([q_1, q_s - q_1, z_r])}{\partial q_1} = \frac{P_3(q_1)}{\gamma_0 (q_1 \gamma_1 + z_r \gamma_2 + 1)^2}; \quad (A.5)$$
where the cubic polynomial

$$P_3(q_1) = A_3 q_1^3 + B_3 q_1^2 + C_3 q_1 + D_3,$$

(A.6)

with $A_3 = -2 \gamma_0^3 \gamma_1^2$, $B_3 = \gamma_0^3 \gamma_1 (q_1 \gamma_1 - 3 z_r \gamma_2 - 4)$, $C_3 = 2 \gamma_0^3 (q_1 \gamma_1 - 1) (z_r \gamma_2 + 1)$, and $D_3 = z_r^2 \gamma_0^2 \gamma_1^2 + q_3 z_r \gamma_0 \gamma_1 \gamma_2 + q_3 z_r \gamma_1 \gamma_2 + z_r \gamma_0 \gamma_1 \gamma_2 - z_r \gamma_1^2 \gamma_2 + q_3 \gamma_0^3$. Since $A_3 < 0$ but $B_3$, $C_3$ and $D_3$ can be positive, negative or zero, we need to consider in general 27 subcases. However, due to the relation among the coefficients, only 15 cases are possible.

- When the signs of $\{B_3, C_3, D_3\}$ are $\{-, \pm, +\}$, $\{-, 0, +\}$, $\{+, +, +\}$, $\{0, +, +\}$, $\{0, +, 0\}$, and $\{\pm, +, 0\}$, using Descartes’ rule of sign [144], $P_3(\cdot)$ has a single strictly positive root. Denote this root as $r_r$. Since for these cases $P_3(0) = D_3 \geq 0$, $P_3(q_1)$ and the derivative in (A.5) are positive for $0 < q_1 < r_r$, negative for $q_1 > r_r$, and zero when $q_1 = r_r$. The function $f_{r_1}([q_1, q_3 - q_1, z_r])$ is then quasiconcave with $r_r$ as its maximizer. The maximizer of $\ell_{IV}$ then depends on the location of $r'$ and $r$ with respect to $r_r$. When $r_3 \leq r' < r$, $f_{r_1}([q_1, q_3 - q_1, z_r])$ is a decreasing function in $\ell_{IV}^{P_3}(z_r)$. Since $f_C(q, q_1^{12})$ is continuously differentiable and $f_{r_2}([q_1, q_3 - q_1, z_r])$ is quasiconcave in $0 \leq q_1 \leq q_3$, $r_6$ must be in $0 < q_1 \leq q_3$, i.e., $r_6$ is in the left segment of $\ell_{IV}^{P_3}(z_r)$ in Fig. 3.2. Hence, $f_C([q_1, q_3 - q_1, z_r], q_1^{12})$ in (3.12) is increasing for $0 < q_1 < r_6$, and decreasing for $r_6 < q_1 < q_3$. The maximizer of $\ell_{IV}$ is then $r_6$ in $\ell_{IV}^{P_3}(z_r)$. Similarly, when $r' < r \leq r_3$, $f_{r_1}([q_1, q_3 - q_1, z_r])$ is an increasing function in $\ell_{IV}^{P_3}(z_r)$ and $r_6$ in the right segment of $\ell_{IV}^{P_3}(z_r)$ ($r \leq q_1 < q_3$) is the maximizer of $\ell_{IV}$. Finally, when $r' < r_3 < r$, $r_3 \in \ell_{IV}^{P_3}(z_r)$ and $f_{r_1}([q_1, q_3 - q_1, z_r])$ is quasiconcave in $\ell_{IV}^{P_3}(z_r)$. Thus, $r_6$ is in $\ell_{IV}^{P_3}(z_r)$ and $r_3 \in \ell_{IV}^{P_3}(z_r)$ is the maximizer of $\ell_{IV}$.

- When the signs are $\{-, -, -,\}$, $\{-, -, 0\}$, $\{-, 0, -\}$ and $\{-, 0, 0\}$, denote the maximizer of $P(q_1)$ in $0 \leq q_1 \leq q_3$ as $r^*$. It can be shown that when $C_3 \leq 0$ and $D_3 \leq 0$, $P(r^*) - z_r < 0$, i.e., $\ell_{IV}(z_r) = 0$, which contradicts the assumption that $P(q_1) > z_r$.

- Finally, when the signs are $\{\pm, +, -\}$ and $\{0, +, -\}$, $P_3(q_1)$ can have either two strictly positive roots or no positive roots. When $P_3(q_1)$ has no positive roots, $f_{r_1}([q_1, q_3 - q_1, z_r])$ is decreasing in $\ell_{IV}^{P_3}(z_r)$ since $P_3(0) = D_3 < 0$, and $r_6$ in the left segment of $\ell_{IV}^{P_3}(z_r)$ ($0 < q_1 \leq r'$) is the maximizer of $\ell_{IV}$. When $P_3(q_1)$ has two strictly positive roots and these roots are equal, $f_{r_1}([q_1, q_3 - q_1, z_r])$ is decreasing in $\ell_{IV}^{P_3}(z_r)$ and $r_6$ is again the maximizer of $\ell_{IV}$. When these roots are not equal, denote the smallest one as $r_3'$ and the greatest one as $r_3$. Since $P_3(0) = D_3 < 0$, (A.5) is negative for $0 \leq q_1 < r_3'$, positive for $r_3' < q_1 < r_3$, and negative again for $q_1 > r_3$. The function $f_{r_1}([q_1, q_3 - q_1, z_r])$ is then decreasing for $0 \leq q_1 < r_3'$,
increasing for \( r' < q_1 < r_3 \), and decreasing again for \( q_1 > r_3 \). The maximizer of \( \ell_{IV} \) then depends on the location of \( r' \) and \( r \) with respect to \( r_3 \) and \( r'_3 \). By analyzing the location of the maximizers of \( P(\cdot) \) and \( P_3(\cdot) \), it can be shown that \( r_3 \) and \( r'_3 \) are in \( 0 \leq q_1 \leq q_s \) and \( r > r_3 \) for these sign cases. The maximizer of \( \ell_{IV} \) then depends only on the location of \( r' \). When \( r' \geq r_3 \), \( f_{P_1}(q_1, q_s - q_1, z_r) \) is a decreasing function in \( \ell^{P_1}_{IV}(z_r) \) and \( r_6 \) in \( \ell^{P_2}_{IV}(z_r) \) \((0 < q_1 \leq r') \) is the maximizer of \( \ell_{IV} \). When \( r'_3 \leq r' < r_3 \), \( f_{P_1}(q_1, q_s - q_1, z_r) \) is quasiconcave in \( \ell^{P_1}_{IV}(z_r) \), \( r_6, r_3 \in \ell^{P_1}_{IV}(z_r) \), and \( r_3 \in \ell^{P_1}_{IV}(z_r) \) is the maximizer of \( \ell_{IV} \). Finally, when \( r' < r'_3 \), \( f_{P_1}(q_1, q_s - q_1, z_r) \) in \( \ell^{P_1}_{IV}(z_r) \) is decreasing for \( r' < q_1 < r'_3 \), increasing for \( r'_3 < q_1 < r_3 \), and decreasing again for \( r_3 < q_1 < r \). Given that the derivative of \( f_{P_1}(q_1, q_s - q_1, z_r) \) at \( r' \) is negative, \( r_6 \) must lie in the left segment of \( \ell^{P_2}_{IV}(z_r) \). In this case, \( \ell_{IV} \) has two maximizers at \( r_6 \) and \( r_3 \).

Note that, except for the last subcase above, \( f_C([q_1, q_s - q_1, z_r], q^*_1) \) in \( \ell_{IV} \) is quasiconcave for \( P(q_1) > z_r \) with a single maximizer either at \( r_3 \) or at \( r_6 \). However, to accommodate for the last subcase, the function values at \( r_6 \) and \( r_3 \) must be compared to find the maximizer of \( \ell_{IV} \). Note also that \( r_6 \) is a feasible maximizer even when it is not in \( P_2 \).

### A.3 Proof of Lemma 3.4

We analyze \( f_C(q, q^*_1) \) in \( P_1 \) and \( P_2 \):

i) When \( q \in P_1 \), \( f_C(q, q^*_1) = f_{P_1}(q_1, q_2, q_t - q_1 - q_2) \). By taking the derivatives of \( f_{P_1}(q_1, q_2, q_t - q_1 - q_2) \) with respect to \( q_1 \) and \( q_2 \), it can be shown that there is a single stationary point inside this region at \( q_1 = -b_2/a_2 \) and \( q_2 = -b_3/a_3 \), where \( a_2 = \gamma_0^2(\gamma_1 - \gamma_2), b_2 = (q_t \gamma_2 + 1) \gamma_0^2 - \gamma_1 \gamma_2, a_3 = \gamma_0^2[\gamma_0^2(\gamma_1 + \gamma_2 + 2q_t \gamma_1 \gamma_2) - \gamma_0(\gamma_1 - \gamma_2)^2 - \gamma_1 \gamma_2(\gamma_1 + \gamma_2)] \) and \( b_3 = -\gamma_0^2(q_t \gamma_2 + 1)(q_t \gamma_1 + 1) + \gamma_0^2(q_t \gamma_1 + 1)(\gamma_1 - \gamma_2) - \gamma_0(\gamma_1 \gamma_2(\gamma_1 - \gamma_2) + \gamma_0^2 \gamma_2^2) \). This point is feasible when \( q_1, q_2 > 0, q_1 + q_2 < q_t \), and \( q_2 - P(q_1) > 0 \). However, the Hessian at the above point is given by \( \nabla^2 f_{P_1}(q_1, q_2, q_t - q_1 - q_2) = \begin{pmatrix} a_4 & b_4 \\ b_4 & 0 \end{pmatrix} \) with \( b_4 \neq 0 \) when feasible. Hence, the Hessian is indefinite and the above point is a saddle and not a local maximizer.

As before, note that \( f_C(q, q^*_1) \) is not quasiconcave in \( R_2 [73] \).

ii) When \( q \in P_2 \), \( f_C(q, q^*_1) = f_{P_2}(q_1, q_2, q_t - q_1 - q_2) \). The derivative of \( f_{P_2}(q_1, q_2, q_t - q_1 - q_2) \) with respect to \( q_2 \) can be written as

\[
\frac{\partial}{\partial q_2} f_{P_2}(q_1, q_2, q_t - q_1 - q_2) = \Delta_1(q_1, q_2) + \Delta_2(q_1, q_2),
\] (A.7)
where

\[
\Delta_1(q_1, q_2) = \frac{(q_1 \gamma_1 + 1)p_5(q_1)}{[q_1 \gamma_1 + (q_t - q_1 - q_2)\gamma_2 + 1]^2},
\]

\[
\Delta_2(q_1, q_2) = \frac{-\sqrt{\gamma_0 \gamma_1 \gamma_2 q_1 (q_1 \gamma_1 + 1)} p_6(q_1, q_2)}{\sqrt{q_2(q_t - q_1 - q_2)[q_1 \gamma_1 + (q_t - q_1 - q_2)\gamma_2 + 1]^2}},
\]

(A.8)

with \(p_5(q_1) = q_1(\gamma_0 \gamma_1 - \gamma_0 \gamma_2 - \gamma_1 \gamma_2) + \gamma_0(q_c \gamma_2 + 1)\) and \(p_6(q_1, q_2) = q_2(2q_1 \gamma_1 - q_1 \gamma_2 + q_2 \gamma_2 + 2) + (q_1 - q_t)(q_1 \gamma_1 - q_1 \gamma_2 + q_2 \gamma_2 + 1)\). By equating \(\Delta_1(q_1, q_2) = -\Delta_2(q_1, q_2)\) and squaring both sides, it can be shown that the derivative in (A.7) is zero for a given \(q_2(q_1)\) is given as in (3.22). Note that \(q_2(q_1) > 0\) and \(q_1 + q_2(q_1) < q_t\) for \(0 < q_1 < q_t\). Substituting \(q_2(q_1)\) in (3.22) into \(f_{p_2}([q_1, q_2, q_t - q_1 - q_2])\), taking the derivative with respect to \(q_1\), and equating it to zero, a single stationary point can be found and is given as \(q_1\) in (3.22). This point can be shown to be in \(0 < q_1 < q_t\). Hence, the pair \([q_1, q_2(q_1)]\) in (3.22) is a stationary point in the interior of \(R_2\). Furthermore, by computing the Hessian, this point can be shown to be a local maximizer. Note that this point is a feasible maximizer even when it is not in \(P_2\).

iii) When \(q_2 = P(q_1)\), we are right at the boundary between \(P_1\) and \(P_2\). Similar to Appendix A.1, \(f_C(q, q_{12}^*)\) is continuously differentiable in \(R_2\). Hence, by continuity and from i) and ii), there is at most one local maximizer when \(q_2 = P(q_1)\).

Thus, \(f_C(q, q_{12}^*)\) is not quasiconcave and has a single local maximizer in the interior of \(R_2\).

### A.4 Proof that \(C_{BF} > C_{OAF}\)

First, note from Appendix A.3 that for a given \(q_1\), \(f_C(q, q_{12}^*) = f_{p_2}([q_1, q_2, q_t - q_1 - q_2])\) is quasiconcave with respect to \(q_2\) with its maximizer given by \(q_2(q_1)\) in in (3.22). Note also that \(f_{p_2}([q_1, q_2, q_t - q_1 - q_2])\) is quasiconcave over the line \(q_2(q_1)\) \((0 \leq q_1 \leq q_t)\) with the maximizer \(q_{1, BF}\) in (3.22). Furthermore, \(q_2(0) = q_t\) and \(q_2(q_t) = 0\). Consider the following cases depending on whether \(q_{1, OAF}\) in (3.24) reaches \(q_t\). When \(q_{1, OAF} = q_t\), \(q_{2, OAF} = z_{2, OAF} = 0\) and the point \([q_{1, OAF}, q_{2, OAF}] = [q_t, 0]\) is in the line \(q_2(q_1)\). Since \(f_{p_2}([q_1, q_2, q_t - q_1 - q_2])\) is quasiconcave over this line with the maximizer \(q_{1, BF} < q_t\), \(f_{p_2}([q_{1, BF}, q_{2, BF}(q_{1, BF}), z_{2, BF}]) > f_{p_2}([q_t, 0, 0])\). When \(q_{1, OAF} < q_t\), \(f_{p_2}([q_{1, OAF}, q_{2, q_t - q_1, OAF}])\) is quasiconcave in \(q_2\) and strictly increasing for \(0 \leq q_2 < q_2(q_{1, OAF})\). Given that \(f_{p_2}([q_1, q_2, q_t - q_1 - q_2])\) is quasiconcave
over the line \( q_2(q_1) \), 
\[
 f_{\mathcal{P}_2}(\lfloor q_1, q_2, q_1, q_2 \rfloor) \geq f_{\mathcal{P}_2}(\lfloor q_1, q_1, q_1, q_2 - q_1, q_2 \rfloor) > f_{\mathcal{P}_2}(\lfloor q_1, q_1, q_1, q_2 \rfloor). 
\]
Appendix B

Proofs for Chapter 4

B.1 Proof of Proposition 4.1

Factoring $a_0$, (4.5a) can be written as $\ln(a_0)$ plus the rate (in nats/s/Hz) of a single-input single-output system with instantaneous SNR $\omega_1/a_0$, which is given by [96, eq. (15.26)]. Also factoring $a_0$, (4.5b) can be written as $\ln(a_0)$ plus the rate of a two-branch MRC combiner with SNRs $\omega_1/a_0$ and $\omega_2/a_0$. This rate is given by [96, eq. (15.33)] when the average SNRs are equal $\mu_1 = \mu_2$ and by [95] for unequal average SNRs $\mu_1 \neq \mu_2$. Finally, (4.6a) and (4.6b) can be respectively calculated from (4.5a) and (4.5b) using the fact that $\lim_{a_0 \to 0} J(a_0/\mu) = -\gamma - \ln(a_0) + \ln(\mu)$.

B.2 Positiveness of (4.17)

Here, we show that $I_{\text{FG}}^{\text{NAF}} - I_{\text{CI}}^{\text{NAF}}$ in (4.17) is positive. When $q_1\phi_1 = z_2\phi_2$, $I_{\text{NAF}}^{\text{FG}} - I_{\text{NAF}}^{\text{CI}} \approx [1 - J(1)]/[2\ln(2)] = 0.2912 > 0$. When $q_1\phi_1 \neq z_2\phi_2$, let $\varsigma = [q_1\phi_1]/[z_2\phi_2]$. Using the fact that $J(x) < \ln(1 + [1/x])$, the difference in (4.17) can be lower bounded when $\varsigma \neq 1$ as

$$I_{\text{NAF}}^{\text{FG}} - I_{\text{NAF}}^{\text{CI}} > \frac{\varsigma \ln(1 + [1/\varsigma]) - \ln(1 + \varsigma)}{2\ln(2)(1 - \varsigma)} = \Delta(\varsigma).$$

Then, from the fact that $\ln(x) > 2([x - 1]/[x + 1])$ for $x > 1$ and $\ln(x) < 2([x - 1]/[x + 1])$ for $x < 1$ [97], it can be shown that $d\Delta(\varsigma)/d\varsigma < 0$ for $\varsigma > 1$ and $d\Delta(\varsigma)/d\varsigma > 0$ for $0 < \varsigma < 1$. Furthermore, it can be easily demonstrated that $\lim_{\varsigma \to 0} \Delta(\varsigma) = 0$ and $\lim_{\varsigma \to \infty} \Delta(\varsigma) = 0$. From these limits and the derivatives, it can be concluded that $\Delta(\varsigma) > 0$ for $\varsigma > 0$. 


Therefore, $I_{NAF}^{FG} - I_{NAF}^{CI} > \Delta(\varsigma) > 0$ for $q_1 \phi_1 \neq z_2 \phi_2$. 
Appendix C

Proofs for Chapter 5

C.1 Properties of $r_i(\cdot)$ with respect to $\mu_3$

Here, we prove three important properties of $r_i(\mu_3)$ in (5.25) with respect to $\mu_3 = 1/[2 \ln(2) \lambda_3]$. For these properties, we assume that $q_1, q_2, \gamma_1,i, \gamma_2,i, \mu_3 > 0$.

i) $\lim_{\lambda_3 \to +\infty} r_i(\mu_3) < 0$. As $\lambda_3 \to +\infty$, $\mu_3 \to 0$ and $P(z_{2,i}, \mu_3) \to P(z_{2,i}, 0) = P_1(z_{2,i})P_2(z_{2,i})$. The roots of $P_j(z_{2,i})$, $r_{1,j}$ and $r_{2,j}$, are given by

$$r_{1,j} = -\frac{q_2 \gamma_{2,i} + q_1 \gamma_{1,i} + 1}{\gamma_{j,i}(q_2 \gamma_{k,i} + 1)} , \quad \text{and} \quad r_{2,j} = -\frac{q_2 \gamma_{2,i} + q_1 \gamma_{1,i} + 1}{\gamma_{j,i}} ,$$

where $j \in \{1,2\}$ and $k = 3 - j$. Given that the above roots are strictly negative, all roots of $P(z_{2,i}, \mu_3)$ in (5.24) approach these negative values and thus $r_i(\mu_3) < 0$.

ii) $\lim_{\lambda_3 \to 0^+} r_i(\mu_3) = +\infty$. Define the quartic $Q_1(z_{2,i}, \mu_3) = [P_1(z_{2,i}) - \mu_3 A_1'][P_2(z_{2,i}) - \mu_3 A_2']$, which has the following roots:

$$p_{1,2} = \frac{-D'_1 \pm \sqrt{D'^2_1 - 4B'_1(E' - \mu_3 A'_1)}}{2B'_1},$$

$$p_{3,4} = \frac{-D'_2 \pm \sqrt{D'^2_2 - 4B'_2(E' - \mu_3 A'_2)}}{2B'_2},$$

where $j \in \{1,2\}$ and $k = 3 - j$. Given that the above roots are strictly negative, all roots of $P(z_{2,i}, \mu_3)$ in (5.24) approach these negative values and thus $r_i(\mu_3) < 0$.

The roots $p_2$ and $p_4$ are strictly positive, while $p_1$ and $p_3$ might be positive or negative. Since $Q_1(z_{2,i}, \mu_3) > P(z_{2,i}, \mu_3)$ for $\mu_3 > 0$, $r_i > \max\{p_1, p_3\}$. Note
from (C.2) that as \( \mu_3 \to \infty, p_1 \to +\infty \) and \( p_2 \to +\infty \). Consequently, \( r_i(\mu_3) \to +\infty \) as \( \lambda_3 \to 0^+ \).

iii) \( r_i(\mu_3) \) is strictly decreasing with \( \lambda_3 > 0 \). Let \( r = \max\{r_{1,1}, r_{1,2}\} \), where \( r_{1,1} \) and \( r_{1,2} \) are the roots in (C.1). Note from (5.24) that \( P(z_{2,i}, \mu_3) \) is a strictly decreasing function of \( \mu_3 > 0 \) for \( z_{2,i} > r \) since \( P_j(z_{2,i}) > 0 \). Given that \( \lim_{z_{2,i} \to +\infty} P(z_{2,i}, \mu_3) = +\infty \), \( r_i(\mu_3) \) must be strictly increasing with \( \mu_3 \) from the continuity of the function. Equivalently, \( r_i(\mu_3) \) is strictly decreasing with \( \lambda_3 > 0 \).

### C.2 Properties of \( r_i(\cdot) \) with respect to the Channel Gains

Here, we discuss some properties of \( r_i(\mu_3) \) in (5.25) with respect to channel gains \( \gamma_{1,i} \) and \( \gamma_{2,i} \). We assume that \( q_1, q_2, \mu_3 > 0 \).

i) \( r_i(\mu_3) < 0 \) when either \( \gamma_{1,i} \) or \( \gamma_{2,i} \) is small. This follows from the fact that when \( \gamma_{1,i} = 0 \) and \( \gamma_{2,i} > 0 \), or when \( \gamma_{2,i} = 0 \) and \( \gamma_{1,i} > 0 \), \( P(z_{2,i}, \mu_3) \) becomes a quadratic with a single real negative root with multiplicity two.

ii) \( r_i(\mu_3) > 0 \) when both \( \gamma_{1,i} \) and \( \gamma_{2,i} \) are large. When \( \gamma_{1,i} \) and \( \gamma_{2,i} \) are large, \( p_1 \) and \( p_3 \) in (C.2) can be approximated as

\[
\begin{align*}
p_1 & \to \frac{1}{2} \left[ -(q_2[\gamma_{2,i}/\gamma_{1,i}] + q_1) + \sqrt{(q_2[\gamma_{2,i}/\gamma_{1,i}] + q_1)^2 + 4\mu_3 (q_2[\gamma_{2,i}/\gamma_{1,i}] + q_1)} \right], \\
p_3 & \to \frac{1}{2} \left[ -(q_1[\gamma_{1,i}/\gamma_{2,i}] + q_2) + \sqrt{(q_1[\gamma_{1,i}/\gamma_{2,i}] + q_2)^2 + 4\mu_3 (q_1[\gamma_{1,i}/\gamma_{2,i}] + q_2)} \right],
\end{align*}
\]

(C.3)

which are clearly positive. Thus, \( r_i(\mu_3) > \max\{p_1, p_3\} > 0 \).

iii) When \( \gamma_{1,i} \) and \( \gamma_{2,i} \) are large, \( r_i(\mu_3) > 0 \) increases as \( \gamma_{2,i}/\gamma_{1,i} \) moves away from one. First, note that \( p_1 \) in (C.3) is a strictly increasing function of \( \gamma_{2,i}/\gamma_{1,i} \), while \( p_3 \) is strictly increasing with \( \gamma_{1,i}/\gamma_{2,i} \). Hence, \( p_1 > p_3 \) when \( \gamma_{2,i} \gg \gamma_{1,i} \). Given that \( r_i(\mu_3) > \max\{p_1, p_3\} = p_1 > 0 \), \( r_i(\mu_3) \) is increasing with \( \gamma_{2,i}/\gamma_{1,i} \). Following the same argument, when \( \gamma_{1,i} \gg \gamma_{2,i} \), \( r_i(\mu_3) > \max\{p_1, p_3\} = p_3 > 0 \) and \( r_i(\mu_3) \) is increasing with \( \gamma_{1,i}/\gamma_{2,i} \).
Appendix D

Proofs for Chapter 6

D.1 Proof of Proposition 6.1

To show that (6.32) is a strictly decreasing function of both $|u_{1,i}|^2$ and $|u_{2,i}|^2$, one needs to prove that

$$
\frac{\partial \Delta_{\pi/2,i}(|u_{1,i}|^2, |u_{2,i}|^2)}{\partial |u_{1,i}|^2} < 0, \quad \text{and} \quad \frac{\partial \Delta_{\pi/2,i}(|u_{1,i}|^2, |u_{2,i}|^2)}{\partial |u_{2,i}|^2} < 0,
$$

for $\rho = P_s/N_0 > 0$, $|u_{1,i}|^2 > 0$ and $|u_{2,i}|^2 > 0$. For convenience, let $\bar{u}_1 = |u_{1,i}|^2$, $\bar{u}_2 = |u_{2,i}|^2$, and $\bar{u} = \|u_i\|^2 = |u_{1,i}|^2 + |u_{2,i}|^2$. After some manipulations and using the fact that $x/(1+x) < \ln(1+x) < x$ for $x > 0$ [97], it can be shown that the derivative of $\Delta_{\pi/2,i}(\bar{u}_1, \bar{u}_2)$ with respect to $\bar{u}_2$ can be upper bounded by

$$
\frac{\partial \Delta_{\pi/2,i}(\bar{u}_1, \bar{u}_2)}{\partial \bar{u}_2} < \max \{ f_1(\bar{u}_1, \bar{u}_2, \rho), f_2(\bar{u}_1, \bar{u}_2, \rho) \}, \quad (D.1)
$$

where

$$
f_1(\bar{u}_1, \bar{u}_2, \rho) = \frac{-16(2\rho + 1)}{(\rho\bar{u} + 4)(4\bar{u}\rho^3 + \rho^2[\bar{u}_1^2 + 16] + 8\rho\bar{u}_1 + 16)},
$$

and

$$
f_2(\bar{u}_1, \bar{u}_2, \rho) = \frac{-16(\rho + 1)}{(\rho\bar{u} + 4)(4\bar{u}\rho^3 + \rho^2[\bar{u}_1^2 + 16] + 8\rho\bar{u}_1 + 16)}.
$$
Since both $f_1(\cdot)$ and $f_2(\cdot)$ are strictly negative in the required range, $\partial \Delta_{\pi/2,i}(\bar{u}_1, \bar{u}_2)/\partial \bar{u}_2 < 0$. Similarly, the derivative of $\Delta_{\pi/2,i}(\bar{u}_1, \bar{u}_2)$ with respect to $\bar{u}_1$ can be upper bounded by

$$\frac{\partial \Delta_{\pi/2,i}(\bar{u}_1, \bar{u}_2)}{\partial \bar{u}_1} < \max \{f_3(\bar{u}_1, \bar{u}_2, \rho), f_4(\bar{u}_1, \bar{u}_2, \rho)\},$$

(D.2)

where

$$f_3(\bar{u}_1, \bar{u}_2, \rho) = \frac{A\rho^4 + B\rho^3 + D\rho^2 + E\rho - 128}{\rho^2 (\bar{u}_1\rho + 4)(\bar{u}\rho + 4)(4\bar{u}\rho^3 + \rho^2[\bar{u}_1^2 + 16] + 8\bar{u}_1\rho + 16)},$$

with $A = 32\bar{u}_2$, $B = -8\bar{u}_1(\bar{u}_1 + 2)$, $D = -8[\bar{u}_1^2 + 8(\bar{u}_1 + 1)]$, $E = -64[\bar{u}_1 + 2]$, and

$$f_4(\bar{u}_1, \bar{u}_2, \rho) = \frac{-16 (\rho^2[2\bar{u} + \bar{u}_1] + \rho[\bar{u}_1 + 12] + 4)}{(\bar{u}_1\rho + 4)(\bar{u}\rho + 4)(4\bar{u}\rho^3 + \rho^2[\bar{u}_1^2 + 16] + 8\bar{u}_1\rho + 16)}.$$

It is not hard to see that both $f_3(\cdot)$ and $f_4(\cdot)$ are strictly negative in medium and low transmitted powers, i.e., in the turbo pinch-off region. Therefore, $\partial \Delta_{\pi/2,i}(\bar{u}_1, \bar{u}_2)/\partial \bar{u}_1 < 0$. 
Appendix E
Proofs for Chapter 8

E.1 Proof of Proposition 8.1

Since all symbols are amplified by $b$ in (2.37), $k_3 = k_4 = \cdots = k_L$ in (8.11). For sufficiently large values of $P_s$, it can be seen from the definitions that $b = O(P_s^{-1/2})$, $c_{\pi/2} = O(P_s)$, $n_1 = O(1)$, $u_l = O(1)$ and $k_3 = O(P_s^\lambda)$, where $O(\cdot)$ denotes the big-$O$ notation as $P_s \to \infty$. Hence, $P_1(\cdot)$ in (8.9) can be approximated as

$$P_1(\alpha_2) = c^2_{\pi/2} \phi_0 \phi_1 P_s \alpha_2 b^2 \left( |u_1|^4 k_3^{L-2} + |u_1|^2 k_2 k_3^{L-3} U_0 + k_1 k_3^{L-3} U_1 + k_1 k_2 k_3^{L-4} U_2 \right) + O(P_s^{\lambda(L-2)+1}),$$

where

$$U_0 = \sum_{l=3}^L |u_{l-1}|^2,$$

$$U_1 = \sum_{m=3}^L |u_2|^2 |u_{m-1}|^2 + |u_m|^2 |u_1|^2 - 2 \text{Re}[u_1^\dagger u_2 u_{m-1} u_m^\dagger],$$

$$U_2 = \sum_{l=3}^{L-1} \sum_{m=l+1}^L |u_l|^2 |u_{m-1}|^2 + |u_m|^2 |u_{l-1}|^2 - 2 \text{Re}[u_{l-1}^\dagger u_l u_{m-1} u_m^\dagger].$$

The terms $U_0$ and $U_1$ in (E.1) are valid when $L \geq 3$, while $U_2$ needs $L \geq 4$. Consider the following cases according to $\lambda$ and $L$. 
Case I: $\lambda = 0$ and $L = 3$. In this case, $k_3 = O(1)$ and
\[
P_1(\alpha_2) = c_{\pi/2}^2 \phi_0 \phi_1 |u_1|^2 P_s b^4 \{ |u_1|^2 (V + N_r) + |u_2|^2 N_r \} \alpha_2^2 + [c_{\pi/2}^2 \phi_0 \phi_1 P_s b^2 N_d U \alpha_2 + O(P_s) = A_0 \alpha_2^2 + B_0 \alpha_2 + C_0,
\]
\[
P_2(\alpha_2) = [N_d N_r P_s^2 b^4 (V + N_r)] \alpha_2^2 + [N_d^2 P_s b^2 (V + 2 N_r)] \alpha_2 + N_d^3 = D_0 \alpha_2^2 + E_0 \alpha_2 + F_0,
\]
with $U = |u_1|^4 + |u_1|^2 u_0 + U_1$. Applying partial fraction expansion to $P_2(\alpha_2)/P_1(\alpha_2)$ and taking the expectation over $\alpha_2$,
\[
\Delta_{\pi/2}(s, \bar{s}) = \frac{D_0}{A_0} + \frac{\omega_1}{\phi_2} J \left( \frac{-p_1}{\phi_2} \right) + \frac{\omega_2}{\phi_2} J \left( \frac{-p_2}{\phi_2} \right),
\]
where $p_{1,2} = [-B_0 \pm \sqrt{B_0^2 - 4A_0C_0}]/[2A_0]$ and
\[
\omega_l = \frac{(E_0 A_0 - B_0 D_0)p_l + (F_0 A_0 - C_0 D_0)}{A_0^2 (p_1 - p_k)},
\]
with $k = 3 - l$. Given that $p_1 = O(P_s^{-1})$ and $p_2 = O(1)$, $J(-p_2/\phi_2) = O(1)$ and $J(-p_1/\phi_2) = \ln(P_s) + O(1)$ from the definition of the exponential integral in (4.4). Then, since $\omega_l = O(P_s^{-2})$ and $D_0/A_0 = O(P_s^{-2})$,
\[
\Delta_{\pi/2}(s, \bar{s}) = \frac{\omega_1}{\phi_2} J \left( \frac{-p_1}{\phi_2} \right) + O(P_s^{-2}) = \frac{\omega_1}{\phi_2} \ln(P_s) + O(P_s^{-2}).
\]
Now, it is easy to show that $\omega_1 = -[F_0/(A_0 p_2)] + O(P_s^{-3})$ and $p_2 = [-B_0/A_0] + O(P_s^{-1})$. Therefore,
\[
\Delta_{\pi/2}(s, \bar{s}) = \frac{F_0 \ln(P_s)}{\phi_2 A_0 (-p_2)} + O(P_s^{-2}) = \frac{F_0 \ln(P_s)}{\phi_2 B_0} + O(P_s^{-2}). \tag{E.2}
\]
Let $\eta = \sum_{i=1}^{L-1} g_i = \sum_{i=1}^{L-1} \sum_{k=1}^{L} |g_{i,k}|^2$. Thus, when $\lambda = 0$, $b = \sqrt{(L-1)/[\eta \phi_1 P_s + O(1)]}$ and we have (8.18) by substituting $B_0$, $F_0$, $c_{\pi/2}$ and $b$ into (E.2), where the fact that
\[
U = |u_1|^4 + |u_1|^2 u_0 + U_1 = \left( \sum_{i=1}^{L} |u_i|^2 \right) \left( \sum_{i=1}^{L-1} |u_i|^2 \right) - \left| \sum_{i=1}^{L-1} u_i u_{i+1}^* \right|^2,
\]
has been used.

**Case II:** $\lambda = 0$ and $L \geq 4$. In this case, $k_3 = O(1)$ and

$$P_1(\alpha_2) = k_3^{L-4}(a_0\alpha_2^3 + b_0\alpha_2 + n_0\alpha_2) + O(P_s),$$

$$P_2(\alpha_2) = k_1 k_2 k_3^{L-2} = k_3^{L-4}(e_0\alpha_2^3 + f_0\alpha_2^2 + g_0\alpha_2 + m_0),$$

where $a_0 = O(P_s^2)$, $b_0 = O(P_s^2)$, $n_0 = c^2_{\pi/2}\phi_0 P_3 k^2 \sum d_i^2 \mathcal{U} = O(P_s^2)$, $e_0 = O(1)$, $f_0 = O(1)$, $g_0 = O(1)$ and $m_0 = N^4 d_i = O(1)$, with $\mathcal{U} = |u_1|^2 + |u_1|^2 \mathcal{U}_0 + \mathcal{U}_1 + \mathcal{U}_2$. The ratio of the polynomials can then be written as

$$\frac{P_2(\alpha_2)}{P_1(\alpha_2)} = \frac{e_0\alpha_2^3 + f_0\alpha_2^2 + g_0\alpha_2 + m_0}{a_0\alpha_2^3 + b_0\alpha_2 + n_0\alpha_2 + d_0} = \frac{e_0}{a_0} + \frac{\omega_1}{\alpha_2 - p_1} + \frac{\omega_2}{\alpha_2 - p_2} + \frac{\omega_3}{\alpha_2 - p_3},$$

where $d_0 = O(P_s)$, $p_i$ are the roots of the cubic polynomial $a_0\alpha_2^3 + b_0\alpha_2 + n_0\alpha_2 + d_0$ and

$$\omega_i = \frac{(f_0 a_0 - e_0 b_0) p_i^2 + (g_0 a_0 - e_0 n_0) p_i + (m_0 a_0 - e_0 d_0) p_i^2}{a_0^2 \prod_{i \neq l}^3 (p_i - p_l)}.$$

Taking the expectation,

$$\Delta_{\pi/2}(s, \tilde{s}) = \frac{e_0}{a_0} + \frac{\omega_1}{\phi_2} \mathcal{J} \left( \frac{-p_1}{\phi_2} \right) + \frac{\omega_2}{\phi_2} \mathcal{J} \left( \frac{-p_2}{\phi_2} \right) + \frac{\omega_3}{\phi_2} \mathcal{J} \left( \frac{-p_3}{\phi_2} \right).$$

Applying the method of dominant balance [75, Ch.3.4], it can be shown that $p_1 = O(P_s^{-1})$, $p_2 = O(1)$ and $p_3 = O(1)$. Then, given that $e_0/a_0 = O(P_s^{-2})$, $\omega_i = O(P_s^{-2})$ and $\mathcal{J}(-p_1/\phi_2) = \ln(P_s) + O(1)$ from (4.4),

$$\Delta_{\pi/2}(s, \tilde{s}) = \frac{\omega_1}{\phi_2} \mathcal{J} \left( \frac{-p_1}{\phi_2} \right) + O(P_s^{-2}) = \frac{\omega_1}{\phi_2} \ln(P_s) + O(P_s^{-2}).$$

It is easy to show that $\omega_1 = [m_0/(a_0 p_2 p_3)] + O(P_s^{-3})$. From the dominant balance method, $p_2$ and $p_3$ are the roots of the quadratic $a_0\alpha_2^2 + b_0\alpha_2 + n_0$ and thus $p_2 p_3 = n_0/a_0$. Therefore,

$$\Delta_{\pi/2}(s, \tilde{s}) = \frac{m_0 \ln(P_s)}{\phi_2 a_0 p_2 p_3} + O(P_s^{-2}) = \frac{m_0 \ln(P_s)}{\phi_2 n_0} + O(P_s^{-2}). \quad (E.3)$$

Hence, when $\lambda = 0$ and $L \geq 4$, $b = \sqrt{(L-1)/[\eta \phi_1 P_s + O(1)]}$ and we have (8.18) by
substituting \(n_0, m_0, c_{\pi/2}\) and \(b\) into (E.3), where we again use the fact that
\[
U = |u_1|^4 + |u_1|^2 U_0 + U_1 + U_2 = \left( \sum_{i=1}^{L} |u_i|^2 \right) \left( \sum_{i=1}^{L-1} |u_i|^2 \right) - \left[ \sum_{i=1}^{L-1} u_i u_{i+1}^* \right]^2.
\]

**Case III:** \(0 < \lambda \leq 1\) and \(L \geq 3\). In this case, \(k_3 = O(P_s^\lambda)\) and
\[
P_1(\alpha_2) = [c_{\pi/2}\phi \phi_1 |u_1|^4 (P_s b^2)]^{L-1} V^{L-2} |\alpha_2|^{L-1} + O(P_s^{\lambda (L-3)+2})
= D_1 \alpha_2^{L-1} + E_1,
\]
\[
P_2(\alpha_2) = [N_d (P_s b^2)]^{L-2} V^{L-2} |\alpha_2|^{L-2} + [N_d N_r (P_s b^2)]^{L-1} V^{L-2} |\alpha_2|^{L-1} + O(P_s^{\lambda (L-3)})
= A_1 \alpha_2^{L-2} + B_1 \alpha_2^{L-1} + C_1.
\]

Thus, applying long division and partial fraction expansion,
\[
\frac{P_2(\alpha_2)}{P_1(\alpha_2)} = \frac{B_1}{D_1} + \frac{1}{D_1} \cdot \frac{A_1 \alpha_2^{L-2} + C_1 - [(E_1 B_1)/D_1]}{\alpha_2^{L-1} + (E_1 / D_1)} = \frac{B_1}{D_1} + \frac{1}{D_1} \sum_{l=1}^{L-1} \frac{\omega_l}{\alpha_2 - p_l},
\]
where \(p_l = \exp \left( \frac{j \pi + 2 \pi (m - 1)!}{L-1} \right) \cdot (E_1 / D_1)^{1/(L-1)} = v_l \cdot (E_1 / D_1)^{1/(L-1)}\). Taking the expectation of the above expression,
\[
\Delta_{\pi/2}(s, \bar{s}) = \frac{B_1}{D_1} + \frac{1}{\phi_2 D_1} \sum_{l=1}^{L-1} \omega_l \mathcal{J} \left( \frac{-p_l}{\phi_2} \right).
\]

Since \((E_1 / D_1) = O(P_s^{-\lambda})\), \(\mathcal{J} (-p_l / \phi_2) = [\lambda / (L - 1)] \ln(P_s) + O(1)\) from (4.4). Furthermore, since \(A_1 (E_1 / D_1)^{(L-2)/(L-1)} = O(P_s^{\lambda (L-2)/L-1})\) and \(C_1 - [(E_1 B_1) / D_1] = O(P_s^{\lambda (L-3)})\),
\[
\omega_l = \frac{A_1 p_l^{L-2} + C_1 - [(E_1 B_1) / D_1]}{\prod_{i \neq l} (p_l - p_i)} = \frac{A_1 (E_1 / D_1)^{(L-2)/(L-1)} v_l^{L-2} + O(P_s^{\lambda (L-3)})}{(E_1 / D_1)^{(L-2)/(L-1)} \prod_{i \neq m} (v_l - v_i)}.
\]
Thus, the \( \Delta_{\pi/2}(s, \bar{s}) \) function simplifies to

\[
\Delta_{\pi/2}(s, \bar{s}) = \frac{B_1}{D_1} + \frac{1}{\phi_2 D_1} \sum_{l=1}^{L-1} A_1 \frac{(E_1/D_1)^{(L-2)/(L-1)} v_l^{L-2} + O(P_s^{\lambda(L-3)})}{(E_1/D_1)^{(L-2)/(L-1)} \prod_{i \neq l}^{L-1}(v_l - v_i)} \left[ \frac{\lambda \ln(P_s)}{L - 1} + O(1) \right]
\]

\[
= \frac{B_1}{D_1} + \frac{\lambda \ln(P_s) A_1}{\phi_2 (L - 1) D_1} \sum_{l=1}^{L-1} \frac{v_l^{L-2}}{\prod_{i \neq l}^{L-1}(v_l - v_i)} + O(P_s^{-2}).
\]

It can be shown that \( \sum_{l=1}^{L-1} \frac{|v_l^{L-2}|}{\prod_{i \neq l}^{L-1}(v_l - v_i)} = 1 \) for \( L \geq 2 \) as follows. Let \( P(x) = \frac{x^{L-2}}{x^{L-1} + 1} \). Using partial fraction expansion,

\[
P(x) = \frac{x^{L-2}}{x^{L-1} + 1} = \frac{x^{L-2}}{\prod_{l=1}^{L-1}(x - v_l)} = \sum_{l=1}^{L-1} \frac{v_l^{L-2}}{(x - v_l) \prod_{i \neq l}^{L-1}(v_l - v_i)}.
\]

Then, for \( L \geq 2 \),

\[
\lim_{x \to \infty} xP(x) = \lim_{x \to \infty} \frac{x^{L-1}}{x^{L-1} (1 + [1/x^{L-1}])} = \lim_{x \to \infty} \sum_{l=1}^{L-1} \frac{xv_l^{L-2}}{x(1 - [v_l/x]) \prod_{i \neq l}^{L-1}(v_l - v_i)}
\]

\[
= \sum_{l=1}^{L-1} \frac{v_l^{L-2}}{\prod_{i \neq l}^{L-1}(v_l - v_i)} = 1.
\]

From the above equality and the fact that \( B_1/D_1 = O(P_s^{-2}) \),

\[
\Delta_{\pi/2}(s, \bar{s}) = \frac{\lambda \ln(P_s) A_1}{\phi_2 (L - 1) D_1} + O(P_s^{-2}). \tag{E.4}
\]

Thus, when \( 0 < \lambda < 1 \), \( b = \sqrt{(L - 1)/(\eta \phi_1 P_s + O(P_s^{\lambda}))} \) and we have (8.18) by substituting \( A_1, D_1, c_{\pi/2} \) and \( b \) into (E.4). When \( \lambda = 1 \), \( b = \sqrt{(L - 1)/(\eta \phi_1 P_s + (L - 2)\beta P_s + O(1))} \) and (8.18) follows again from (E.4).
E.2 Proof of Proposition 8.2

When $0 < \lambda < 1$, $A_2/C_2 = O(P_s^{\lambda-1})$, $B_2/(\phi_2 C_2) = O(P_s^{-1})$ and $(B_2 C_2 - B_2 A_2)/(\phi_2 C_2^2) = O(P_s^{-1})$. Thus, $J(B_2/[\phi_2 C_2]) = \ln(P_s) + O(1)$ from (4.4). Hence, (8.15) can be written as

$$\Delta_{\pi/2}(s_i, \bar{s}_i) = \frac{4\beta}{\phi_1 |\epsilon_i|^2} P_s^{\lambda-1} + \frac{4N_d}{\phi_1 \phi_2 (P_s b^2)|\epsilon_i|^2} P_s^{-1} \ln(P_s) + O(P_s^{-1}).$$

Since $b = \sqrt{1/[\phi_1 P_s + O(P_s^\lambda)]}$ from (2.37),

$$\Delta_{\pi/2}(s_i, \bar{s}_i) = \frac{4\beta}{\phi_1 |\epsilon_i|^2} P_s^{\lambda-1} + \frac{4N_d}{\phi_2 |\epsilon_i|^2} P_s^{-1} \ln(P_s) + O(P_s^{-1}),$$

and (8.27) with $0 < \lambda < 1$ holds. When $\lambda = 1$, $A_2/C_2 = O(1)$, $B_2/(\phi_2 C_2) = O(P_s^{-1})$, $(B_2 C_2 - B_2 A_2)/(\phi_2 C_2^2) = O(P_s^{-1})$ and hence $J(B_2/[\phi_2 C_2]) = \ln(P_s) + O(1)$. Then, (8.15) can be written as

$$\Delta_{\pi/2}(s_i, \bar{s}_i) = \frac{4\beta}{4\beta + \phi_1 |\epsilon_i|^2} + \frac{4N_d\phi_1 |\epsilon_i|^2}{\phi_2 (P_s b^2)(4\beta + \phi_1 |\epsilon_i|^2)^2} P_s^{-1} \ln(P_s) + O(P_s^{-1}).$$

Given that $b = \sqrt{1/[\phi_1 P_s + \beta P_s + O(1)]}$ from (2.37),

$$\Delta_{\pi/2}(s_i, \bar{s}_i) = \frac{4\beta}{4\beta + \phi_1 |\epsilon_i|^2} + \frac{4N_d\phi_1 |\epsilon_i|^2 (\phi_1 + \beta)}{\phi_2 (4\beta + \phi_1 |\epsilon_i|^2)^2} P_s^{-1} \ln(P_s) + O(P_s^{-1}),$$

and (8.27) with $\lambda = 1$ holds. Finally, when $\lambda = 0$, $A_2/C_2 = O(P_s^{-1})$, $B_2/(\phi_2 C_2) = O(P_s^{-1})$ and $(B_2 C_2 - B_2 A_2)/(\phi_2 C_2^2) = O(P_s^{-1})$. Thus, $b = \sqrt{1/[\phi_1 P_s + O(1)]}$ and (8.27) with $\lambda = 0$ follows.
References


