CROSS-LAYER DESIGN FOR WIRELESS COOPERATIVE NETWORKS

by

Yao Xiao

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Winter 2014

© 2014 Yao Xiao
All Rights Reserved
CROSS-LAYER DESIGN FOR WIRELESS COOPERATIVE NETWORKS

by

Yao Xiao

Approved: ________________________________
Kenneth E. Barner, Ph.D.
Chair of the Department of Electrical and Computer Engineering

Approved: ________________________________
Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: ________________________________
James G. Richards, Ph.D.
Vice Provost for Graduate and Professional Education
I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ________________________________
Leonard J. Cimini, Jr., Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ________________________________
Chien-Chung Shen, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ________________________________
Javier Garcia-Frias, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ________________________________
John M. Walsh, Ph.D.
Member of dissertation committee
ACKNOWLEDGEMENTS

During the course of my graduate study at University of Delaware, I received help and support from many people, and this Ph.D. dissertation would not have been possible without them. First of all, I would like to thank my advisor Prof. Leonard Cimini for his great support and guidance throughout the years. His boundless energy, continual encouragement, and devotion to his students have been truly inspirational. Prof. Cimini always showed great faith in my abilities and allowed me to work quite independently, but at the same time provided invaluable guidance at the necessary times.

I am also grateful to Prof. Chien-Chung Shen for serving as my associate advisor. Prof. Shen has always been a wonderful reference and supporter for my research. I have learned a great deal of knowledge, especially on wireless networking, through the collaborative work with Prof. Shen and his group. I would also like to thank Professors Javier Garcia-Frias and John Walsh for serving on my dissertation committee. Prof. Garcia-Frias has been an inspiring resource and teacher. His courses on information theory and channel coding greatly broadened my depth of knowledge of that field. Prof. Walsh has been leading the ongoing collaboration between the Delaware and Drexel groups on overhead-performance analysis. His valuable comments on my proposal have been very helpful in preparing this dissertation.

I am also deeply indebted to current and past members of the Wireless Systems Lab, including Hao Feng, Qi Wang, Gubong Lim, Chenzi Jiang, Xiantao Sun, Hongzheng Wang, Wooyoung Ryu, Erdem Bala, Lu Zhang, Bo Gui, Guangyi Liu, Li Li, Bohan Zhang, Haina Ye, and Chengming Zhou. I am also thankful to my collaborators in the DEGAS Networking Group at Delaware, Prof. Garcia-Frias’s group at Delaware, and the Adaptive Signal Processing and Information Theory Research Group at Drexel.
including Yang Guan, Gwanmo Ku, Jie Ren, Bradford Boyle, Marcos Portnoi, Bo Lu, Lu Li, Zequn Huang, Wei Chen, and Mohamed Hassanin.

Of course, I have to thank my friends, inside and outside of University of Delaware, for making these years so enjoyable. I would also like to thank my parents for inspiring a love of learning, particularly in math and engineering. I can’t imagine coming so far without their love and dedication.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xiv</td>
</tr>
</tbody>
</table>

## Chapter

### 1 INTRODUCTION

1.1 Background and Motivation

1.2 Cross-Layer Design for Wireless Cooperative Networks

  1.2.1 Overhead-Performance Trade-off
  1.2.2 Interference Management

1.3 Outline of Dissertation

### 2 PERFORMANCE-OVERHEAD TRADE-OFF: SPECTRAL EFFICIENCY

2.1 Background and Related Work

2.2 System Model

2.3 Overhead Analysis

  2.3.1 Channel Estimation Overhead
  2.3.2 Relay Selection Overhead
  2.3.3 Transmission Coordination Overhead

2.4 Spectral Efficiency

  2.4.1 Timer-Based Best-Select Relaying (TBBS)
  2.4.2 Dis-STC Cooperative Relaying (Dis-STC)
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.3</td>
<td>$M$-Group Dis-STC Cooperative Relaying ($M$-group)</td>
<td>26</td>
</tr>
<tr>
<td>2.5</td>
<td>Simulation Results</td>
<td>29</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Impact of Relay Selection Overhead</td>
<td>30</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Impact of Channel Estimation Overhead</td>
<td>34</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Impact of Coordination Overhead</td>
<td>34</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Performance Comparison</td>
<td>35</td>
</tr>
<tr>
<td>2.5.5</td>
<td>A Note on Randomly Distributed Relays</td>
<td>39</td>
</tr>
<tr>
<td>2.6</td>
<td>Conclusions</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>PERFORMANCE-OVERHEAD TRADE-OFF: ENERGY EFFICIENCY</td>
<td>43</td>
</tr>
<tr>
<td>3.1</td>
<td>Background and Related Work</td>
<td>43</td>
</tr>
<tr>
<td>3.2</td>
<td>System Model</td>
<td>44</td>
</tr>
<tr>
<td>3.3</td>
<td>Energy Efficiency</td>
<td>45</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Ideal Best-Select Relaying</td>
<td>46</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Timer-Based Best-Select Relaying (TBBS)</td>
<td>47</td>
</tr>
<tr>
<td>3.3.3</td>
<td>$M$-Group Dis-STBC All-Select Relaying ($M$-group)</td>
<td>48</td>
</tr>
<tr>
<td>3.4</td>
<td>Simulation Results</td>
<td>49</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Impact of Selection Process</td>
<td>49</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Impact of Energy Consumption Profiles</td>
<td>51</td>
</tr>
<tr>
<td>3.5</td>
<td>Conclusions</td>
<td>54</td>
</tr>
<tr>
<td>4</td>
<td>COOPERATIVE ROUTING WITH MINIMUM OVERHEAD</td>
<td>56</td>
</tr>
<tr>
<td>4.1</td>
<td>Background and Related Work</td>
<td>56</td>
</tr>
<tr>
<td>4.2</td>
<td>Basic Concept of LACR</td>
<td>58</td>
</tr>
<tr>
<td>4.3</td>
<td>Analysis of Design Issues</td>
<td>59</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Cooperative Transmission Strategy</td>
<td>59</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Capability to Cooperate</td>
<td>60</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Performance Metric</td>
<td>64</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Selection Method</td>
<td>64</td>
</tr>
<tr>
<td>4.4</td>
<td>Protocol Description</td>
<td>65</td>
</tr>
</tbody>
</table>
4.5 Simulation Results ........................................... 67
   4.5.1 Setting .................................................. 67
   4.5.2 Results and Discussion .............................. 67
      4.5.2.1 Probability of Successful Route Discovery .... 67
      4.5.2.2 Packet Delivery Ratio ......................... 69
      4.5.2.3 End-to-End Delay .............................. 70
4.6 Conclusions ................................................. 71

5 INTERFERENCE MANAGEMENT: SPATIAL REUSE SCHEDULING ............. 73
   5.1 Background and Related Work .......................... 73
   5.2 Multi-Hop Cooperative Networks ...................... 74
   5.3 Dis-STC Based Cooperation ............................ 76
      5.3.1 Impact of Cooperation ............................ 76
      5.3.2 Impact of Spatial Reuse ......................... 77
      5.3.3 Optimal Spatial Reuse Scheduling ............... 79
   5.4 Incorporation of Cooperative Beamforming .......... 80
      5.4.1 Impact of Cooperative Beamforming ............. 81
      5.4.2 Optimal Reuse Scheduling ....................... 83
      5.4.3 Impact of Channel Estimation Overhead and Imperfect CSI ...... 84
   5.5 Performance Evaluation and Discussion ............. 87
      5.5.1 Impact of Cooperation ............................ 87
      5.5.2 Impact of Interference Mitigation ............... 90
      5.5.3 Impact of Imperfect CSI .......................... 91
   5.6 Conclusions ............................................. 93

6 CONTRIBUTIONS AND FUTURE WORK .................................. 95
   6.1 Contributions ............................................ 95
      6.1.1 Overhead-Performance Trade-off .................. 95
      6.1.2 Cooperative Routing with Minimum Overhead .... 95
6.1.3 Optimal Spatial Reuse for Wireless Networks .......................... 96

6.2 Future Work .................................................................................. 97

6.2.1 Overhead-Performance Trade-off Analysis: Information Theoretic Approach ............................................................... 97
6.2.2 Efficient Heterogeneous Cooperative Networks ...................... 97

BIBLIOGRAPHY .................................................................................. 99
LIST OF TABLES

3.1 Transceiver power consumption (mW) . . . . . . . . . . . . . 46
4.1 Cooperative Transmission Range . . . . . . . . . . . . . . . . 64
**LIST OF FIGURES**

1.1 Cooperation creates a VMISO link ........................................... 3

1.2 Cooperative relaying (packet relaying) increases data rate ............ 4

1.3 Best-select cooperative relaying chooses the best node to relay the source information ......................................................... 6

1.4 Illustration of performance-overhead trade-off ............................ 7

2.1 Two-phase relaying scheme: In the first phase, the source broadcasts its message to all cooperating nodes; in the second phase, the selected relay(s) forward(s) the source information to the destination. 14

2.2 Suppose the source broadcasts the signal at time 0. Due to different distances to the source, $R_1$ receives the source message at $t_{d_1}$ while $R_2$ receives it at $t_{d_2}$. Since $R_1$ is the “better” relay, it has the “shorter” timer that expires at $t_{d_1} + t_1$. Then, $R_1$ begins preparing the to-be-transmitted message, which takes $t_{proc}$ to complete, and then rebroadcasts the source information. After another $t_{prop}$, $R_2$ receives the signal broadcasted by $R_1$ and backs off. 18

2.3 Collision probability as a function of $\mu$ with different sizes of the decoded set (solid curves with $L = 5$ and dashed curves with $L = 10$) and different numbers of training symbols (different markers that are corresponding to $N_t = 1, 5, 10$). 31

2.4 Expected selection time as a function of $\mu$ with different sizes of the decoded set (solid curves with $L = 5$ and dashed curves with $L = 10$) and different numbers of training symbols (different markers that are corresponding to $N_t = 1, 5, 10$). 32

2.5 Spectral efficiency of TBBS, including relay selection overhead, as a function of $N$ with different values of $\mu$ ................................. 33
2.6 Impact of channel estimation overhead: spectral efficiency as a function of the number of symbols used in one training sequence ($\theta = 3$ dB, $N = 5$) .......................................................... 35

2.7 Impact of channel estimation overhead: spectral efficiency as a function of the number of symbols used in one training sequence ($\theta = 5$ dB, $N = 5$) .......................................................... 36

2.8 Spectral efficiency as a function of $\theta$ ($N_t = 2, N = 5$) .......................... 37

2.9 Spectral efficiency as a function of $N$ ($N_t = 2, \theta = 5$ dB) ............... 38

2.10 Spectral efficiency as a function of $N$ with randomly distributed relays ($N_t = 2, \theta = 5$ dB) ............................ 39

3.1 Energy efficiency of 2-group and TBBS as a function of $N$, the number of potential relays .................................................. 50

3.2 Energy efficiency of 2-group and TBBS as a function of $T_s/T$, the ratio of selection time to total transmission time. (The energy efficiencies are normalized by the efficiency for ideal Best-Select with $N = 3$.) ................................. 51

3.3 Energy efficiency of 2-group and TBBS as a function of the ratio of $P_r$ to $P_t$ ($P_{non} = 0$) ............................................. 52

3.4 Energy efficiency of 2-group and TBBS as a function of the ratio of $P_r$ to $P_t$ ($P_{non} = 0.5P_t$) ............................................. 53

3.5 Energy efficiency of 2-group and TBBS as a function of the ratio of $P_{non}$ to $P_t$ ($P_r = 0.5P_t$) ............................................. 54

4.1 The impact of distance on outage probability when 2-group Dis-STC is used as cooperative transmission strategy ....................... 63

4.2 Probability of successful route discovery as a function of the number of nodes ................................................................. 68

4.3 Route setup delay as a function of the number of nodes .................... 69

4.4 Packet delivery ratio as a function of the number of nodes ............... 70

4.5 Average end-to-end delay as a function of the number of nodes ....... 71
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>A multi-hop linear wireless cooperative network</td>
<td>75</td>
</tr>
<tr>
<td>5.2</td>
<td>Spatial reuse in linear cooperative networks</td>
<td>78</td>
</tr>
<tr>
<td>5.3</td>
<td>Normalized network throughput as a function of spatial reuse factor</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>(Dis-STC is used)</td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>Normalized network throughput as a function of spatial reuse factor</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>with randomly distributed nodes (Dis-STC is used)</td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>Normalized network throughput as a function of spatial reuse factor</td>
<td>90</td>
</tr>
<tr>
<td>5.6</td>
<td>Normalized network throughput as a function of $M$ ($N=12$, ZFB is</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>used)</td>
<td></td>
</tr>
<tr>
<td>5.7</td>
<td>Normalized network throughput as a function of spatial reuse factor</td>
<td>93</td>
</tr>
</tbody>
</table>
Cooperative communications has been shown to be effective in combating fading in wireless channels. In order to realize the potential benefits promised by cooperative communications in practical wireless networks, careful cross-layer design is essential. In particular, investigations of the interactions among the physical, link, and network layers are critical in developing cooperation-enabled MAC or routing protocols. Recently, cross-layer design for cooperative networks has become a very active research area; many efficient and elegant cooperative networking techniques that provide significant performance gains have been proposed.

Cooperation, however, also introduces new challenges, including increased overhead and interference, that must be addressed in order to efficiently implement cooperative networks. Although cooperation is promising in improving performance, it requires much more overhead to implement compared with conventional point-to-point communications. In addition, with cooperation, the interference environment will change; this means that new interference management techniques are required.

In this dissertation, we focus on the overhead-performance trade-off and interference management for wireless cooperative networks. The overhead in implementing cooperation, particularly the overhead to estimate the channels, select the relay(s), and coordinate the transmissions, is investigated. Taking into account the overhead, the spectral and energy efficiencies of several different relaying schemes are studied. Through analyses and simulations, we demonstrate the impact of the overhead on these efficiencies, and provide guidelines for determining the appropriate cooperative scheme for specific applications.
In order to realize the promised benefits of cooperative communications in practical wireless networks, cooperation-enabled routing algorithms are essential. Motivated by the analysis of the performance-overhead trade-off for cooperative relaying, we propose a novel cooperative routing algorithm that reduces the amount of overhead incurred in implementations and hence provides a significant performance gain. Specifically, we describe Location-Aware Cooperative Routing (LACR), a routing protocol for multi-hop wireless networks, which incorporates cooperative transmissions into geographical routing. Simulation results show that LACR performs well, providing a high probability of discovering a route with low overhead; these advantages are particularly apparent when the network is sparse.

In addition to requiring more overhead, employing cooperation in wireless networks affects the performance of existing interference management techniques. Focusing on multi-hop linear cooperative networks, we investigate the impact of cooperation on spatial reuse scheduling. The impact of enabling spatial reuse and incorporating cooperation on the network throughput is studied. Through analyses and simulations, we show that appropriate reuse is critical for efficient networks, and incorporating cooperation is effective in improving the network performance when the direct links are suffering high outages. Furthermore, the reuse factor that maximizes the network throughput is derived, the performance of which is illustrated via simulations.
Chapter 1
INTRODUCTION

1.1 Background and Motivation

Wireless communications has evolved very rapidly during the past three decades to the point where it is now considered an essential component of people’s daily lives. It is well known that high-speed wireless communications is challenging because of the harsh characteristics of the medium, including significant attenuation, time variation, and interference [1–3].

In typical wireless environments, there are many reflectors and scatters. Therefore, the signal that arrives at the receiver is generally a superposition of signals that have traveled through many different propagation paths. These signals can add constructively or destructively, causing fluctuations in the amplitude of the received signal; this effect is called fading [3]. When a channel is in a deep fade, the received signal can easily fall below the threshold where it can be reliably decoded. Diversity combining is one powerful technique to combat this impairment. Essentially, the basic idea is to combine signals that encounter statistically independent fading, for example, the signals received at two antennas which are sufficiently separated in space. Because of the extremely low probability that all the independently faded signals are simultaneously experiencing deep fading, diversity combining can significantly improve the probability of successful reception. Some well-known forms of diversity are spatial diversity, time diversity, and frequency diversity [1–5].

Using multiple antennas at the transmitter and/or receiver side is a standard method for achieving spatial diversity to combat fading [6–8]. Diversity in such multiple-antenna systems, which are also known as Multiple-Input-Multiple-Output (MIMO)
systems, can be obtained through antenna selection [9], the use of space-time codes [10, 11], or the intelligent use of channel state information (CSI) at the transmitter (e.g., maximum ratio transmission [12]). Employing multiple antennas can also significantly increase the capacity via spatial multiplexing [13]. Of course, a combination of diversity and multiplexing is also possible [14]. Due to the well-recognized significant diversity and multiplexing gains, MIMO techniques have been widely deployed, or are envisioned to be implemented, in practical systems such as 3GPP LTE and IEEE 802.11 [15–18].

In some scenarios, however, due to the size and/or cost limitations, deploying multiple antennas may not be feasible, for example, for mobile handsets. Cooperative communications has been proposed, in this case, to provide spatial diversity and/or multiplexing gain by constructing a virtual antenna array from multiple single-antenna nodes [19–25]. For example, we consider a system with one source (S) – destination (D) pair, as shown in Fig. 1.1. The traditional way to send information from S to D is by a direct link. If both terminals are equipped with a single antenna, this is a Single-Input-Single-Output (SISO) link; in this case, if the S–D separation is large and/or the channel is in a deep fade, there is a high probability that the source information cannot successfully be received at the destination. To increase the probability of successful transmission, S can first broadcast the message to its neighboring nodes, which are close enough and thus have good communication quality; then, together with the neighboring nodes, S can form a virtual multiple-antenna terminal to transmit the information to the destination. This cooperation among S and its neighboring nodes creates a virtual Multiple-Input-Single-Output (VMISO) channel and can provide significant diversity gain.

Similar to MIMO techniques, cooperative communications has been shown to be effective in providing diversity gain, which, in turn, can improve the reliability of the communication, increase the data rate, and/or reduce the required transmitted power [26–28]. To fully realize the potential performance gains promised by cooperative communications, it is essential to investigate cooperative communications in a network
context. In this work, we apply a cross-layer approach in studying the design of wireless cooperative networks.

1.2 Cross-Layer Design for Wireless Cooperative Networks

Cross-layer design for wireless networks, where the useful interactions among different layers across the protocol stack are exploited to improve the system performance, has been an active research area [29–32]. Recently, to efficiently implement cooperative communications in wireless networks, there has been a growing interest in investigating cooperation in a systematic manner, in particular, using cross-layer approaches [33–45].

In conventional wireless systems, any unicast communication involves only two parties. In order to enhance performance by applying cooperative communications, new medium access control (MAC) protocols are required. In [37], based upon the IEEE 802.11 distributed coordination function (DCF) mode, the authors propose and implement a new MAC protocol called relay-enabled DCF, where packet relaying (i.e., cooperative relaying) is applied to mitigate the throughput bottleneck caused by low-data-rate links. For example, as shown in Fig. 1.2, suppose the direct link from $S$ to $D$ can only support a maximum data rate of 1 Mbps (due to the large $S$-$D$ separation and/or deep fading); utilizing the relay node $R$ which has much better channels (from $S$ to $R$ and from $R$ to $D$) improves the data rate between $S$ and $D$ to 5.5 Mbps. A similar relay-based MAC protocol named CoopMAC is proposed in [38]; here, the authors take a step further to investigate the impact of cooperation on inter-flow interference.
In [39], the authors focus on the design of efficient relay selection schemes that adapt to the channel characteristics and payload length. It has been shown in [37–39], through simulations and experiments, that a cooperation-enabled MAC significantly increases the network throughput, while also reducing the end-to-end delay.

![Diagram of cooperative relaying](image)

**Figure 1.2:** Cooperative relaying (packet relaying) increases data rate

In addition to a MAC protocol that enables cooperation, a cooperative routing algorithm is also essential in achieving the promised benefits of cooperation [40–45]. In [40, 41], the authors formulate cooperative routing as an optimization problem, and propose routing algorithms that are based on global optima; these algorithms are inherently centralized and thus not scalable. In [43], the authors divide the cooperative routing problem into two tasks: (1) find a primary SISO route and (2) shorten the SISO route by selecting relays to construct a cooperative path. As demonstrated through simulations and experiments, compared with SISO routing, cooperative routing has a higher route discovery rate, a higher packet delivery ratio, and a smaller end-to-end delay.

Cross-layer design for wireless cooperative networks, especially considering the interactions among the physical, MAC, and network layers, has been shown to be critical in efficiently implementing cooperative communications in practical networks [46–49]. On the one hand, as discussed above, significant performance gains have been observed when applying carefully designed cooperative schemes (e.g., cooperative MAC
and cooperative routing). On the other hand, as pointed out in [35, 49], incorporating cooperation also introduces potential drawbacks that might degrade the network performance. In particular, compared with conventional point-to-point communications, cooperation inherently requires more overhead for acquiring CSI as well as coordinating transmissions. In addition, the inclusion of cooperative communications changes the interference characteristics, which requires modifications or even a fully new design of the currently applied interference mitigation techniques (e.g., spatial reuse). Therefore, the focus of our work is on studying the impact of overhead and interference on cross-layer design for wireless cooperative networks.

1.2.1 Overhead-Performance Trade-off

One well-recognized and well-studied overhead that is incurred in cooperative communications is the requirement of one extra transmission in sending one packet from the source to the destination. Specifically, since the cooperating nodes are distributed in space, as shown in Fig. 1.1, it requires two transmissions to send one packet, where the first is used to disseminate the source information among the cooperating nodes and the second is used to transmit from the cooperating nodes to the destination. In [50], factoring in this overhead, the authors study the spectral efficiency of conventional point-to-point communications and cooperative communications, aiming to answer the question “When is cooperative relaying more efficient than direct communication?”

In addition, cooperation requires more CSI and more control signals, which implies a larger amount of channel estimation overhead and transmission coordination overhead. The actual amount of overhead depends on the applied cooperation techniques; we take best-select cooperative relaying as an example in this discussion. As shown in Fig. 1.3, best-select relaying (BSR) chooses the best relay, that is, the one which has the highest channel power gain to the destination, to forward the source information. One possible centralized BSR works as follows: every relay transmits a training sequence to the destination after decoding the source signal; then the destination determines the best node (based on the collected CSI) and feeds back the index
of the chosen relay. Clearly, the number of required channel estimation sequences increases linearly with the number of relays. Also, control signals are needed to coordinate the transmission of the training sequences from the relays to the destination, and in feeding back the index of the chosen relay. Therefore, in order to reveal the true performance gain (that is, the performance gain taking into account the impact of the incurred overhead) provided by BSR, it is essential to quantify the impact of the incurred overhead.

![Diagram](image)

**Figure 1.3:** Best-select cooperative relaying chooses the best node to relay the source information

Conceptually, there exists a trade-off between the overhead (the resources consumed in acquiring the required information) and the true performance with consideration of overhead; one possible illustration is shown in Fig. 1.4. When no/low overhead is allowed, that is, no/little CSI can be collected, intuitively, only fixed-relay or random relay selection can be applied, which usually has poor performance since no diversity can be achieved (see the leftmost region in Fig. 1.4). On the contrary, allowing high overhead, e.g., collecting all CSI at the destination, guarantees that the best relay gets selected; however, the high overhead degrades the performance because it consumes an overwhelming amount of resources. So, the true performance with high overhead might also be very poor (see the rightmost region in Fig. 1.4). In this work, we aim
to understand the impact of overhead in a quantitative manner, then investigate the
performance-overhead trade-off, and finally determine the optimal operating region
(e.g., the middle section in Fig. 1.4) and propose practical algorithms that can achieve
the optimal performance.

![Illustration of performance-overhead trade-off](image)

**Figure 1.4:** Illustration of performance-overhead trade-off

We believe that a comprehensive understanding of the overhead-performance
trade-off is essential to better determine the true system performance. Hence, under-
standing this trade-off is critical for system designers to architect efficient cooperative
networks. In Chapters 2 and 3, we conduct a quantitative analysis of the overhead as
well as the overhead-performance trade-off, focusing on wireless cooperative networks,
and demonstrate the impact of this trade-off on the network performance.

### 1.2.2 Interference Management

It is well-recognized that efficient interference management is important in
achieving high-throughput wireless networks. In practical wireless infrastructures,
frequency/spatial reuse is used to reduce/mitigate the interference. As pointed out
in [35, 51], however, the increased interference caused by cooperation limits the opportunities for spatial reuse of the spectrum, which then significantly degrades the network throughput. To avoid this performance degradation, we need to apply interference-aware power control [52–55] and/or a cooperation-aware spatial reuse design.

In wireless networks, spatial reuse is essential to balance the interference level and the channel utilization [56]. On the one hand, spatial reuse improves the channel utilization by enabling concurrent transmissions, which increases network throughput. On the other hand, spatial reuse introduces co-channel interference, which degrades performance. Therefore, determining the optimal spatial reuse is critical for high-throughput wireless networks. Although reuse in wireless networks, especially multi-hop linear networks, is a well-studied topic [57–62], the impact of incorporating cooperation on spatial reuse scheduling is an open problem.

As discussed in [63, 64], cooperative communications can be effective in mitigating interference. In [64], the authors design and implement 802.11n+, an enhanced version of 802.11n that utilizes cooperative beamforming [65] and interference alignment [66] to cancel the co-channel interference, and hence enables concurrent transmissions without harming the ongoing communications. A significant performance gain is achieved, compared with the existing 802.11n networks. Inspired by this work, we include cooperative beamforming into our spatial reuse design to study the optimal interference-aware utilization of cooperation in networks. For example, one fundamental trade-off is the allocation of the available degrees of freedom between beamforming and diversity. Specifically, on the one hand, beamforming mitigates/reduces the interference, which has the potential to improve the network performance; on the other hand, the degrees of freedom that are consumed by beamforming can be otherwise utilized to provide a diversity gain.

In Chapter 5, we investigate the impact that cooperative communications has on the spatial reuse design for wireless cooperative networks, and derive the optimal reuse factor that maximizes the network throughput. In order to actively mitigate the interference, the incorporation of cooperative beamforming is also studied. A joint
optimization problem, with consideration of the overhead incurred in implementing beamforming, is formulated. Through analyses and simulations, we show the importance of an appropriate spatial reuse for efficient networks, and provide guidelines in determining reuse scheduling that maximizes the network throughput.

1.3 Outline of Dissertation

In this dissertation, we investigate cross-layer design for practical cooperative networks, focusing on the overhead and interference management; we present the analysis of the overhead-performance trade-off, as well as the optimal spatial reuse for wireless linear multi-hop networks. We start with an investigation of the impact of overhead. Based on this quantitative analysis, we present the overhead-performance trade-off for cooperative relaying and demonstrate how to utilize this trade-off in designing practical cooperative systems. In addition, aiming to balance the interference and channel utilization, we derive the optimal spatial reuse scheduling for multi-hop linear cooperative networks. These results reveal some important insights of cooperative networking design, and bring closer the realization of high-throughput wireless networks. The rest of this dissertation is organized as follows.

In Chapter 2, we investigate the overhead-performance trade-off for cooperative relaying with a focus on spectral efficiency. The quantitative analysis of the overhead in implementing cooperative relaying is presented; in particular, the overhead required to acquire essential CSI, to select the relay(s), and to provide the required coordination is included. Factoring in the impact of overhead, the spectral efficiency of cooperative relaying schemes is then studied. Extensive simulation results are provided to illustrate the impact of overhead on the spectral efficiency. More importantly, based on the analyses and simulations, we provide guidelines for determining the appropriate cooperative relaying scheme for specific scenarios.

In Chapter 3, we focus on the impact of the overhead-performance trade-off on energy efficiency. Although the results in Chapter 2 suggest that $M$-group All-Select relaying is spectrally efficient, it consumes a large amount of energy in the second phase
because of the fact that all the nodes in the decoded set participate in forwarding the signal. In order to evaluate the energy efficiency of All-Select relaying and Best-Select relaying, we quantitatively analyze the impact of overhead on the efficiency and compare several practical cooperative relaying schemes. Through analyses and simulations, we identify the critical system parameters that affect the efficiency and provide guidelines for determining a suitable relaying scheme for specific applications.

Motivated by the analysis of the overhead-performance trade-off, in Chapter 4, we propose cooperative routing protocols that incur the minimum amount of overhead. Specifically, we incorporate cooperative communications into geographic routing and propose Location-Aware Cooperative Routing, which provides cooperative diversity with a minimum amount of overhead incurred. Through simulations, we show that our proposed cooperative routing protocol provides a higher route discovery ratio as well as a higher throughput compared with SISO geographic routing and SISO-based cooperative routing.

In Chapter 5, we study interference management for wireless cooperative networks, in particular, spatial reuse scheduling for multi-hop linear cooperative networks. We start by investigating the impact of cooperation on the throughput of multi-hop linear networks, and then derive the reuse factor that maximizes the network throughput. Simulation results are provided to illustrate the importance of choosing an appropriate reuse factor, and to demonstrate the significant performance gains achieved when using the optimal reuse scheduling.

Finally, Chapter 6 summarizes the contributions of this dissertation and discusses future potential research topics.
Chapter 2
PERFORMANCE-OVERHEAD TRADE-OFF: SPECTRAL EFFICIENCY

It is well known that cooperative relaying has the potential to improve the performance of wireless communications. However, compared with point-to-point communications, cooperative relaying requires much more overhead to implement. In this chapter, we quantitatively analyze the overhead in implementing cooperative relaying, including the overhead to acquire essential CSI, to select the relay(s), and to provide the required coordination. Factoring in the impact of overhead, the spectral efficiency of cooperative relaying schemes is then investigated. A comparison of three cooperative relaying schemes (Timer-Based Best-Select, Distributed Space-Time-Coded, and \( M \)-group) is presented. Numerical results are provided to illustrate the impact of overhead on the spectral efficiency. Finally, based on the results of the analyses and simulations, we provide guidelines for determining the appropriate cooperative relaying scheme for specific scenarios.

2.1 Background and Related Work

Cooperative relaying, which can provide spatial diversity by constructing a virtual antenna array from multiple single-antenna nodes, has been shown to be effective in combating fading on wireless channels [19, 20, 22, 23, 25]. In order to enable communication from a source to a destination over a poor quality link, for example, one with high outage, one or more nodes can be employed to relay the source message to the destination. At the relays, different relaying schemes can be performed [23]; here, we assume a Decode-and-Forward relaying scheme. In addition, for the sake of scalability and to minimize the overhead, we focus on distributed cooperative relaying
algorithms. Two questions must be answered in designing such cooperative relaying schemes: (i) How to choose the forwarding relay(s) out of all available cooperating nodes? and (ii) How to coordinate the transmission(s) from the selected relay(s) to the destination? Several cooperative relaying strategies have been proposed and evaluated (for example, see [21, 52, 67–69]).

Timer-Based Best-Select relaying (TBBS) [67] is a popular selective cooperative relaying approach, where only the best node is selected to forward the source information. Based on the performance metric, every eligible cooperating node sets up an individual timer so that the best node has the shortest one. Once the first timer expires (at the best node), it begins transmitting, and the other nodes back off. For example, a node that successfully decodes the source information sets its timer inversely proportional to its channel power gain to the destination. By doing so, the best node, that is, the one with decoded source information and the highest gain to the destination, will be selected.

In contrast to a selective relaying scheme such as TBBS, Distributed Space-Time-Coded relaying (Dis-STC) [21] recruits all the eligible cooperating nodes and applies STC to coordinate the simultaneous transmissions from the multiple relays. Notice that the critical step in TBBS is the relay selection (the first question), while for Dis-STC, it is the coordination of the multiple transmissions (the second question). However, it has been shown that TBBS and Dis-STC achieve comparable performance, and, in particular, both can achieve the full diversity order [21, 67, 70].

The benefits provided by cooperative relaying have already been well established [19, 23, 70, 71], but, the overhead incurred in implementing cooperative relaying is usually overlooked. In practice, both the relay selection and the transmission coordination require control information. For example, in TBBS, the relay selection process introduces possible collisions and selection time consumption [67, 72]. In Dis-STC, on the other hand, the relay selection is by default since all the cooperating nodes are selected, but the coordination using STC requires significant overhead [72]. In addition, when more nodes are involved, more CSI is required, incurring more overhead for
channel estimation.

A few papers have focused on the overhead in implementing cooperative relaying. In [68], $M$-group Dis-STC relaying ($M$-group), a special case of randomized Dis-STC [69], is proposed to minimize the amount of control information for Dis-STC. In $M$-group, all the eligible cooperating nodes are employed, but in a random manner instead of the conventional STC used in Dis-STC. This random coordination approach enables a distributed implementation of cooperative relaying, dramatically reducing the overhead. In [73], the overhead in TBBS is investigated and one optimal timer-based selection scheme is proposed with the purpose of bounding the selection overhead. In [72, 74], we present an initial study of the impact of the overhead required for relay selection and channel estimation on the performance of cooperative relaying. These preliminary results show that the impact of this overhead is not negligible, but rather is critical in determining the true performance improvement provided by cooperative relaying.

In this chapter, we investigate the overhead incurred in implementing cooperative relaying, particularly, that required for channel estimation, relay selection, and coordinating the simultaneous multiple transmissions. To study the impact of different overhead components, we investigate the spectral efficiencies, taking into account the overhead, of three cooperative relaying schemes, namely, TBBS, Dis-STC, and $M$-group. Through analyses and simulations focusing on the performance-overhead trade-offs, we show (i) that the consideration of overhead in designing cooperative relaying schemes is necessary to achieve good performance in practical implementations; and (ii) how to choose among TBBS, Dis-STC, and $M$-group to maximize the true performance, given the system parameters that determine the overhead level (e.g., network size).

The rest of the chapter is organized as follows. In Section 2.2, the system model is presented, followed by detailed descriptions of the three cooperative relaying schemes: TBBS, Dis-STC, and $M$-group. In Section 2.3, the overhead to acquire the CSI, select the relay(s), and coordinate the cooperative communications is quantitatively analyzed.
Then, the impact of the overhead on the spectral efficiency is investigated. Results that illustrate the various trade-offs are presented and discussed in Section 2.5. Section 2.6 concludes this chapter.

2.2 System Model

We consider a system with one source-destination pair and $N$ cooperating nodes, where the direct link from the source to the destination is assumed to be unreliable due to the large source-destination separation and/or the presence of deep fading. Two-phase cooperative relaying is applied to facilitate the communication, as shown in Fig. 2.1. In the first phase, the source broadcasts the signal and all the cooperating nodes listen. If relay $i$, $1 \leq i \leq N$, has a higher signal-to-noise ratio (SNR) than the specified threshold $\gamma_{\text{th}}$, we assume that it can successfully decode the source message and it becomes a potential relay. All the potential relays comprise the decoded set $\mathcal{D}$, which forwards the source message to the destination in the second phase.

![Figure 2.1: Two-phase relaying scheme: In the first phase, the source broadcasts its message to all cooperating nodes; in the second phase, the selected relay(s) forward(s) the source information to the destination.](image)

We denote the channel coefficient of the link from the source to relay $i$ as $h_{si}$, and that from relay $i$ to the destination as $h_{id}$. In this work, we assume that all the cooperating nodes are located close to the midpoint between the source and the
destination, and the distance from the source to relay \(i\) and that from relay \(i\) to the destination is approximately the same. This assumption is made to isolate our discussion from stochastic geometry [75], which is beyond the scope of this work. Moreover, it has been shown to be a reasonable assumption in implementing cooperative relaying in practical wireless networks [43]. Nevertheless, a note on the impact of the randomly distributed relays is presented in Section 2.5.5. Rayleigh fading is assumed, and, without loss of generality, the channel coefficients are assumed to be i.i.d. complex Gaussian random variables with zero mean and unit variance (recalling that every link has the same physical distance).

In TBBS, each node in the decoded set \(D\) sets up an individual timer that is inversely proportional to the estimated channel power gain of the link from itself to the destination. For example, the timer \(t_i\) at relay \(i, i \in D\), could be configured as \(t_i = \lambda/|\hat{h}_{id}|^2\), where \(\lambda\) is a constant system parameter and \(\hat{h}_{id}\) is the estimated channel coefficient. Once the timer expires, the node immediately transmits the source message to the destination, and all the competitors back off when they overhear this transmission.

If Dis-STC or \(M\)-group is used, instead of selecting the best potential relay, all the nodes in \(D\) will transmit the source signal simultaneously to the destination. In Dis-STC, each node in \(D\) is assigned one unique column of an \(L\)-column STC matrix to transmit, where \(L\) is the cardinality of \(D\). Obviously, a central controller or full inter-node information exchange is required to obtain (i) \(L\), the size of the decoded set, and (ii) the unique assignment of columns among \(D\); this causes significant overhead (quantitative analysis is presented in Section 2.4.2). In contrast, for \(M\)-group, all the nodes in \(D\) randomly choose one column out of a given \(M\)-column STC matrix; in this case, knowledge of \(L\) and the unique column assignment are not required. By doing so, although the achievable cooperative diversity is reduced, the overhead is minimized.

Clearly, different cooperative relaying schemes have diverse performances and require different amounts of overhead. An investigation of the impact of this overhead is critical in evaluating the true potential of cooperation and in providing guidelines
for choosing the appropriate cooperative relaying method for specific applications.

2.3 Overhead Analysis

We now investigate the overhead that is required to implement cooperative relaying, specifically, the overhead incurred by channel estimation, relay selection, and the coordination of the multiple transmissions.

2.3.1 Channel Estimation Overhead

A training sequence is generally used to estimate the channel. Assuming that the actual channel, \( h \), and the estimated channel, \( \hat{h} \), are jointly ergodic and stationary Gaussian processes, we can write

\[
h = \hat{h} + e
\]

where \( e \) is the estimation error, usually modeled as a complex Gaussian random variable with zero mean and variance \( \sigma_e^2 \). In [80], the authors showed that, assuming a linear minimum mean square error estimator, \( \sigma_e^2 = \frac{1}{1 + \bar{\gamma} N_t} \), where \( N_t \) is the number of training symbols and \( \bar{\gamma} \) is the average SNR for the training sequence. As expected, more training symbols and/or higher SNR reduce the estimation error.

Assuming the same transmit power \( P \) for the training sequence and the information stream, the SNR, including the estimation error, is [79]

\[
\gamma = \frac{P |h|^2}{P \sigma_e^2 + P_N} = \frac{\bar{\gamma} |\hat{h}|^2}{1 + \bar{\gamma} N_t}
\]

where \( P_N \) is the power of the noise and \( \bar{\gamma} = P/P_N \). Obviously, applying more training symbols reduces the data transmission time, and, a degradation in the performance (for example, spectral efficiency) is expected. On the other hand, a larger \( N_t \) implies a higher SNR, which improves the performance. Therefore, as is well-understood, there is a trade-off between the number of training symbols used (the overhead) and the achieved performance. Moreover, the amount of channel estimation overhead also depends on the CSI requirement. As we saw in Section 2.2, different cooperative
relaying schemes require different amounts of CSI, and thus introduce different amounts of channel estimation overhead. This is further described in Section 2.4.

2.3.2 Relay Selection Overhead

In selective cooperative relaying, for example TBBS, a relay selection process is required; while, in All-Select relaying, for example Dis-STC, since all the nodes in the decoded set are recruited to forward the source message, the selection is by default, and does not introduce any overhead. Among all the proposed selection schemes, timer-based selection [67] is a fully distributed algorithm, and, potentially, incurs the minimum overhead.

Ideally, the timer-based relay selection that is described in Section 2.2 can always be successful; however, in practice, the selection may fail due to collisions, as discussed in [67, 72, 73]. For example, suppose there are two potential relays in the decoded set, $R_1$ and $R_2$. Let $R_1$ be the better of the two based on the channel power gain to the destination, so that it has a shorter timer $t_1$. As shown in Fig. 2.2, we denote the end-to-end delay from the source to $R_i$ as $t_{di}$, $i \in \{1, 2\}$, the processing time at $R_1$ to switch from receive mode to transmit mode and prepare the packet for transmission as $t_{proc}$ and the propagation time of a signal from $R_1$ to $R_2$ as $t_{prop}$. Suppose that $t_2 < t_1 + \delta + t_{proc} + t_{prop}$, where $\delta = t_{d1} - t_{d2}$, so $R_2$ will start transmitting the source information before it realizes that $R_1$ has been already selected; a collision will result and the transmission from the source to the destination will be unsuccessful.

This means that $R_2$ cannot recognize that $R_1$ has been selected until a time $t_1 + t_g = t_1 + \delta + t_{proc} + t_{prop}$ after the beginning of its own timer $t_2$. Therefore, to avoid collisions and achieve a successful selection, the second smallest timer must be greater than the summation of the smallest timer and a guard time interval $t_g$ to ensure that the timers at all the other nodes do not expire before overhearing the transmission from the best node. Otherwise, more than one node will be selected and a collision will happen. The value of the required $t_g$ depends on the capabilities of the system [73].
Figure 2.2: Suppose the source broadcasts the signal at time 0. Due to different distances to the source, $R_1$ receives the source message at $t_{d1}$ while $R_2$ receives it at $t_{d2}$. Since $R_1$ is the “better” relay, it has the “shorter” timer that expires at $t_{d1} + t_1$. Then, $R_1$ begins preparing the to-be-transmitted message, which takes $t_{proc}$ to complete, and then rebroadcasts the source information. After another $t_{prop}$, $R_2$ receives the signal broadcasted by $R_1$ and backs off.

Note that no capture effect is taken into account in timer-based selection, that is, if more than one relay rebroadcasts the signal simultaneously without coordination, the destination cannot decode the message correctly. But, in practice, it is possible that the destination can still succeed in decoding the message if one of the received versions of the message is much stronger than the others. However, with an appropriate timer configuration, the transmissions that collide usually have comparable signal strengths at the destination because they have similar timers, i.e., similar gains. In this case, in order to avoid the requirement of tight frequency and time synchronization, we assume that the transmission will fail if more than one relay is selected.

Suppose that there are $L$ nodes in the decoded set, that is, $|\mathcal{D}| = L$, where $| \cdot |$
denotes the cardinality of a set, and the ordered sequence, from the smallest to the largest, of the i.i.d. timers is \( t(1) \leq t(2) \leq \cdots \leq t(L) \). Given \( L, 1 \leq L \leq N \), as shown in \([67]\), the collision probability is

\[
p_{\text{coll}}(L) = Pr\left(t(2) < t(1) + t_g\right) = 1 - L(L - 1) \int_{t_g}^{+\infty} \frac{f(x)F(x - t_g)}{(1 - F(x))^{2-L}} \, dx \tag{2.3}
\]

where \( f(x) \) is the probability density function of \( t_i \) and \( F(x) \) is the corresponding cumulative distribution function.

Since \( t_i = \lambda/|\hat{h}_{id}|^2 \) and \( \lambda \) is a constant, the distribution of \( t_i \) is determined by the distribution of \( |\hat{h}_{id}|^2 \). In \([67,72]\), the channel estimation is assumed to be perfect and no channel estimation overhead is considered. Here, taking into account the channel estimation error, we have \( \hat{h}_{id} = h_{id} - e_{id} \sim \mathcal{CN}(0, 1 - \frac{1}{1+\gamma N_t}) \), where \( \mathcal{CN}(\cdot) \) stands for the complex Gaussian probability distribution. This implies that the estimated channel power gain of the link from the potential relay \( i \) to the destination, \( |\hat{h}_{id}|^2 \sim \text{Exp}(\beta) \), where \( \text{Exp}(\cdot) \) denotes an exponential probability distribution and \( \beta = \frac{1+\gamma N_t}{\gamma N_t} \). Therefore, we have

\[
F(x) = Pr(t_i \leq x) = Pr\left(\frac{\lambda}{|\hat{h}_{id}|^2} \leq x\right) = \begin{cases} 
Pr\left(|\hat{h}_{id}|^2 \geq \frac{\lambda}{x}\right) = e^{-\frac{\beta \lambda}{x}} & x > 0 \\
0 & x \leq 0
\end{cases} \tag{2.4}
\]

and

\[
f(x) = \frac{\partial F(x)}{\partial x} = \begin{cases} 
\beta \lambda x^{-2}e^{-\frac{\beta \lambda}{x}} & x > 0 \\
0 & x \leq 0
\end{cases} \tag{2.5}
\]

Using (2.4) and (2.5) in (2.3), we get

\[
p_{\text{coll}}(L) = 1 - L(L - 1)\beta \int_{t_g}^{+\infty} (\lambda x^{-2}e^{-\frac{\beta \lambda}{x}}) e^{-\frac{\beta \lambda}{x-t_g}} \left[1 - e^{-\frac{\beta \lambda}{2}}\right]^{L-2} \, dx \tag{2.6}
\]

In addition to the possible collisions, a duration \( T_s = t(1) \) is consumed, during which there is no data transmission. Given \( L \), the cumulative distribution function of
The smallest timer, is

\[ Pr(t_{(1)} \leq x) = 1 - Pr(t_{(1)} > x) = 1 - \prod_{i \in \mathcal{D}} Pr(t_i > x) = 1 - [1 - F(x)]^L \]

\[ = \begin{cases} 
1 - (1 - e^{-\frac{\beta \lambda}{x}})^L & x > 0 \\
0 & x \leq 0 
\end{cases} \tag{2.7} \]

and, the corresponding probability density function is

\[ f_{(1)}(x) = \begin{cases} 
L(1 - e^{-\frac{\alpha}{x}})^{L-1}e^{-\frac{\alpha}{x} \beta \lambda x^{-2}} & x > 0 \\
0 & x \leq 0 
\end{cases} \tag{2.8} \]

Therefore, the expected selection time \( T_s \), with imperfect CSI, is

\[ E(T_s) = E(t_{(1)}) = \int_0^{+\infty} xf_{(1)}(x)dx = L\beta \lambda \int_0^{+\infty} \frac{e^{-\frac{\alpha}{x}}}{x} (1 - e^{-\frac{\alpha}{x}})^{L-1}dx \tag{2.9} \]

where \( E(\cdot) \) represents expectation.

As noted in [67], the topology of connectivity also has an effect on the selection process. Up to this point, we have assumed that the inter-node communication is perfect in the sense that every potential relay can hear every other node’s transmission. However, in realistic networks, it is possible that some potential relays cannot hear the broadcast from the best relay and will start transmitting the source information once their own timers expire; this will introduce more collisions. One possible approach to overcome this is for the best relay to first send a notification to the destination, and then the destination can broadcast the selection decision. Note that, with the assumption of reciprocal channels, the nodes that cannot hear the selection decision broadcasted by the destination are also not useful for relaying the source information. For this approach, more overhead is required, in particular, a larger timer selection period is required because, in addition to \( t_{(1)} \), another waiting period is necessary to allow the nodes to receive the selection decision from the destination.

### 2.3.3 Transmission Coordination Overhead

If selective relaying such as TBBS is used, the transmission from the relay to the destination is the same as traditional point-to-point communications and does
not need extra coordination; however, if more than one node is utilized in forwarding the message to the destination, as in Dis-STC or $M$-group, some coordination of the multiple transmissions is necessary.

There are several methods for coordinating the simultaneous transmissions from multiple transmitters; among these, STC is attractive because it provides full spatial diversity and only requires CSI at the destination. Both Dis-STC and $M$-group apply STC to coordinate the transmissions from the relays to the destination. In Dis-STC, to obtain the required CSI at the destination, upon decoding the source message, relay $i, i \in \mathcal{D}$, transmits a training sequence to the destination, and the destination estimates the channel from relay $i$ to itself. In order to prevent collisions that might happen when more than one relay transmits the training sequence at the same time, a scheduling scheme is required, e.g., ALOHA-like random access. Alternatively, the relays in $\mathcal{D}$ could send the training sequences in an orthogonal manner, e.g., TDMA with a preassigned order. The detailed implementation is beyond the scope of this work; here, we assume that the time consumed in acquiring the essential CSI equals the product of the duration of one training sequence and the number of relay-to-destination links. In this case, assuming $T_i$ is the time duration of one training sequence, $LT_i$ is consumed to provide the CSI at the destination in Dis-STC, and $MT_i$ is consumed in $M$-group.

In addition to the overhead incurred in obtaining the CSI, the column assignment in Dis-STC also consumes a portion of the transmission time. Since a unique assignment is required, upon reception of the training sequences from all the relays in the decoded set, the destination assigns a unique column of an $L$-column STC matrix to each relay in $\mathcal{D}$, which can be done by broadcasting the assignment; this is assumed to take another $T_i$ because both the column assignment and training sequence are relatively short control signals. Note that $M$-group utilizes a random column assignment, that is, each potential relay chooses its own column individually and randomly, and thus does not require any overhead for this purpose.
In summary, we have addressed three key components of the overhead in implementing cooperative relaying: channel estimation, relay selection, and transmission coordination. Note that the analysis of the channel estimation overhead is generic and can be applied to other cooperative relaying algorithms as long as pilot symbols are used for channel estimation; but the relay selection overhead is specific to timer-based selection and the coordination overhead may only apply to STC-based cooperation. The impact of the overhead for these three tasks on the performance of cooperative relaying is complicated because the tasks are coupled. For example, the performance and the amount of overhead for relay selection depend on the channel estimation performance since the timer is inversely proportional to the estimated channel power gain. We have also shown that different cooperative relaying schemes require significantly different amounts of overhead. For example, the relay selection overhead plays a major role in determining the total overhead in implementing TBBS; while the coordination overhead is the main component of the overhead in Dis-STC. In the following, the performance-overhead trade-offs for different cooperative relaying schemes are studied to answer the two questions: (i) What is the true performance including overhead? and (ii) Which among TBBS, Dis-STC, and M-group is the most efficient relaying scheme?

2.4 Spectral Efficiency

To analyze the performance-overhead trade-offs and compare the cooperative relaying techniques, in what follows, we focus on the spectral efficiency, which is defined as the successfully delivered bits from the source to the destination per second per Hz,

$$\eta_s = \frac{1}{2} \frac{r}{B} T_e p_{suc}$$

(2.10)

where $r$ is the bit rate of the data transmission, $T_e$ is the effective data transmission time, $B$ is the bandwidth, $T$ is the total transmission time, $p_{suc}$ is the probability that the destination succeeds in receiving and decoding the signal initiated from the source, and the factor of $1/2$ comes from the two-phase transmission.
It is important to note that the three types of overheads that are addressed in Section 2.3 play different roles in determining the true spectral efficiency. Obviously, applying longer training sequences (more training symbols) reduces $T_e$, degrading the spectral efficiency. On the other hand, more training symbols provides a higher SNR, which has the potential to improve the spectral efficiency. So, there is a trade-off between the channel estimation overhead and the spectral efficiency. For relay selection, $T_e$ is reduced (selection time consumes a portion of the transmission time) and collisions are introduced that decrease the chance of successful transmission. Similarly, coordinating multiple transmissions consumes a portion of the transmission time and reduces $T_e$. However, relay selection in TBBS and transmission coordination in Dis-STC are essential to achieve the diversity benefit, which, in turn, improves the spectral efficiency. The objective of this section is to quantify these various performance-overhead trade-offs.

TBBS, Dis-STC, and $M$-group share the same first phase, in which the source broadcasts a signal, and all $N$ cooperating nodes listen. Assume that $L$ out of $N$ nodes have a higher SNR than the given threshold $\gamma_{th}$, that is, these $L$ nodes can successfully decode the source message, and comprise the decoded set $D$. Note that a training sequence is needed to provide the channel information $h_{si}$ to decode the source message at node $i$. This overhead, for one training sequence, is identical for all three cooperative relaying schemes, and is included in the following analysis by subtracting $T_t$, the duration of one training sequence, from the total transmission time $T$.

For the link from the source to node $i$, the SNR $\gamma_{si}$ is given in (2.2). Then, the outage is

$$p_{out}^{si} = Pr\{\gamma_{si} < \gamma_{th}\} = Pr\left\{ |\hat{h}_{si}|^2 < \frac{\gamma_{th}}{\gamma} \frac{1 + \bar{\gamma}(1 + N_t)}{1 + \bar{\gamma}N_t} \right\}$$

(2.11)

Note that $|\hat{h}_{si}|^2 \sim Exp(\beta)$, where $\beta = \frac{1 + \bar{\gamma}N_t}{\bar{\gamma}N_t}$, therefore,

$$p_{out}^{si} = 1 - \exp \left[ -\frac{\gamma_{th}}{\bar{\gamma}} \frac{1 + \bar{\gamma}(1 + N_t)}{\bar{\gamma}N_t} \right]$$

(2.12)

For large $N_t$, the outage approaches $1 - e^{-\frac{\gamma_{th}}{\bar{\gamma}}}$, the result with perfect channel estimation.
Since the $h_{si}$’s are i.i.d., $p_{si}^{out}$ is the same for any $i$, and the probability that the cardinality of the decoded set equals $L$ is (letting $p_{si}^{out} = p$ for notational simplicity)
\[
Pr\{|\mathbf{D}| = L\} = \binom{N}{L} (1 - p)^L p^{N-L} \tag{2.13}
\]

### 2.4.1 Timer-Based Best-Select Relaying (TBBS)

In the second phase of TBBS, the best relay in $\mathcal{D}$ is selected, in the previously described timer-based manner, to forward the source message. In order to set up the timer, $t_i = \lambda/|\hat{h}_{id}|^2$, the estimated channel power gain is required at the potential relay $i, i \in \mathcal{D}$ ($\lambda$ is a pre-assigned system parameter); the gain is obtained by broadcasting a training sequence from the destination [67]. Once the best relay is selected and forwards the source message to the destination, another training sequence is included in the packet to provide the knowledge of the channel from the selected relay to the destination. In total, counting the training sequence applied in the first phase, there are three training sequences used in TBBS. Recall that timer-based best selection requires a selection time $T_s$. Therefore, the effective data transmission time is $T - 3T_t - T_s$.

Clearly, the source-to-destination transmission in TBBS is successful as long as (i) there is no outage in the first phase (the decoded set is not empty), (ii) the selection is successful (no collision), and (iii) the SNR at the destination is higher than the threshold. If there is no collision, the relay in $\mathcal{D}$ with the highest gain to the destination is selected to forward the source message. So, the outage in the second phase for a given decoded set is $p^L$, where $p$ is the outage probability for the link from one relay to the destination and is the same as $p_{si}^{out}$ in (2.11). Therefore, the probability of successful communication between the source and the destination is
\[
p_{suc} = \sum_{L=1}^{N} Pr\{|\mathbf{D}| = L\} [1 - p_{coll}(L)] (1 - p^L) \tag{2.14}
\]
where $p_{coll}(L)$, the collision probability for a given $L$, is given in (2.6). Then, the
spectral efficiency of TBBS can be written as (with $\delta = \frac{T_s}{T}$)

$$\eta_{BS} = \sum_{L=1}^{N} Pr\{|D| = L\} \frac{1}{2B} \frac{r}{T} \frac{T - 3T_i - T_s}{T} [1 - p_{coll}(L)] (1 - p^L)$$

$$= \frac{r}{2B} \sum_{L=1}^{N} \binom{N}{L} (1 - p)^L p^{N-L} (1 - p^L) \times (1 - 3\delta - \frac{T_s}{T}) [1 - p_{coll}(L)] \quad (2.15)$$

where $T_s = t_{(1)}$, whose expectation is given in (3.2).

### 2.4.2 Dis-STC Cooperative Relaying (Dis-STC)

As discussed in Section 2.3.3, $LT_i$ is consumed in the second phase of Dis-STC for the destination to obtain the CSI of all the channels from each of the $L$ potential relays to the destination, and $T_t$ is consumed in broadcasting the column assignment from the destination. Therefore, including the $T_t$ consumed in the first phase, the effective data transmission time becomes $T_e = T - (L + 2)T_t$. It is important to note that the effective data transmission time decreases linearly with the size of the decoded set; in the simulations in Section 2.5, this is shown to be one major drawback of implementing Dis-STC in practical systems.

If the conventional coherent detection algorithm for STC is used, the SNR at the destination is the sum of the SNR’s of all the links. Hence, the outage in the second phase is

$$p_{ST,II} = Pr\left\{ \sum_{i \in D} \gamma_{id} < \gamma_{th} \right\} = Pr\left\{ \sum_{i \in D} |\widehat{h}_{id}|^2 < \frac{\gamma_{th} \frac{1 + \gamma (1 + N_t)}{\bar{\gamma}}}{1 + \gamma N_t} \right\} \quad (2.16)$$

To have a fair comparison, a total power constraint is applied, and, to minimize the control overhead, uniform power allocation is assumed for those relays that forward the source message. Note that every relay that forwards the source message has the knowledge of the size of $D$ (the destination can insert $L$ in the column assignment message) and can easily set its transmit power as $P/L$. This means that the $|\widehat{h}_{id}|^2$'s
are i.i.d. exponentially distributed with mean $\beta/L$, that is, $|\hat{h}_{id}|^2 \sim \text{Exp}(L\beta)$. Then, 

$$\sum_{i \in D} |\hat{h}_{id}|^2 \sim \text{Erlang}(L, L\beta)$$

and the outage becomes

$$p_{ST,II}(L) = 1 - \sum_{j=0}^{L-1} \exp \left\{-\frac{\gamma_{th} L [1 + \gamma (1 + N_t)]}{\gamma N_t} \right\} \frac{\left(\frac{\gamma_{th} L [1 + \gamma (1 + N_t)]}{\gamma N_t}\right)^j}{j!}$$

(2.17)

Since the selection is by default and there is no collision, the probability of successful transmission is

$$p_{suc} = \sum_{L=1}^{N} Pr\{|D| = L\} [1 - p_{ST,II}(L)]$$

(2.18)

Therefore, the spectral efficiency of Dis-STC becomes

$$\eta_{ST} = \sum_{L=1}^{N} Pr\{|D| = L\} \frac{r_c}{2B} \frac{T - (L + 2)T_t}{T} [1 - p_{ST,II}(L)]$$

$$= \frac{r}{2B} \sum_{L=1}^{N} \binom{N}{L} (1-p)^L p^{N-L} \frac{[L/2] + 1}{2[L/2]} [1 - (L + 2)\delta]$$

$$\times \sum_{j=0}^{L-1} \exp \left\{-\frac{\gamma_{th} L [1 + \gamma (1 + N_t)]}{\gamma N_t} \right\} \frac{\left(\frac{\gamma_{th} L [1 + \gamma (1 + N_t)]}{\gamma N_t}\right)^j}{j!}$$

(2.19)

where $r_c = \frac{[L/2] + 1}{2[L/2]}$ is the maximum code rate for orthogonal STC [82].

### 2.4.3 M-Group Dis-STC Cooperative Relaying (M-group)

To reduce the overhead incurred in Dis-STC, M-group utilizes an underlying M-column STC matrix, which does not change with the decoded set, and lets every potential relay randomly choose one column to transmit. So, the decoded set is divided into M groups, each of which transmits a unique column out of the M possible columns (note that empty groups might exist). Although the achievable diversity order is upper-bounded by M (achieved when there are no empty groups), which is smaller than the diversity order ($N$) of Dis-STC, neither central control nor full inter-node communication is required; hence, much less overhead is needed. In addition, since the potential relays in each group comprise a virtual channel to the destination, only $M$
training sequences are needed for channel estimation. In this case, the effective data
transmission time is \( T_e = T - (M + 1)T_t; \) \( T_t \) is consumed in the first phase and \( MT_t \)
in the second phase.

Clearly, the outage performance depends on the grouping result. We denote all
the possible groupings as a set \( G \), where each element \( g_j, j \in \{1, 2, \ldots, |G|\} \) is an \( M \)
dimensional vector indicating the numbers of the relays that choose different columns.
For example, assuming \( M = 3 \) and \( N = 6 \), \( g_j = [3 \ 2 \ 1] \) means that 3 nodes out of the 6
relays choose column 1 to transmit, 2 nodes choose column 2, and 1 node chooses the
last column. Given a grouping, say \( g_j \), by applying the conventional coherent detection
algorithm for STC, the outage is

\[
p_{M,II}(g_j) = Pr \left\{ \sum_{m=1}^{M} |\hat{h}'_m|^2 < \frac{\gamma_{th} 1 + \gamma(1 + N_t)}{\bar{\gamma} 1 + \gamma N_t} \right\} \tag{2.20}
\]

where \( |\hat{h}'_m|^2 \) is the power gain of the virtual channel from the relays that use column
\( m \) to transmit to the destination. Due to the independence among the \( \hat{h}_{ij} \)’s, \( |\hat{h}'_m|^2 \) is
distributed as \( Exp\left(\frac{L\beta}{g_j[m]}\right) \), where \( g_j[m] \) is the \( m \)th element of \( g_j \), i.e., the number of the
relays that choose column \( m \). Denoting the probability of one specific grouping result
as \( Pr(g_j) \), we have

\[
p_{M,II} = \sum_{g} Pr(g_j) Pr \left\{ \sum_{m=1}^{M} |\hat{h}'_m|^2 < \frac{\gamma_{th} 1 + \gamma(1 + N_t)}{\bar{\gamma} 1 + \gamma N_t} \right\} \tag{2.21}
\]

Because a rate-1 Alamouti Code [10] can be used with \( M = 2 \), and because
the required overhead is minimal, in most cases, 2-group (\( M = 2 \)) provides better
performance than \( M \)-group with \( M \geq 3 \). Therefore, in the following, we focus on the
case where \( M = 2 \). When \( L = 1 \), that is, there is one node in \( D \), the outage in the
second phase is \( p \). When \( L \geq 2 \), suppose \( k, 1 \leq k \leq L - 1 \), relays choose the first
column to transmit and the remaining \( (L - k) \) relays use the second column; then, we
have \( |\hat{h}'_1|^2 \sim Exp\left(\frac{L\beta}{k}\right) \) and \( |\hat{h}'_2|^2 \sim Exp\left(\frac{L\beta}{L-k}\right) \). Note that uniform power allocation is
again assumed to present a fair comparison. The practical implementation of uniform
power allocation in \( M \)-group is not discussed here. In the extreme cases where \( k = 0 \)
or \( k = L \), one group is empty while the other has all \( L \) components, that is, one of \( \{|\hat{\mathbf{h}}_1|^2, |\hat{\mathbf{h}}_2|^2\} \) is 0 and the other one has an exponential distribution, \( \text{Exp}(\beta) \). Similar to (2.20), we have

\[
p_{M,II} = Pr\left\{ |\hat{\mathbf{h}}_1|^2 + |\hat{\mathbf{h}}_2|^2 \leq \frac{\gamma th}{\bar{\gamma}} \left( \frac{1 + \bar{\gamma}(1 + N_t)}{1 + \bar{\gamma}N_t} \right) \right\} \tag{2.22}
\]

If \( k = 0 \) or \( k = L \) (\( L \geq 2 \)), then \( |\hat{\mathbf{h}}_1|^2(|\hat{\mathbf{h}}_2|^2) = 0 \) and \( |\hat{\mathbf{h}}_2|^2(|\hat{\mathbf{h}}_1|^2) \sim \text{Exp}(\beta) \). Therefore,

\[
p_{M,II} = p \tag{2.23}
\]

If \( 1 \leq k \leq L - 1 \) (\( L \geq 2 \)), then \( |\hat{\mathbf{h}}_1|^2 \sim \text{Exp}(\frac{L\beta}{k}) \) and \( |\hat{\mathbf{h}}_2|^2 \sim \text{Exp}(\frac{L\beta}{L-k}) \). Therefore,

\[
p_{M,II} = \int_{x_1 + x_2 < \frac{\gamma th}{\bar{\gamma}} \frac{1 + \bar{\gamma}(1 + N_t)}{1 + \bar{\gamma}N_t}} f(x_1, x_2) dx_1 dx_2 \tag{2.24}
\]

where \( f(x_1, x_2) = \frac{L\beta}{k} \exp\left(-\frac{L\beta x_1}{k}\right) \frac{L\beta}{L-k} \exp\left(-\frac{L\beta x_2}{L-k}\right) \). Combining (2.23) and (2.24),

\[
p_{M,II}(L) = \sum_{k=1}^{L-1} \frac{L!}{(L-k)! k!} \int_{x_1 + x_2 < \frac{\gamma th}{\bar{\gamma}} \frac{1 + \bar{\gamma}(1 + N_t)}{1 + \bar{\gamma}N_t}} f(x_1, x_2) dx_1 dx_2 + \frac{p}{2L-1} \tag{2.25}
\]

To make the comparison clearer and provide more insight, we assume uniform grouping (the effect of random grouping is illustrated in the simulations in Section 2.5), that is, if \( L \) is even, each group has the same number of relays; otherwise, one group has \( \lfloor L/2 \rfloor \) nodes and the other \( \lceil L/2 \rceil \), and the uniform power allocation is slightly modified so that the total power is equally allocated between the two groups. In this case, \( |\hat{\mathbf{h}}_1|^2 + |\hat{\mathbf{h}}_2|^2 \sim \text{Erlang}(2,2\beta) \), and hence, substituting \( L = 2 \) into (2.17), the outage is

\[
p_{M,II}(L) = 1 - \exp\left\{ -\frac{\gamma th}{\bar{\gamma}} \frac{2[1 + \bar{\gamma}(1 + N_t)]}{\bar{\gamma}N_t} \right\} \left\{ 1 + \frac{\gamma th}{\bar{\gamma}} \frac{2[1 + \bar{\gamma}(1 + N_t)]}{\bar{\gamma}N_t} \right\} \tag{2.26}
\]

Similar to Dis-STC, the source-destination transmission is successful as long as (i) the decoded set is not empty and (ii) there is no outage in the second phase. So, we have

\[
p_{\text{suc}} = \sum_{L=1}^{N} Pr\{|\mathcal{D}| = L\} \left[ 1 - p_{M,II}(L) \right] \tag{2.27}
\]
Therefore, the spectral efficiency of $2$-group ($M = 2$) with uniform grouping is

$$\eta_M = \sum_{L=1}^{N} Pr\{|D| = L\} \frac{1}{2B} \frac{r}{T} \frac{T - 3T_t}{T} [1 - p_{M,II}(L)]$$

$$= \frac{r}{2B} \sum_{L=1}^{N} \binom{N}{L} (1 - p)^L p^{N-L} (1 - 3\delta) \exp \left\{ -\frac{\gamma_{th} 2 [1 + \bar{\gamma} (1 + N_t)]}{\bar{\gamma} N_t} \right\}$$

$$ \times \left\{ 1 + \frac{\gamma_{th}}{\bar{\gamma}} \frac{2 [1 + \bar{\gamma} (1 + N_t)]}{\bar{\gamma} N_t} \right\}$$

(2.28)

Among these three schemes, TBBS and Dis-STC provide full diversity in the number of cooperating nodes, while the diversity order of $M$-group is upper bounded by $M$. On the other hand, the efficiency of TBBS suffers selection time consumption and possible collisions; the coordination overhead dramatically degrades the efficiency of Dis-STC. These issues together make the comparison complicated. In the case of perfect CSI without channel estimation overhead, it has been shown in [72] that $2$-group is generally more spectrally efficient than TBBS.

### 2.5 Simulation Results

In this section, we present numerical results illustrating the impact of overhead on the spectral efficiency, and compare the performance for the three cooperative relaying schemes described above. In addition to TBBS, Dis-STC, and $2$-group, the ideal TBBS without relay selection overhead ($p_{col} = 0, T_s = 0$) and $2$-group with uniform grouping (UG) are also simulated to more clearly reveal the impact of overhead.

We consider a wireless system with moderate mobility such that the coherence time is on the order of 10-100 msec [83], and the end-to-end transmission duration is set as $T = 1$ msec. The symbol length is assumed to be 0.01 msec, which implies that $\delta = T_t / T = N_t / 100$. Clearly, if the symbol length is smaller and/or the transmission duration is larger, given the same $N_t$, $\delta$ becomes smaller, which means, as expected, the channel estimation takes a smaller portion of total transmission time. So, the influence of the channel estimation overhead is highly dependent on the number of transmitted symbols in one end-to-end transmission (100 is assumed here); the more symbols there are in one transmission, the smaller the impact of the channel estimation overhead.
Moreover, since the absolute value of the spectral efficiency is not the focus here, we set \( r/B \) to 1 bits/s/Hz, and hence, the highest achievable spectral efficiency is 0.5 bits/s/Hz due to the factor of 1/2 from the two phases. To simplify the notation, we term the ratio of the average SNR, \( \bar{\gamma} \), to the fixed SNR threshold, \( \gamma_{th} \), as the SNR margin \( \theta = \bar{\gamma}/\gamma_{th} \).

### 2.5.1 Impact of Relay Selection Overhead

As discussed in [72], we can set \( t_g = \Delta_t + t_{prop} + t_{proc} \). \( \Delta_t \) results from the different propagation distances from the source to the cooperating nodes and \( t_{prop} \) is the duration for the broadcasted message from the best relay to propagate among the decoded set. Generally, all the potential relays are close to each other; so, \( \Delta_t \) and \( t_{prop} \) are negligible (much smaller than 1 \( \mu \)sec), compared with \( t_{proc} \), the processing delay. Note that, in realistic networks, all signals are transmitted as data packets and the time consumption to prepare the packets, i.e., \( t_{proc} \), must be clearly addressed. As a lower bound, \( t_{proc} = t_{sw} + t_p + t_h \), where the first term is the switching time from receive mode to transmit mode and the last two account for a packet’s preamble and header. We assume \( t_{sw} < 2 \mu \)sec, \( t_p = 32 \mu \)sec, and \( t_h = 8 \mu \)sec; these values are consistent with the specifications defined in the IEEE 802.11 wireless LAN standard [18]. Therefore, the typical guard interval \( t_g \) should be on the order of 10 – 100 \( \mu \)sec; here, we set \( t_g = 50 \mu \)sec.

To obtain general insight without restriction on the specific value of \( t_g \), the impact of \( \mu = \lambda/t_g \), recalling that \( \lambda \) is the constant parameter in the timer setting, on the collision probability is plotted in Fig. 2.3. As expected, it shows that, given \( L \) (the size of the decoded set) and \( N_t \) (the number of training symbols used), \( p_{coll} \) is decreasing with respect to \( \mu \), which increases the spectral efficiency. However, the expected selection time increases with \( \mu \), as shown in Fig. 2.4, which degrades the spectral efficiency. Therefore, there is an important trade-off in the choice of \( \mu \). Additionally, with increasing \( L \), the probability of collision goes up, while the expected selection time goes down. The reason is that the contention is more intense with a
larger decoded set, while, at the same time, the expected value of the minimum timer becomes smaller.

![Collision Probability Graph](image)

**Figure 2.3:** Collision probability as a function of $\mu$ with different sizes of the decoded set (solid curves with $L = 5$ and dashed curves with $L = 10$) and different numbers of training symbols (different markers that are corresponding to $N_t = 1, 5, 10$).

In addition, the impact of the channel estimation overhead on the relay selection performance is also illustrated in Figs. 2.3 and 2.4. Recall that $t_i = \lambda / |\hat{h}_{id}|^2$ and $|\hat{h}_{id}|^2 \sim Exp(\beta)$, where $\beta = \frac{1+\gamma N_t}{\gamma N_t}$. Therefore, given $L$ and $\mu$, the selection time consumption $T = t_{(1)}$, which is the smallest among all the timers, decreases when $N_t$ increases because of the increased mean of $|\hat{h}_{id}|^2$. In contrast, the collision probability increases with $N_t$. Because the estimated channel power gains are higher, for the same $\lambda$, the difference between the smallest two timers becomes smaller, which implies a higher probability that $t_{(2)} < t_{(1)} + t_g$. For the cases of interest ($p_{coll}$ is low, say,
the impact of channel estimation on the collision probability is negligible, while the expected selection time decreases with $N_t$. This means that, as expected, the more training symbols that are used, the better the relay selection performance.

Figure 2.4: Expected selection time as a function of $\mu$ with different sizes of the decoded set (solid curves with $L = 5$ and dashed curves with $L = 10$) and different numbers of training symbols (different markers that are corresponding to $N_t = 1, 5, 10$).

As discussed in Section 2.3.2, the selection process in TBBS introduces selection time consumption and possible collisions that degrade the spectral efficiency. It has been shown, in Figs. 2.3 and 2.4, that $\mu$ is the main determining factor for the collision probability and the expected selection time. Therefore, in Fig. 2.5, we present the effect of $\mu$ on the spectral efficiency of TBBS. Since the focus here is the impact of relay selection overhead, perfect CSI without channel estimation overhead is assumed; the simulation results including channel estimation overhead are provided in Section
2.5.2. As expected, the performance with an adaptive $\mu$ is the best. When the number of nodes in the network is small, that is, $N$ is small, the selection time consumption plays a more important role than the collision probability in determining the efficiency; so a small $\mu$ is preferred. In contrast, in a large network, the main factor is the collision probability, which means a large $\mu$ provides a higher efficiency. This phenomenon is a consequence of the fact that $p_{\text{coll}}$ increases with respect to $N$ and decreases with $\mu$, while $T_s$ decreases with respect to $N$ and increases with $\mu$. Note that an appropriate “fixed $\mu$” might provide acceptable performance (for example, $\mu = 100$ in Fig. 2.5); however, the optimal or acceptable value is highly dependent on the system parameters (such as $T$ and $\theta$), and is difficult to determine in a distributed manner.

![Figure 2.5: Spectral efficiency of TBBS, including relay selection overhead, as a function of $N$ with different values of $\mu$](image)

Figure 2.5: Spectral efficiency of TBBS, including relay selection overhead, as a function of $N$ with different values of $\mu$
2.5.2 Impact of Channel Estimation Overhead

To illustrate the impact of the channel estimation overhead, we show, in Figs. 2.6 and 2.7, the spectral efficiency as a function of the number of symbols used in one training sequence. On the one hand, the more training symbols that are used, the more accurate the CSI, and thus, the smaller the outage, which, in turn, improves the efficiency. On the other hand, using more training symbols implies more overhead, and a degradation in efficiency. Therefore, there exists an optimal number of training symbols.

It has been shown, in Section 2.3.3, that the number of required training sequences in Dis-STC is $L + 2$, while the other two schemes require just 2 or 3 sequences. So, the efficiency of Dis-STC decreases fastest as $N_t$ (the number of symbols used in one sequence) increases. Another observation is that the optimal number of training symbols decreases when the SNR margin, $\theta$, increases. Recall that the channel estimation error has variance $(1 + \bar{\gamma} N_t)^{-1}$, hence, when $\theta = \bar{\gamma}/\gamma_{th}$ increases, a smaller $N_t$ is required for the same estimation performance. Therefore, for cases with moderate-to-high SNR margins, say, $\theta \geq 3 \text{ dB}$, a relatively small number of training symbols ($N_t \leq 4$) provides the best performance for all schemes. This means that, when the averaged SNR is larger than twice the SNR threshold, the channel estimation overhead is the major factor in determining the spectral efficiency, compared with the estimation error.

2.5.3 Impact of Coordination Overhead

In Dis-STC and $M$-group, coordination is essential for enabling the simultaneous multiple transmissions. As shown in Section 2.3.3, the associated overhead can be represented by the training sequences required: $(L + 1)T_t$ for Dis-STC and $2T_t$ for 2-group. Clearly, the impact of the coordination overhead in Dis-STC increases when there are more nodes in the network, which usually suggests a larger decoded set.
2.5.4 Performance Comparison

Including all three types of overheads, namely, relay selection, channel estimation, and coordination, the spectral efficiencies for different SNR margins, given the number of training symbols \( N_t = 2 \) and the number of cooperating nodes \( N = 5 \), are shown in Fig. 2.8. At first, we notice that ideal TBBS is always the best, since it can achieve full diversity in the number of cooperating nodes without the large amount of overhead required for Dis-STC. When the SNR margin is small, TBBS performs better than 2-group since the determining factor is the diversity order (outage is high when the SNR margin is small, and the diversity can reduce the outage dramatically), and the selection overhead is small (small SNR margin means a small decoded set, and TBBS can be accomplished with a small amount of overhead). On the other hand,
Figure 2.7: Impact of channel estimation overhead: spectral efficiency as a function of the number of symbols used in one training sequence ($\theta = 5$ dB, $N = 5$)

2-group is more efficient when the SNR margin is large ($\theta \geq 7$ dB in Fig. 2.8). The reason is that when the margin is large, the determining factor becomes the selection overhead. Since the outage is fairly small and the decoded set is large, the advantage of the higher diversity order for TBBS is vanishing, and a large amount of overhead is required to select the best relay.

The impact of grouping on the performance of 2-group is also shown in Fig. 2.8. With random grouping, it is possible that all the potential relays transmit the same column, and then the diversity order is only 1. This is why random grouping is worse than uniform grouping. When $\theta$ is large, which implies a large decoded set, the probability of an empty group is quite small, and hence the performance of 2-group is almost the same as that of 2-group with uniform grouping. Dis-STC has poor
performance, especially when $\theta$ is large (decoded set is large), because (i) applying orthogonal STC limits the code rate and (ii) the number of required training sequences increases linearly with the size of the decoded set.

![Figure 2.8: Spectral efficiency as a function of $\theta$ ($N_t = 2, N = 5$)](image)

The comparison among the cooperative relaying schemes with different numbers of cooperating nodes is illustrated in Fig. 2.9. Similar to the results in Fig. 2.8, the performance of Dis-STC drops rapidly when the number of cooperating nodes increases because of the increasing amount of coordination overhead. Therefore, we then focus on the performance of the two practical distributed cooperative relaying schemes: TBBS and 2-group. Note that, even if there is only one node in the decoded set, the relay selection overhead is still incurred for TBBS, particularly, the selection time consumption. This is the reason why the performance of TBBS is slightly worse.
when the number of cooperating nodes is small ($N = 1$ in Fig. 2.9). When $N$ increases, TBBS provides more benefit in the sense of higher diversity order, but also degrades the efficiency more due to the increased relay selection overhead. Therefore, an optimal number of cooperating nodes for TBBS in the sense of maximum spectral efficiency can be observed ($N = 9$ in Fig. 2.9). In contrast to TBBS, 2-group requires a constant amount of overhead with increasing $N$. This is why 2-group is more efficient than TBBS when $N$ is large ($N \geq 9$ in Fig. 2.9). From Figs. 2.8 and 2.9, we conclude that the size of the decoded set is the major factor in determining whether TBBS or 2-group is more efficient. With the parameters used in this simulation, 2-group outperforms TBBS when $\theta \geq 7$ dB with $N = 5$ or $N \geq 9$ with $\theta = 5$ dB.

![Figure 2.9: Spectral efficiency as a function of $N$ ($N_t = 2, \theta = 5$ dB)](image-url)
2.5.5 A Note on Randomly Distributed Relays

In the previous discussions, we assumed that the cooperating nodes are located close to the midpoint between the source and the destination. Although it has been shown to be a reasonable assumption in practical networks, it is instructive to investigate the validity of our conclusions when the relays are randomly distributed. Keeping the other system settings the same as in the previous simulations, we assume that the source is located at (0, 0) and the destination is at (400, 0). We consider the simple, yet instructive, case where the cooperating nodes are uniformly distributed along the line from the source to the destination, specifically, between (200 − x, 0) and (200 + x, 0). Note that, in general, the relays that are further to the destination than the source are not helpful; so, we assume $x \in [0, 200]$ (the units are meters in all cases).

![Figure 2.10: Spectral efficiency as a function of $N$ with randomly distributed relays ($N_t = 2, \theta = 5 \text{ dB}$)](image-url)
As discussed in Section 2.5.4, TBBS and 2-group are efficient and practical; hence, in the following, we focus on these two schemes. In Fig. 2.10, we show the performance of TBBS and 2-group when $x = 0$ (the relays are fixed and at the midpoint), $x = 100$ (the relays are between $(100, 0)$ and $(300, 0)$), and $x = 200$ (the relays are between $(0, 0)$ and $(400, 0)$). For 2-group, the performance decreases due to the randomness in the locations and the resulting differences in path loss; but the degradation becomes small as the number of nodes increases. For TBBS, however, with randomly distributed relays, the performance increases if there are a sufficient number of nodes. For example, when $x = 200$ and $N \geq 8$, TBBS achieves a slightly better performance than that achieved when $x = 0$. The reason is that more randomness in the channel power gains from the relays to the destination separates the timers more, which reduces the relay selection overhead (since collisions are less likely to happen), and, hence, improves the efficiency. More importantly, focusing on the comparison between TBBS and 2-group, we can see that TBBS improves relative to 2-group when the relays are randomly distributed. Specifically, when $\theta = 5$ dB, with fixed relays, 2-group is better when $N \geq 6$; while, with randomly distributed relays ($x = 100$), 2-group is preferable only if $N \geq 8$, and when $x = 200$, 2-group is better only if $N \geq 14$. More generally, due to the increasing amount of overhead incurred in TBBS, 2-group becomes more efficient when $N$ is large; the only difference with the inclusion of randomly distributed relays is where the crossover occurs.

In summary, as shown and discussed in the analysis and the results, Dis-STC is not practical because of the amount of overhead required in the centralized implementation. The distributed cooperative relaying schemes, that is, TBBS and 2-group, are more attractive. TBBS provides full diversity order but suffers the relay selection overhead, while 2-group is fully distributed but with a diversity order upper-bounded by 2. Therefore, TBBS is better when the SNR margin is small (high diversity order is important to combat the outage) and the network size is small (low relay selection overhead); while 2-group is preferable when the SNR margin is large and/or the network is large.
These results provide a guideline for selecting a suitable cooperative relaying scheme based on the statistical system parameters, namely, the SNR level and the network size. In addition, in practical implementations, determining the best timer scheme for TBBS is still a challenging problem, especially, in a distributed manner; while for 2-group, an efficient distributed power allocation algorithm is critical in closing the gap to 2-group with uniform grouping.

2.6 Conclusions

In this chapter, we investigated the overhead in implementing cooperative relaying, particularly, the overhead to estimate the channels, select the relay(s), and coordinate the transmissions. Taking into account the overhead, the spectral efficiencies of three different relaying schemes, that is, TBBS, Dis-STC, and $M$-group, were studied. Through analyses and simulations, we have addressed the impact of the overhead on spectral efficiency, and provided some guidelines in determining the appropriate relaying scheme for specific applications. In general, Dis-STC is not practical due to the centralized implementation that incurs significant overhead; 2-group is preferable when the network is large and/or the SNR margin is large, while, TBBS is better in small networks.

In addition to the analysis of the spectral efficiency, the energy efficiency is also an issue with a lot of practical interest. Although there has been some related work on energy efficiency of cooperative relaying [53, 84–86], a comprehensive study considering the overhead is still an open area of investigation. In Chapter 3, we extend our overhead-performance trade-off analysis to investigate the energy efficiency of cooperative relaying, taking into account the energy consumed by the transceivers in Transmit, Receive, Idle, and Sleep modes.

Based on the conclusions that the network size is one critical parameter to determine which relaying scheme (TBBS or 2-group) is more efficient, it is natural to derive the optimal size of the decoded set for TBBS. Specifically, if TBBS is chosen as the cooperative relaying technique for a network, e.g., because the network is small
and/or the wireless channels are experiencing rapid fluctuations, we need to determine how many cooperative relays the source should recruit to forward the signal to the destination. The optimal number of cooperating relays is determined by the trade-off between the achieved diversity order and the relay selection overhead. The optimal size of the decoded set can provide important guidelines for efficient implementation of selective relaying. Therefore, in continuing work, we plan on analyzing the impact of the size of the decoded set and deriving the threshold where a switch between TBBS and 2-group is needed to maximize the efficiency.

Furthermore, a more “realistic” comparison can be achieved when interference is included, and using a more holistic analysis of cooperative relaying in wireless networks, taking into account the interplay with higher layers in the protocol stack. For example, the practical implementation of the coordination for Dis-STC is mainly a MAC layer issue, and a “perfect” scheme is instead assumed in the above analysis. Clearly, a holistic analysis is helpful in figuring out the implementation difficulties and providing a comprehensive picture of how the algorithms work in practice. As discussed Chapter 1, incorporating cooperation changes the performance of existing interference management techniques. In Chapter 5, we investigate interference management in cooperative networks, and conduct a holistic analysis and simulation of cooperative networks, focusing on the overhead-performance trade-off and efficient interference management.
Chapter 3

PERFORMANCE-OVERHEAD TRADE-OFF: ENERGY EFFICIENCY

Recently, the energy efficiency of wireless networks has become a growing concern [87–89]. For network operators, especially, cellular operators, improving energy efficiency is an important component of maximizing profitability; for mobile users, improving the energy efficiency is equivalent to prolonging the battery life, a critical feature for future power-hungry applications. Cooperative communications has been shown to be effective in improving the energy efficiency, and hence has attracted more and more academic and industrial research interest [53, 84, 85, 90–92]. As discussed in Chapter 2, however, implementing cooperation requires more overhead.

The performance-overhead trade-off for cooperative relaying has been investigated in Chapter 2, focusing on the spectral efficiency alone. In [84], the authors investigate the impact of the application of cooperative beamforming on the energy efficiency. It is shown that the efficiency may not increase with the number of cooperating nodes due to the overwhelming amount of overhead incurred in implementing cooperation. In this chapter, we present our research work on the overhead-performance trade-off analysis focusing on the energy efficiency of cooperative relaying systems.

3.1 Background and Related Work

Distributed cooperative relaying is an attractive method to combat fading in wireless communications because of its performance advantages, simplicity, scalability, and low overhead. In Chapter 2, we introduced three different distributed relaying schemes, namely, TBBS, Dis-STC, and \( M \)-group, and investigated the spectral efficiency with consideration of the overhead incurred in implementations. From the perspective of spectral efficiency, \( M \)-group outperforms TBBS when the network is
large and/or the SNR margin is large (due to the larger amount of overhead required in TBBS); however, since every node in the decoded set retransmits the source message, much more power will be consumed in $M$-group. In this chapter, the energy efficiency of $M$-group and TBBS is studied to provide a more comprehensive guide for system designers to determine which strategy fits a specific application.

There has been extensive previous work on the energy efficiency, or power efficiency, of cooperative communications (see [53, 84] and references therein), but most of this research has considered only the transmission power consumption with little or no attention paid to the power consumed by the nodes that are not transmitting, which can be significant [88, 89]. In this discussion, power consumption is addressed by taking into account the power consumed in all possible modes: Transmit, Receive, Idle, and Sleep.

The rest of this chapter is organized as follows: The system model is presented in Section 3.2. In Section 3.3, energy efficiency is defined and then derived for $M$-group and TBBS. Results are presented and discussed in Section 3.4. Section 3.5 concludes this chapter.

### 3.2 System Model

Similar to that in Chapter 2, we consider a system with one source-destination pair and $N$ potential relays. The direct link between the source and the destination is assumed to be unreliable due to the large source-destination separation and/or the presence of deep fading. A two-phase relaying scheme is utilized to facilitate the direct link communication. We assume that the two phases are completed in a time period $T$ that is much shorter than the coherence time of the wireless channel; so, the channels from the source to each potential relay and those from each potential relay to the destination remain constant during one end-to-end transmission time. Uniform time allocation is considered here, that is, the data transmission time is equally divided into two phases. We indicate the channel coefficient of the link from the source to potential relay $i$, $i \in \{1, 2, \cdots, N\}$ as $h_{si}$, and the channel from potential relay $i$ to
the destination as \( h_{id} \). For the sake of simplicity, we assume all \( h_{si} \)'s and \( h_{id} \)'s are i.i.d random variables following a complex Gaussian distribution; without loss of generality, the mean is set to 0 and the variance is 0.5.

### 3.3 Energy Efficiency

We define the energy efficiency as

\[
\eta_e = \frac{1}{2} \frac{r T_e}{E} p_{suc}
\]  

(3.1)

where \( r \) is the bit rate, \( T_e \) is the effective transmission time, that is, the time consumed by the data transmission, \( p_{suc} \) is the probability of successful transmission from the source to the destination, and \( E \) is the total energy consumption during one end-to-end transmission. The factor of one-half results from the two-phase transmission process. Note, in comparison with (2.10) for the spectral efficiency, the denominator in (3.1) is the energy consumed during the two-phase relaying.

Note that, in practice, not only do the transmitting nodes consume energy, but the nodes in Receive mode, and even those in Idle or Sleep modes also consume energy. For example, Table 3.1 lists the power consumption of a commercial IEEE 802.11 wireless transceiver in different modes [88]. Note that the power consumption in Transmit mode is not the transmission power, but rather the power consumed by the transceiver in transmitting the signal. These values show that the power consumption of the non-transmitting nodes should also play an important role. Therefore, in the computation of the total energy consumption, instead of considering only the transmitting nodes, here, we take into account all possible modes: Transmit, Receive, Idle, and Sleep.

The power consumptions of the nodes in Transmit, Receive, Idle, and Sleep mode are indicated as \( P_t, P_r, P_i, \) and \( P_s \), respectively, and the four values represent an energy consumption profile. Intuitively, turning the idle nodes completely off (Sleep mode) can save a significant amount of energy. However, power is also required to “wake up” the sleeping nodes. Therefore, scheduling algorithms to conserve energy are needed, but these are often difficult to design and implement. To include this consideration
Table 3.1: Transceiver power consumption (mW)

<table>
<thead>
<tr>
<th>MODE</th>
<th>802.11b</th>
<th>802.11a</th>
<th>802.11g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep</td>
<td>132</td>
<td>132</td>
<td>132</td>
</tr>
<tr>
<td>Idle</td>
<td>544</td>
<td>990</td>
<td>990</td>
</tr>
<tr>
<td>Receive</td>
<td>726</td>
<td>1320</td>
<td>1320</td>
</tr>
<tr>
<td>Transmit</td>
<td>1089</td>
<td>1815</td>
<td>1980</td>
</tr>
</tbody>
</table>

in our analysis, the power consumption of the nodes that are neither transmitting nor receiving is indicated as $P_{\text{non}}$, which ranges from the power consumption in Sleep mode to that in Idle mode and depends on the efficiency of the scheduling algorithms.

3.3.1 Ideal Best-Select Relaying

First, we assume that the selection of the best potential relay is ideal in the sense that it is always successful and costs negligible time compared to $T$, that is, $T_e = T$. In this case, the end-to-end transmission is successful as long as (1) there is no outage in the first phase, that is, the decoded set is not empty, and (2) there is no outage in the second phase. An outage occurs if the received SNR is smaller than a given threshold. In this analysis, we assume the same SNR threshold $\gamma_{th}$ at all nodes, which means all nodes in the network have the same sensitivity.

In the first phase, for the link from the source to potential relay $i$, since $|h_{si}|^2 \sim \text{Exp}(1)$, the outage probability is

$$p = \Pr \left( \frac{P_T|h_{si}|^2}{P_N} < \gamma_{th} \right) = 1 - e^{-\delta_{th}} \quad (3.2)$$

where $P_T$ is the transmission power (not $P_t$, which is the power consumption of the node in Transmit mode), $P_N$ is the noise power, and $\delta_{th} = \frac{\gamma_{th}P_N}{P_T}$. Since all $h_{si}$’s are i.i.d., that is, no path loss is considered in the channel model, all potential relays have the same probability to decode the source signal. Then, the probability that the size of the decoded set $\mathcal{D}$ is $L$ can be written as

$$p_L = \Pr(|\mathcal{D}| = L) = \binom{N}{L} (1 - p)^L p^{N-L} \quad (3.3)$$
where $|\cdot|$ denotes the cardinality of a set. Since the source transmits and all $N$ potential relay nodes listen, the energy consumed in this phase is $(P_t + NP_r + P_{non})\frac{T}{2}$.

In the second phase, for a given non-empty $\mathcal{D}$, the outage probability is

$$p_{BS,II} = \Pr(\arg\max_{1 \leq i \leq L} |h_{id}|^2 < \delta_{th})$$

$$= \prod_{i=1}^{L} \Pr(|h_{id}|^2 < \delta_{th}) = (1 - e^{-\delta_{th}})^L$$

The energy consumption in this phase is $(P_t + P_r + NP_{non})\frac{T}{2}$ since the selected node transmits and the destination listens.

Therefore, the energy efficiency of ideal Best-Select is

$$\eta_{BS} = \sum_{1 \leq l \leq N} \frac{p_l \ rT}{2 \ E_{BS}} (1 - p_{BS,II})$$

where the total energy consumption is

$$E_{BS} = (P_t + NP_r + P_{non})\frac{T}{2} + (P_t + P_r + NP_{non})\frac{T}{2}$$

$$= [2P_t + (N + 1)(P_r + P_{non})]\frac{T}{2}$$

3.3.2 Timer-Based Best-Select Relaying (TBBS)

In practice, the implementation of best selection is not ideal because of possible collisions and a non-zero selection time. In TBBS, every node in the decoded set sets up its own timer, $t_i = \lambda/|h_{id}|^2$, where $\lambda$ is a constant system parameter. The node with the best channel to the destination has the smallest timer $t_{(1)}$ and will transmit the source information first.

As discussed in Section 2.3.2, a collision probability $p_{coll}$ is introduced in TBBS and a selection time $T_s$ is required. In the first phase, the source transmits and all $N$ potential relay nodes listen. So, the energy consumption is $(P_t + NP_r + P_{non})\frac{T_s}{2}$. In the selection process, all the nodes in the decoded set $\mathcal{D}$ are in the Receive mode to detect the signal broadcasted by the best relay; the energy consumed in this process is $[LP_r + (N - L + 2)P_{non}]T_s$. In the second phase, the best node transmits and the destination listens; the energy consumed is $(P_t + P_r + NP_{non})\frac{T_s}{2}$. 
Therefore, the efficiency becomes
\[
\eta'_{\text{BS}} = \sum_{1 \leq k \leq N} \frac{p_L}{2} \frac{r(T - T_s)}{E'_{\text{BS}}}(1 - p_{\text{BS,II}})(1 - p_{\text{coll}})
\]  
(3.7)
and the total consumed energy is now
\[
E'_{\text{BS}} = (P_t + NP_r + P_{\text{non}})\frac{T_e}{2} + (P_t + P_r + NP_{\text{non}})\frac{T_e}{2}
\]
\[
+ [LP_r + (N - L + 2)P_{\text{non}}]T_s
\]
\[
= E_{\text{BS}} + T_s\Delta E_{\text{BS}}
\]  
(3.8)
where \(\Delta E_{\text{BS}} = -P_t + (L - \frac{N+1}{2})P_r + (N - L + 2 - \frac{N+1}{2})P_{\text{non}}\).

### 3.3.3 M-Group Dis-STBC All-Select Relaying (M-group)

In M-group, all the nodes in the decoded set transmit together to forward the source information. Each of them randomly chooses one column in the underlying M-column STBC matrix to transmit. Hence, there is no time consumed in the relay selection, that is, \(T_e = T\). (Note that, if an orthogonal STBC matrix with more than two columns is used, the rate penalty \([11]\) must be taken into account.) To simplify the analysis, 2-group Dis-STBC is considered. Given the decoded set \(D\), the successful transmission probability has been derived in Section 2.4.3,
\[
p_{\text{AS,II}} = \sum_{k=1}^{L-1} \frac{\binom{L}{k}}{2L} \int_{x_1+x_2<\delta_{\text{th}}} 1 \frac{1}{k} e^{-x_2/k} \frac{1}{L-k} e^{-x_1/L-k} dx_1 dx_2
\]
\[
+ \frac{1}{2L-1} (1 - e^{-\frac{L-1}{T}})
\]  
(3.9)
In the first phase, the source transmits and all \(N\) potential relay nodes listen. In the second phase, all the nodes in \(D\) transmit, and the destination listens. Therefore, given \(D\), the total energy consumption is \((P_t + NP_r + P_{\text{non}})T^2 + [LP_r + P_r + (N - L + 1)P_{\text{non}}]T^2\).

Therefore, the energy efficiency is
\[
\eta_{\text{AS}} = \sum_{L=1}^{N} \frac{p_L}{2} \frac{rT}{E_{\text{AS}}}(1 - p_{\text{AS,II}})
\]  
(3.10)
where
\[
E_{AS} = \frac{T}{2}[(L + 1)P_t + (N + 1)P_r + (N - L + 2)P_{non}]
\] (3.11)

3.4 Simulation Results

The end-to-end transmission duration is set as \(T = 10\) msec, which is roughly 10% of the typical coherence time of a wireless system with low mobility [83], and the guard interval \(t_g = 50\) µsec. Since the absolute value of the energy efficiency is not the focus here, we set the data rate as \(r = \frac{2}{T}\) bits/s, that is, \(\frac{rT}{2} = 1\) bit, to normalize the comparison of these cooperative relaying strategies.

3.4.1 Impact of Selection Process

As discussed before, timer-based selection introduces collisions that occur if more than one relay is selected, as well as selection time consumption that reduces the data transmission time. These two factors degrade the performance of Best-Select relaying. In this section, the impact of the timer-based selection process is investigated. The energy consumption profile is set according to IEEE 802.11g in Table 3.1, and \(P_{non}\) is set equal to the power consumption in the *Idle* mode; that is, no scheduling is used to make the idle nodes sleep.

It has been shown in Section 2.5.1 that the value of \(\mu\), the ratio of \(\lambda\) to \(t_g\), is critical in determining the collision probability \(p_{coll}\) and selection time \(T_s\). Therefore, it has an important impact on the energy efficiency of TBBS. As shown in Fig. 3.1, TBBS with different values of \(\mu\) have significantly different performances. Although adaptive \(\mu\) provides the optimal performance, it is not practical to implement in a distributed way. We can see that TBBS with adaptive \(\mu\) has very similar performance to 2-group; and, when \(N\) is small, the former can even be worse than the 2-group. If a fixed \(\mu\) is used (more practical for distributed algorithms), the performance of TBBS can be much worse than 2-group.
A larger $\mu$ means that the individual timers are separated further from each other; this gives a smaller $p_{coll}$ and a larger $T_s$. A smaller $p_{coll}$ increases the end-to-end successful transmission probability, and hence, improves the energy efficiency; a larger $T_s$ decreases the effective data transmission time, and thus degrades the energy efficiency. When the network is small, $T_s$ is the main determining factor for the efficiency, and thus a smaller $\mu$ is preferred. When $p_{coll}$ is the main factor ($N$ large), a larger $\mu$ is required to keep $p_{coll}$ small.

To study the impact of the selection time $T_s$, we assume $p_{coll}$ is negligible and compare ideal Best-Select, TBBS, and 2-group; results are shown in Fig. 3.2, where $N$ is the number of potential relays. Note that the energy efficiencies are normalized by that for ideal Best-Select with $N = 3$. Clearly, if $T_s = 0$, TBBS is ideal; if $T_s = T$, that is, all the transmission time is consumed by the selection process, the efficiency...
reduces to 0. As expected and already shown in Fig. 3.1, the energy efficiencies of both TBBS and 2-group decrease with the number of potential relays. For the given energy consumption profile, when $N = 3$, the performance of TBBS is worse than 2-group if the ratio $T_s/T$ is larger than 0.23, and this value becomes 0.3 and 0.34, respectively, when $N = 6$ and $N = 9$. It means that the probability that 2-group is better than TBBS in the sense of energy efficiency is lower, that is, a higher $T_s/T$ is needed when $N$ is larger.

![Figure 3.2](image-url): Energy efficiency of 2-group and TBBS as a function of $T_s/T$, the ratio of selection time to total transmission time. (The energy efficiencies are normalized by the efficiency for ideal Best-Select with $N = 3$.)

### 3.4.2 Impact of Energy Consumption Profiles

Clearly, the amount of power consumed in different modes is critical for the energy efficiency. This impact is quantified next. As discussed before, the value of $P_{non}$ depends on the efficiency of the scheduling algorithms. At first, we assume “perfect”
scheduling is utilized, that is, $P_{\text{non}} = 0$, and study the impact of the ratio of $P_r$ to $P_t$ (Fig. 3.3). Then, we assume no scheduling algorithm is used and the nodes that are not transmitting or receiving (Idle mode) consume a significant amount of energy, i.e., $P_{\text{non}} = 0.5P_t$ (Fig. 3.4). Finally, the impact of the ratio of $P_{\text{non}}$ to $P_t$ is presented for a fixed value of $P_r/P_t$ (Fig. 3.5, $P_r/P_t = 0.5$). In all the figures, where $N$ is the number of potential relays, the energy efficiencies are normalized by that for ideal Best-Select.

![Figure 3.3: Energy efficiency of 2-group and TBBS as a function of the ratio of $P_r$ to $P_t$ ($P_{\text{non}} = 0$).](image)

For both TBBS and 2-group, the degradations in energy efficiency compared to ideal Best-Select relaying increase with the number of potential relays. The reason is that the additional consumed power compared to the ideal Best-Select relaying goes up with $N$. Another observation is that 2-group is worse than TBBS when both $P_{\text{non}}$ and $P_r$ are small, especially, when $N$ is large. In this case, the extra consumed transmission energy in the second phase of All-Select is the majority of the total energy consumption.
Figure 3.4: Energy efficiency of 2-group and TBBS as a function of the ratio of $P_r$ to $P_t$ ($P_{non} = 0.5P_t$).

On the other hand, if $P_{non}$ and $P_r$ are large, as shown in Fig. 3.4, 2-group can achieve better efficiencies than TBBS, especially when $N$ is small. The reason is that with high power consumption in Receive and Idle modes, the selection process in TBBS consumes a significant amount of energy, which degrades the energy efficiency.

The same observations apply for the results in Fig 3.5: the degradations compared to ideal Best-Select relaying increase with $N$; and 2-group is more worse than TBBS when $P_r$ and $P_{non}$ are smaller and $N$ is larger. A specific example is shown in Fig. 3.1, where the power consumption profile is set according to IEEE 802.11g in Table 3.1, and $P_{non} = P_r$. When $N$ is small, 2-group has slightly higher energy efficiencies than TBBS, even with adaptive $\mu$. Recall that the optimal value of $\mu$ depends strongly on the size of the network, the coherence time of the wireless channel, and the required guard time. Therefore, in general, TBBS is not as robust as 2-group.
3.5 Conclusions

The energy efficiency of two distributed relaying strategies is investigated in this chapter. The impact of the relay selection process on the performance of TBBS is clearly addressed, and the impact of the energy consumption profiles is also presented. Through analyses and simulations, we showed that, although Best-Select relaying conserves much more transmitting power than All-Select relaying, it is not always better in the sense of energy efficiency because (1) the implementation of the selection process degrades the performance, and (2) in some cases, the energy consumed by the nodes, not in Transmit mode, is comparable to or even higher than that consumed by those nodes that are transmitting.

The power model used here needs refinement. For example, in Best-Select relaying and All-Select relaying, the destination uses different processing to decode the

![Figure 3.5: Energy efficiency of 2-group and TBBS as a function of the ratio of $P_{\text{non}}$ to $P_t$ ($P_{r} = 0.5P_t$).]
received signal for the two strategies, and hence, the consumed power will be different. A precise power consumption model is essential to obtain accurate results; this is one obvious direction of the future work. In addition, power allocation algorithms would improve the energy efficiency, however, distributed allocation algorithms are difficult to design and implement. Scheduling algorithms that can minimize $P_{\text{non}}$ by decreasing the duration that nodes are in $\text{Idle}$ mode would also save a significant amount of energy.

In practice, the protocols applied in the upper-layers, e.g., MAC layer, could also affect the efficiencies of the relaying schemes. For example, the probability of collisions that occur when more than one relay is selected as the “best” would reduce dramatically if IEEE 802.11 RTS/CTS is utilized to reserve the channel in the selection process of Best-Select relaying. However, the implementation of these protocols consumes energy. Therefore, a holistic analysis at the system level is necessary to reveal the “realistic” efficiencies of distributed cooperative relaying.
Chapter 4

COOPERATIVE ROUTING WITH MINIMUM OVERHEAD

Geographic routing is a routing strategy in wireless networks; this approach is scalable because it makes use of position information, and every node can make a decision based on local information. Cooperative routing, which transfers the benefit of cooperative communications into the network layer, is proposed to provide better routing performance by utilizing cooperative diversity to combat the effects of fading. In this chapter, we incorporate geographic routing into cooperative routing and propose Location-Aware Cooperative Routing (LACR), in which any node that has received Route Request (RREQ) will make a decision on whether to rebroadcast the information it received, as well as how to rebroadcast, based on its position information and its availability to cooperate. Note that LACR is a fully distributed routing algorithm that incurs a minimum amount of overhead. A theoretical analysis about the impact of cooperative communications on the transmission range extension is presented to provide a guide for measuring the potential capability of each node in routing the source information. Through simulation results, we show that LACR performs better in the sense of a smaller probability of failing to find a route and a higher throughput compared to SISO routing and SISO-based cooperative routing.

4.1 Background and Related Work

Geographic routing, also known as position-based routing, is a promising routing strategy in wireless networks, especially when the size of the network is large [93, 94]. The most attractive advantage of geographic routing is its scalability, which has a two-fold meaning here. First, compared to proactive/table-driven routing, e.g., global state routing (GSR) [95], geographic routing does not have to maintain a routing table
at each node to track the global topology since every node makes a decision based on local topology (usually, the position information of its own one-hop neighbors). Second, compared to flooding-based reactive/demand-driven routing, e.g., dynamic source routing (DSR) [96], geographic routing significantly reduces the redundancy caused by flooding and thus saves energy. It also converges much faster since fewer backoffs are needed to avoid collisions.

Generally speaking, geographic routing makes routing decisions based on a distance-based metric, which captures the long-term performance of wireless links. However, this metric cannot keep track of the instantaneous status of the wireless channel because of the lack of information about fading and shadowing. Despite that, the simplicity and scalability make geographic routing a good candidate for multi-hop wireless networks.

Cooperative routing makes use of cooperative diversity from the physical layer to benefit the network layer by improving the performance of routing in wireless networks. It is well understood that cooperation among single antenna nodes can provide diversity to combat fading in wireless communications [19, 20], which results in a much lower required SNR to maintain the same outage probability. The reduced requirement for received signal power can be used to extend the transmission range and/or increase the transmission rate.

Several cooperative routing protocols have been proposed [40, 41, 43, 45]. In [40, 41], cooperative routing is formulated as optimization problems, and algorithms based on global optima or some heuristics to find locally optimal cooperative paths are proposed. Since the basic idea is to find the optimal path among all feasible ones, these algorithms are inherently not scalable and difficult to implement in practice. In [43], the authors divide the cooperative routing problem into two tasks: first find a primary SISO route, and then shorten the SISO route by selecting relays to construct the cooperative path. In [45], cooperation-related information, such as the number of neighbors, is added to the RREQ messages. An existing ad hoc routing protocol such as DSR is used to find the route, except that the destination uses the collected
cooperation-related information to make decisions about the cooperative route.

Since both [43] and [45] rely on SISO routing protocols, they cannot make full use of the benefits provided by cooperative transmissions. For example, there might be a void in the network which cannot be easily spanned with SISO transmissions, but which, with the increased range of cooperative transmissions, can easily be bridged. However, neither the algorithm in [43] nor that in [45] can find a cooperative path even though one may exist.

LACR is the first routing protocol, to the best of our knowledge, to incorporate cooperation into geographic routing. This approach can significantly improve the routing performance with very low overhead. The basic idea is explained in Section 4.2, followed by an analysis of some of the design issues in Section 4.3. In Section 4.4, we describe LACR in detail, and its performance is presented in Section 4.5. Section 4.6 concludes this chapter.

4.2 Basic Concept of LACR

Greedy Forward Routing (GFR), one form of geographic routing, tries to deliver RREQ as close as possible (greedy) to the destination. There exist many variations of GFR [97,98]. LACR uses the same motivation as GFR, that is, the closer to the destination, the better. However, there are two main differences that distinguish LACR from any GFR protocol: First, LACR allows only the best contender ¹ to take charge of the rebroadcast of RREQ. Second, the selection criterion is based not only on position information, but also on the potential benefits of the cooperative transmissions.

In this work, we assume every node in the network has access to its own up-to-date position information, and when a route request is initiated, the source already knows the position of the destination and inserts its own position and the position of the destination into RREQ. Once a node receives RREQ, it calculates the distance from itself to the destination, which is one critical metric for measuring the performance of

¹ The nodes that received RREQ from the previous hop are the contenders in the current hop.
this contender. The other metric is the number of cooperating nodes, which determines the potential benefit of cooperative transmissions. Based on these two metrics, each contender can individually calculate its own performance; this is explained in detail in Section 4.3.3. Then, a timer-based selection scheme [67] is used to select the best contender. After that, the best contender will collaborate with its neighboring nodes to broadcast RREQ. This process continues until RREQ arrives at the destination. Finally, the destination initiates a new RREQ back towards the source; this contains information about the route that was determined from the source to itself. This new RREQ, at the same time, serves as the Route Reply (RREP) to the RREQ from the source. Once this new RREQ arrives at the source, the routing process is completed, and both the forward route (from the source to the destination) and the backward route (destination to source) are discovered.

4.3 Analysis of Design Issues

4.3.1 Cooperative Transmission Strategy

It has been shown that Dis-STC cooperative transmissions can offer full spatial diversity in the number of cooperating nodes [21]. In Dis-STC, each cooperating node is assigned one column of the STC matrix to transmit. To implement Dis-STC cooperative transmissions, either a central controller or full inter-node information exchange is needed to build the STC matrix based on the number of cooperating nodes and then assign the columns among all the cooperating nodes. However, this centralized approach suffers from signaling overhead, which is not scalable for large networks. To exploit cooperative diversity in a distributed manner with less overhead, in [68] a decentralized strategy called M-group Dis-STC is proposed. In contrast to Dis-STC, once a node successfully receives a signal, it automatically becomes a cooperating node and randomly selects one column of the underlying STC matrix to transmit. In this way, cooperative communications is implemented in a distributed manner and incurs minimum overhead.
As pointed out in [72, 74, 85, 99], 2-group is much more spectral- and energy-efficient than Dis-STC in most cases. To realize cooperative communications in a distributed fashion with minimum overhead, which is important to make a routing protocol scalable, for LACR, we adopt 2-group Dis-STC as the cooperative transmission strategy. In this case, there is no rate penalty [82] and the use of an Alamouti Code [10] is fairly easy to implement.

4.3.2 Capability to Cooperate

LACR utilizes cooperative diversity to extend the transmission range, defined as the area within which the communication between the transmitting node and an arbitrarily deployed node has an expected outage probability below some given threshold, $p_{\text{out}}^{th}$. In this subsection, we analyze the impact of the number of cooperating nodes on the transmission range and provide an approximation to guide the design of performance metric of the contenders.

To get the number of cooperating nodes, every contender has to notify its nearby nodes that it has already received the signal and is ready to cooperate. Doing this incurs many inter-node communications, which requires a complicated scheduling scheme to coordinate all the contenders’ notifications. To eliminate this overhead and complexity, we utilize the number of neighbors to approximate the number of cooperating nodes. Since the probability that every neighbor receives the signal from the previous hop increases with a smaller neighborhood, in LACR, we reduce the size of the neighbor by increasing the SNR threshold of the received HELLO packet.

In the rest of this subsection, we analyze the impact of the capability to cooperate (that is, the number of neighbors) on the transmission range. Let $h_{ij}$ be the instantaneous channel coefficient of the link from node $i$ to node $j$, which is a complex Gaussian random variable, with the real and imaginary components having zero mean and variance $\sigma^2$. Then the power gain of the channel, $|h_{ij}|^2$, is an exponential random
variable with mean $\beta_{ij}$, which is

$$\beta_{ij} = \frac{2\sigma^2 G}{(d_{ij}/d_0)^\alpha}$$

(4.1)

where $d_{ij}$ is the distance from node $i$ to node $j$, $d_0$ is a reference distance (assumed here to be 1 meter), $\alpha$ is the path-loss exponent, and $G$ is a constant related to the frequency used and the features of the transceivers [3]. To simplify the analysis, we set $2\sigma^2 G = 1$.

For any realization of $M$-group Dis-STC cooperative communication among $N$ cooperating nodes, one subset of these nodes, which consists of the nodes selecting the $k$th column of the underlying STC matrix, construct a virtual channel from themselves to the receiver [100]. The virtual channel coefficient can be expressed as

$$h'_k = \sum_{i \in g_k} h_{id}$$

(4.2)

where node $d$ is the receiver, $k \in \{1, 2, ..., M\}$ and $g_k$ is the set of the cooperating nodes that select the $k$th column to transmit. Due to the independence of the coefficients $h_{id}$, $|h'_k|^2$ is an exponential random variable with mean

$$\beta'_k = \sum_{i \in g_k} \beta_{id} = \sum_{i \in g_k} \frac{1}{d_{id}^\alpha}$$

(4.3)

By applying the conventional coherent detection algorithm for STC, we have

$$\gamma_d = \frac{P_t \sum_{k=1}^{M} |h'_k|^2}{P_N}$$

(4.4)

where $P_t$ is the transmitting power, $P_N$ is the noise power and $\gamma_d$ is the SNR at the receiver. The noise power $P_N$ is assumed to be the same for any node and, without loss of generality, we set $P_N = 1$. An outage will occur if $\gamma_d$ is below some given SNR threshold, $\gamma_{th}$. For the long-term performance, we can write the expected outage probability as

$$p_{out} = \int_{\mathbf{H}} f(\mathbf{X}) Pr\{\sum_{k=1}^{M} P_t|h'_k|^2 \leq \gamma_{th}\}d\mathbf{X}$$

(4.5)
where $\mathbf{X}$ is the vector $[|h'_1|^2, |h'_2|^2, ..., |h'_M|^2]^T$, $\mathbf{H}$ is the feasible region of $\mathbf{X}$ and $f(\mathbf{X})$ is the probability density function of $\mathbf{X}$.

Denote all the possible realizations of $M$-group Dis-STC cooperative communication among $N$ decoded nodes as $\Delta$. One element of it is $\delta^{(u)} = [\delta_1^{(u)}, \delta_2^{(u)}, ..., \delta_M^{(u)}]^T$, where $\delta_k^{(u)}$ indicates the number of nodes that select the $k$th column and $u \in \{1, 2, ..., |\Delta|\}$. Clearly, for any $u$, $\sum_{k=1}^M \delta_k^{(u)} = N$. To simplify the analysis, we assume that every cooperative node has the same distance, $d$, from itself to the receiver. Therefore,

$$\beta'_k = \frac{\delta_k^{(u)}}{d^\alpha}$$

and (4.5) can be rewritten as

$$p_{\text{out}} = \sum_{\Delta} Pr(\delta^{(u)}) \int \prod_{k=1}^M \frac{d^\alpha}{\delta_k^{(u)}} e^{-\frac{x_k d^\alpha}{\delta_k^{(u)}}} d\mathbf{X}$$

where $x_k = |h'_k|^2 \sim \text{exp}(\frac{\delta_k^{(u)}}{d^\alpha})$. $Pr(\delta^{(u)})$ is the probability that $M$-group Dis-STC cooperative communication is realized as $\delta^{(u)}$ and $\theta = \frac{2\mu}{P_t}$.

When $M = 2$, that is, 2-group Dis-STC cooperative communication is used, (4.7) becomes

$$p_{\text{out}}^{M=2} = \frac{2}{2^N} \int_{x_1 \leq \theta} \frac{d^\alpha}{N} e^{-\frac{x_1 d^\alpha}{N}} dx_1 + \sum_{l=1}^{N-1} \left( \begin{array}{c} N \\ 2^N \end{array} \right) \int_{x_1 + x_2 \leq \theta} \frac{d^\alpha}{l} e^{-\frac{x_1 d^\alpha}{l}} \frac{d^\alpha}{N - l} e^{-\frac{x_2 d^\alpha}{N - l}} dx_1 dx_2$$

If $N = 1$, then we have

$$p_{\text{out}}^{M=2} = 1 - e^{-\theta d^\alpha}$$

When $N > 1$

$$p_{\text{out}}^{M=2} = \frac{1}{2^{N-1}}(1 - e^{-\frac{d^\alpha}{N}}) + \sum_{l=1}^{N-1} \left( \begin{array}{c} N \\ l \end{array} \right) \left[ 1 + \frac{l}{N - 2l} e^{-\frac{d^\alpha}{N - 2l}} - \frac{N - l}{N - 2l} e^{-\frac{d^\alpha}{N - 2l}} \right], \text{ N odd}$$

(4.10)
\[ p_{out}^{M=2} = \frac{1}{2^{N-1}} (1 - e^{-\frac{\theta d}{N}}) + \sum_{l=\frac{1}{2}}^{N-1} \frac{(N)}{l} e^{-\frac{\theta d}{N}} - \frac{N - l}{N - 2l} e^{-\frac{\theta d}{N}} \]

\[ + \frac{(N)}{2^{N}} \left[ 1 - \frac{2d^\alpha}{N} (\theta + \frac{N}{2d^\alpha} e^{-\frac{2\theta d^\alpha}{N}}) \right], \text{ N even} \quad (4.11) \]

A straight-forward illustration of the effect of the number of neighbors, \(N\), on the cooperative transmission range, is presented in Fig 4.1. The outage probability is a monotonically decreasing function of \(N\) for a given distance, as expected. Therefore, for a given \(p_{out}^{th}\), the transmission range extends further with more cooperating nodes. Assuming \(p_{out}^{th} = 10^{-3}\), the normalized cooperative transmission range with \(N \in [1, 8]\) is presented in Table 4.1; when \(N = 1\), this corresponds to the SISO transmission range.

**Figure 4.1:** The impact of distance on outage probability when 2-group Dis-STC is used as cooperative transmission strategy
4.3.3 Performance Metric

Ideally, at each hop, the contender that can deliver the signal closest to the destination will be selected to take charge of the rebroadcast of the message. Then, a performance metric is provided as follows,

\[ \nu_i = d_{iD} - g(N_i) r_{SISO} \]  

(4.12)

where \( d_{iD} \) is the distance from node \( i \) to the destination \( D \). \( g(N_i) \) is the mapping between the number of cooperating nodes, \( N_i \), and the transmission range extension factor with respect to SISO transmission range, \( r_{SISO} \). This mapping can be determined from (4.8), or be set up using Table 4.1 as an approximation.

4.3.4 Selection Method

The selection of the best contender, that is, the one with the smallest \( \nu \), is accomplished by a timer-based selection scheme [67]. Each contender \( i \) sets up a timer \( t_i \) according to its locally calculated performance metric, \( \nu_i \), in such a manner that the timer at the best node will expire first. After that, the best node will notify the other contenders that it has been selected, and then the other nodes will back off.

Basically, there are two ways for the best node to notify the other nodes: (1) broadcast a notification, and (2) send a notification to the destination and the destination then broadcasts the selection result. The former may suffer from hidden
contenders, while the latter mode is not suitable in multi-hop routing since the next hop is unknown to the current contenders.

In LACR, we use the timer-based selection scheme as follows. First, normalize $\nu_i$ as

$$\nu_{i}^{nr} = \frac{d_{SD} - \nu_i}{d_{SD} + g(N_{max})r_{SISO}}$$  \hspace{1cm} (4.13)$$

where $d_{SD}$ is the distance between $S$ and $D$, $N_{max}$ is the maximal number of cooperating nodes. Although more cooperators can provide better performance, the energy consumption is also increasing. In addition, according to Table 4.1, the incremental improvement is decreasing with respect to $N$. In LACR, we set $N_{max} = 8$, i.e., at most 7 neighbors will transmit along with the best contender. Obviously, $-g(N_{max})r_{SISO} \leq \nu_i \leq d_{SD}$ and so, $\nu_{i}^{nr} \in [0, 1]$. Then, the timer at one contender is

$$T_i = (1 - \nu_{i}^{nr})T_{max}$$  \hspace{1cm} (4.14)$$

where $T_{max}$ is the maximal selection time.

4.4 Protocol Description

In this section, we describe the LACR protocol in detail. In LACR, each node has access to its own up-to-date location information using some location services [101, 102], e.g., GPS devices. Periodically exchanged HELLO packets are applied to make sure that each node has the information of the number of its neighbors. LACR works as follows:

- **Initiating RREQ.** The source will initiate RREQ if it has data to send to a destination, but does not have a route. The RREQ will be broadcasted after being inserted into the position information of the source and the destination.

- **Selecting the best contender.**

  Action upon reception of RREQ: Once a node receives a RREQ, it first checks whether it has already processed this RREQ. If this is not the first time it has
received this RREQ, the node discards it. Otherwise, it automatically becomes a contender and then calculates its own performance metric according to (4.12). The parameters needed to calculate it are: (1) $d_{iD}$, the distance from itself to the destination, which can be easily calculated based on its own location and the position of the destination included in the received RREQ; and (2) $N_i$, the number of neighbors, which can be obtained from the HELLO packets. Once the contender computes its own performance metric, it immediately sets up a timer according to the method described in Section 4.3.4.

Selection based on timer: After setting up the timer, each contender waits for the expiration of its own timer. Once the timer expires, the contender broadcasts a notification (NTF) message that includes its own ID. Any other contender that receives this NTF checks whether it is a neighbor of the node that sent the NTF. If it is, the node will become a cooperating node; if it is not, it discards the RREQ and becomes idle. Note that the NTF can be sent out with a higher power or using a more robust coding technique to reduce the probability of hidden contenders.

- **Transmitting RREQ cooperatively.** A group of contenders are selected to cooperatively transmit the RREQ. Each of them broadcasts the RREQ using one randomly selected column of the underlying Alamouti Code. In this way, cooperative transmission is implemented using 2-group Dis-STC.

- **Sending the RREP back.** The process of “selecting the best contender” and “transmitting the RREQ cooperatively” continues until the RREQ arrives at the destination. When the destination receives the RREQ, it immediately broadcasts a NTF message, and any current contender discards the received RREQ. After that, the destination initiates a new RREQ, which includes the information of the discovered route from the source to itself, to discover a reverse route from the destination to the source, proceeding exactly the same way as sending the RREQ from the source to the destination. Obviously, this new RREQ serves as the RREP at the same
time. The reason for this reverse path discovery is that, by incorporating cooperative transmission, links along the discovered route become asymmetric, so that the found route cannot be used to feed back the RREP. Note that this reverse route is also used to feed back the ACK when data packets are transmitted.

- **Transmitting data packets.** The routing process is finished when the RREP, that is, the new RREQ initiated from the destination, arrives at the source. Then the data packets flow along the forward route found by LACR and the necessary ACK’s flow to the source through the established reverse route.

### 4.5 Simulation Results

#### 4.5.1 Setting

We implement LACR in Qualnet 3.7 [103] and compare the performance with DSR [96]. We use a uniform placement strategy for nodes; we divide the 1000 m by 1000 m terrain into grids and place a node randomly within each grid cell. In this way, we avoid, with a high probability, voids in the network, which are bottlenecks to geographic routing. The source and the destination are located at the ends of the diagonal.

We assume that the wireless channel remains constant during the transmission of a single packet. The transmitting power is configured so that the SISO transmission range is around 250 m. IEEE 802.11 is used as the MAC protocol, and $M$-group DisSTC is implemented to support cooperative transmission. A Constant-Bit-Rate (CBR) application, with each message of size 512 bytes, is used to generate the traffic. Each of the results presented in this section is the average of 500 runs and the error bars show the 95% confidence intervals.

#### 4.5.2 Results and Discussion

##### 4.5.2.1 Probability of Successful Route Discovery

In this subsection, we show the probability that a feasible route is discovered by initiating a limited number of RREQs. For each incoming packet from the application
layer, one RREQ will be initiated if there is no discovered route. Specifically, 100 packets are produced to transmit, and the number of nodes is varied to evaluate the effect of network density.

![Graph showing the probability of successful route discovery as a function of the number of nodes. The graph compares DSR and LACR, with LACR performing better with higher probability, especially when the network is sparse. Conversely, DSR finds a route with less setup time. Since both protocols discover a route only when there are more than 30 nodes in the network, the plots start from 30 nodes. Since one RREQ is initiated every second, the setup delay can, at the same time, serve as a measurement of the number of required RREQs to discover a route. In addition to fewer initiated RREQs, LACR eliminates the flooding phenomenon.]

**Figure 4.2:** Probability of successful route discovery as a function of the number of nodes

The results in Fig. 4.2 indicate that LACR performs better than DSR in the sense that it can find a route with higher probability, especially when the network is sparse. The reason is that LACR makes use of cooperative transmissions to extend the transmission range. Now, the unconnected nodes using the SISO transmissions can communicate with each other. In addition, LACR can find a route with less setup time, which is shown in Fig. 4.3. Since LACR and DSR have almost the same probability to discover a route only when there are more than 30 nodes in the network, the plots in Fig. 4.3 start from 30 nodes. Since one RREQ is initiated every second, the setup delay can, at the same time, serve as a measurement of the number of required RREQs to discover a route. In addition to fewer initiated RREQs, LACR eliminates the flooding phenomenon.
of RREQ and, therefore, it can successfully discover a route with higher probability and less overhead.

![Route setup delay as a function of the number of nodes](image)

**Figure 4.3:** Route setup delay as a function of the number of nodes

### 4.5.2.2 Packet Delivery Ratio

In this subsection, we focus on the packet delivery ratio, that is, the percentage of successfully received packets. A packet will be dropped if (1) there is no route found, or (2) after 4 retransmissions (which is a typical maximum retransmission count used in practical systems), and the source receives no ACK from the destination. In Fig. 4.4, we see that LACR performs better than DSR when the network is sparse. However, LACR’s performance is a little worse than DSR when the network is dense, since DSR has more feasible paths to choose from. It is also more difficult to coordinate the cooperative transmissions. In addition, $M$-group Dis-STC cannot guarantee the full cooperative diversity due to its random selection of STC columns.
Although building a cooperative path based on a SISO route discovered by DSR (for example, see [43]) can improve the packet delivery ratio, this approach cannot achieve a higher probability of successful routing discovery than DSR. It also suffers from the overhead from both the SISO route discovery and the cooperative route construction. To overcome these drawbacks, LACR is designed to, with less overhead, enhance the chance that a feasible route can be found.

4.5.2.3 End-to-End Delay

Although LACR has a slightly lower packet delivery ratio than DSR when the network is dense, LACR always enjoys a lower average end-to-end delay, as shown in Fig. 4.5. The end-to-end delay is the time required to successfully transmit one packet from the source to the destination. For the same packet delivery ratio, a smaller end-to-end delay means a higher throughput. The smaller end-to-end delay of LACR results
from (1) less time needed to discover a route, which is shown in Fig. 4.3, and (2) a smaller hop-count due to the longer cooperative transmission range.

![Figure 4.5: Average end-to-end delay as a function of the number of nodes](image)

### 4.6 Conclusions

In this chapter, we described LACR, a routing protocol for multi-hop wireless networks, which incorporates cooperative transmissions into geographical routing. To make the routing protocol scalable, the $M$-group Dis-STC cooperative technique was applied, and an analysis of the cooperative transmission range was provided to guide the design of LACR. Simulation results show that LACR performs well, providing a high probability to discover a route with low overhead; these advantages are particularly apparent when the network is sparse.

An interesting topic for further work is to compare LACR with SISO-based cooperative routing schemes with respect to the performance-overhead trade-off. A
routing algorithm named cooperative source routing (CSR), which is based on Dis- STC, is proposed in [104]; a natural follow-up is to compare the performance of the CSR and LACR, along with the SISO-based cooperative routing schemes [43].
Chapter 5
INTERFERENCE MANAGEMENT: SPATIAL REUSE SCHEDULING

Interference management has been shown to be critical for the design of efficient wireless networks. Among potential interference management techniques, spatial reuse is widely employed in practical wireless networks. It is also well known that cooperative communications has the potential to significantly improve network performance. How does cooperation affect spatial reuse design in wireless networks? To answer this question, as a start, we investigate the impact of reuse on the throughput of multi-hop linear cooperative networks. The reuse that maximizes the network throughput is then derived. Simulation results are provided to illustrate the importance of choosing an appropriate reuse factor.

5.1 Background and Related Work

In wireless multi-hop networks, spatial reuse is essential to balance the interference level and the channel utilization [56]. On the one hand, spatial reuse improves the channel utilization by enabling concurrent transmissions, which, in turn, can improve the network throughput. On the other hand, spatial reuse introduces co-channel interference that degrades the performance. Therefore, determining the optimal spatial reuse is critical for efficient multi-hop wireless transmission. There has been some research investigating optimal spatial reuse, particularly in linear networks [58,61,62]. In addition, several practical schemes have been proposed to implement adaptive reuse by tuning the carrier-sensing thresholds [59,105].

Cooperative communications has been shown to be effective in combating fading on wireless channels, and thus has the potential to dramatically improve performance [23]. To exploit the potential benefits of cooperation in wireless networks, a
variety of cross-layer designs have been proposed to efficiently implement cooperation \[33,35,48\]. For example, in [33], the authors propose cooperative MAC protocols to enable cooperative relaying; in \[43,104,106\], several cooperative routing algorithms are proposed to establish wireless multi-hop cooperative routes. It has been shown that incorporating cooperation dramatically improves the network throughput.

Although reuse in multi-hop wireless networks is a well-studied topic, the impact of incorporating cooperation on spatial reuse scheduling is an open problem. In this paper, focusing on multi-hop wireless networks, we investigate the impact of applying cooperative communications, and derive the reuse scheduling that maximizes the network throughput. We show, through analyses and simulations, that cooperation with the appropriate reuse can significantly improve the network throughput.

In addition to applying cooperative communication for spatial diversity only, cooperation can also be utilized to mitigate interference. Hence, we further investigate incorporating cooperative beamforming to actively manage interference. It is well known that the performance of cooperative beamforming strongly depends on the accuracy of the collected CSI \[107\]. Therefore, we jointly optimize the transmission mode for beamforming and the reuse factor, with and without perfect CSI. Simulation results are provided to illustrate the performance improvement provided by applying cooperative beamforming, as well as the impact of imperfect CSI.

The rest of this chapter is organized as follows: In Section 5.2, we describe the system model. The analysis of spatial reuse in multi-hop cooperative networks is presented in Section 5.3 and 5.4. In Section 5.5, we provide simulation results. Section 5.6 concludes this chapter.

5.2 Multi-Hop Cooperative Networks

We focus on one-dimensional wireless networks, also called linear or chain networks. Although linear networks are over-simplified, they are widely-used models to initiate the study of complicated network problems and have been shown to be capable of revealing insights that inspire practical solutions \[108,109\]. In Section 5.5, we
provide simulations for non-linear, more realistic networks.

Multi-hop linear cooperative networks are constructed by applying cooperative routing algorithms (for example, see [43]). As shown in Fig. 5.1, there is a SISO route between the source (S) and the destination (D), that is, \( S \rightarrow 1 \rightarrow \cdots \rightarrow i \rightarrow \cdots \rightarrow D \) with \( L \) hops. We apply cooperative communication to improve the performance, that is, the neighboring nodes of node \( i \) simultaneously transmit to the next hop, which essentially forms a VMISO (Virtual Multiple-Input-Single-Output) link. The cooperating nodes, e.g., node \( i \) and its neighbors, form a cluster \( C_i \) and the nodes that belong to the SISO route, e.g., node \( i \), are the cluster-heads, which take charge of disseminating the information among the neighboring nodes.

![Figure 5.1: A multi-hop linear wireless cooperative network](image)

The data transmission from \( S \) to \( D \) works as follows: The source, \( S \), broadcasts the message, which its neighboring nodes receive and decode. Since the distances among the neighboring nodes are usually quite small, we assume that, by applying error correction, the local information dissemination is perfect in the sense that all the nodes in the cluster will have a correct copy. The cluster, including \( S \) itself and all its neighbors, then cooperatively forwards the signal to the next cluster, specifically, node 1. Upon receiving and decoding the message without outage, node 1 disseminates the message among its own neighborhood and, together with its neighboring nodes, forms a VMISO link to forward the message to the next cluster. This process continues until the message reaches the destination, \( D \).
Clearly, every hop along this cooperative network is a VMISO link, which reduces the outage significantly compared with a SISO channel. However, it also complicates the transmission by introducing an extra local message dissemination process, which doubles the per-hop transmission duration and thus halves the channel utilization (assuming the broadcast inside the cluster and the transmission between clusters use the same data rate). Therefore, a factor of $1/2$ is included in the following analysis to represent this efficiency loss.

### 5.3 Dis-STC Based Cooperation

There are several methods for coordinating the simultaneous transmissions from multiple transmitters; among these, Dis-STC [21] is attractive because it provides full spatial diversity and only requires CSI at the receiver. In this section, we assume Dis-STC is applied for cooperation and focus on the throughput of the network, which is defined as the number of bits per second correctly received at destination $D$. First, without spatial reuse, we investigate the impact of incorporating cooperation. Then, the impact of spatial reuse on the network throughput is studied and the reuse factor that maximizes the throughput is derived.

#### 5.3.1 Impact of Cooperation

Focusing on one hop, say hop $i$ from cluster $C_i$ to $C_{i+1}$, denoting the channel coefficient from node $j, j \in C_i$, to node $i + 1$, the cluster-head of $C_{i+1}$, as $h_j$, the SNR of the SISO link from node $j$ to node $i + 1$ is

$$\gamma_j = \frac{P_j|h_j|^2}{P_N(d_j/d_0)^\alpha}$$  \hspace{1cm} (5.1)

where $P_j$ is the transmit power of node $j$, $P_N$ is the noise power at the receiver, $d_j$ is the distance to the receiver, $d_0$ is the reference distance (assumed here to be 1 meter), and $\alpha$ is the path-loss exponent. At the receiver, using conventional coherent detection for STC, the SNR of the VMISO link is the sum of the SNRs of the individual SISO links, that is, $\gamma = \sum_{j \in C_i} \gamma_j$. 

76
We assume all the nodes inside a cluster use the same transmit power $P_T$ and transmission bit rate $r$. We also assume the signal experiences independent and identically distributed Rayleigh fading; thus all the $|h_j|^2$ are exponentially distributed with mean 1. In general, the distances among neighboring nodes are much smaller than the per-hop separation. Hence, we assume that the distance from each node in node $i$’s neighborhood to node $i + 1$ is approximately the same (denote as $d$), which has been shown to be a reasonable assumption in practical wireless networks [43]; the impact of randomly deployed relays is illustrated in Section 5.5.

An outage occurs if the required bit rate exceeds the channel capacity; so, assuming additive white Gaussian noise, we have

$$p_{\text{out}} = \Pr \{ B \log(1 + \gamma) < r \}$$

$$= \Pr \left\{ B \log \left( 1 + \sum_{j=1}^{N} \frac{P_T|h_j|^2}{P_Nd^\alpha} \right) < r \right\}$$  \hspace{1cm} (5.2)

where $B$ is the available bandwidth, and $N$ is the size of the cluster. Note that $N = 1$ corresponds to a SISO link without cooperation. Since the $h_j$’s are independent and $|h_j|^2 \sim \text{Exp}(1)$, we have $\sum_{j=1}^{N} |h_j|^2 \sim \text{Erlang}(N, 1)$ [110]. Therefore

$$p_{\text{out}} = \Pr \left\{ \sum_{j=1}^{N} |h_j|^2 < \frac{(2\pi - 1)P_Nd^\alpha}{P_T} \right\}$$

$$= 1 - \sum_{j=0}^{N-1} \frac{1}{j!} e^{-\frac{(2\pi - 1)P_Nd^\alpha}{P_T}} \left[ \frac{(2\pi - 1)P_Nd^\alpha}{P_T} \right]^j$$  \hspace{1cm} (5.3)

where $\eta = r/B$ is the normalized bit rate (spectral efficiency). It has been shown, for example in [111], that applying cooperation can significantly reduce the outage; and, as expected, more cooperating nodes, i.e., larger $N$, results in better performance.

### 5.3.2 Impact of Spatial Reuse

Given the per-hop data rate $r$, the per-hop throughput can be written as $\frac{1}{2}r(1 - p_{\text{out}})$, which reflects the rate of successfully delivered bits. If no spatial reuse is applied, that is, at any time, there is only one active hop in the network, the throughput of the
network is $1/L$ of the per-hop throughput. This is very inefficient, especially when $L$, the number of hops, is large.

To improve the throughput, a spatial reuse factor $K$ ($2 \leq K < L$) should be introduced so that nodes $K$ hops away can transmit simultaneously, as shown in Fig. 5.2. It is well understood that, in spite of the interference introduced by reuse, an appropriate reuse scheme could dramatically improve the network throughput. On the one hand, with a smaller $K$, more nodes can transmit concurrently, which increases the channel utilization, and, hence, has the potential to improve the throughput. On the other hand, a smaller $K$ suggests more aggressive reuse, which introduces more interference and might degrade the throughput.

![Spatial reuse in linear cooperative networks](image)

**Figure 5.2:** Spatial reuse in linear cooperative networks

In order to obtain a tractable, yet characteristic-conserving, model to investigate the impact of spatial reuse, we assume that (i) the interference coming from the adjacent two interferers is dominant (this is reasonable, especially when the path-loss exponent is large); and (ii) the network is sufficiently large that almost all nodes have more than one interferer, that is, the number of the cluster-heads that are close to the source or the destination and have only one interferer is much smaller than the number of all the intermediate cluster-heads. With these assumptions, only the interference coming from $K$ hops to the left of the transmitting cluster and $K$ hops to the right are considered. Focusing on one hop, say hop $i$, the outage of using Dis-STC is then
\[ p_{\text{out}}^{\text{ST}} = \Pr \left\{ \log \left( 1 + \frac{P_T g_1^2 d^{-\alpha}}{(K-1)\alpha} + \frac{P_T g_2^2 d^{-\alpha}}{(K+1)\alpha} + P_N \right) < \eta \right\} \]

\[ \approx \Pr \left\{ |g_1|^2 < (2^{\eta} - 1) \left[ \frac{|g_2|^2}{(K-1)\alpha} + \frac{|g_3|^2}{(K+1)\alpha} \right] \right\} \]  

(5.4)

where \(|g_1|^2\) is the gain of the VMISO link under consideration, and \(|g_2|^2\) and \(|g_3|^2\) are the gains for the interfering links, that is, the links from the two dominant interferers to the intended receiver, node \(i+1\). The noise power is assumed negligible compared with the interference; the impact of \(P_N\) is included in the simulations.

Since Dis-STC is used, full diversity gain is achieved and \(|g_1|^2 \sim \text{Erlang}(N, 1)\); for the interfering links, no coordination is applied, so \(|g_2|^2 \sim \text{Exp}(1/N)\) and \(|g_3|^2 \sim \text{Exp}(1/N)\). Therefore, it is easy to show that

\[ p_{\text{out}}^{\text{ST}} = \int_{x_1 < a_1 x_2 + a_2 x_3} \frac{x_1^{N-1} e^{-x_1} e^{-\frac{x_2}{N}} e^{-\frac{x_3}{N}}}{N} \, dx_1 \, dx_2 \, dx_3 \]  

(5.5)

where \(x_m = |g_m|^2\), \(m = 1, 2, 3\), \(a_1 = \frac{2^\eta - 1}{(K-1)\alpha}\), and \(a_2 = \frac{2^\eta - 1}{(K+1)\alpha}\). As derived in [111], by applying the approximations \(a_1 \approx \frac{2^\eta - 1}{K\alpha}\) and \(a_2 \approx \frac{2^\eta - 1}{K\alpha}\), the outage becomes

\[ p_{\text{out}}^{\text{ST}} \approx \frac{a N (a + 1 + \frac{1}{N})}{(a + \frac{1}{N})^{N+1}} \]  

(5.6)

where \(a = \frac{2^\eta - 1}{K\alpha}\).

Clearly, \(a\) increases when \(\eta\) increases and/or \(K\) decreases. From (5.6), given the fact that \(N \geq 1\) and \(a > 0\), \(p_{\text{out}}^{\text{ST}}\) increases when \(a\) increases. Therefore, with fixed \(N\), \(p_{\text{out}}^{\text{ST}}\) increases when \(K\) decreases (more interference) and/or \(\eta\) increases (higher channel quality requirement). In addition, with fixed \(K\) and \(\eta\), the outage decreases when \(N\) increases, indicating the benefits of cooperation.

### 5.3.3 Optimal Spatial Reuse Scheduling

As discussed before, in a network without reuse, the throughput is \(\frac{r(1-p_{\text{out}})}{2L}\), which represents the number of bits successfully delivered from \(S\) to \(D\) per second. If reuse factor \(K\) is used, the throughput is \(\frac{L r(1-p_{\text{out}}^{\text{ST}})}{2L}\), where \(\frac{L}{K}\) is the average number of
active hops and $p_{out}^{ST}$ is the outage including the interference. To determine the optimal reuse factor, we formulate the following problem

$$\max_K \frac{r(1 - p_{out}^{ST})}{2K}$$

s.t. $K \in \mathbb{Z}$, $2 \leq K < L$ \hspace{1cm} (5.7)

Substituting (5.6) into (5.7), we have

$$\max_K \frac{r}{2K} \left[ 1 - \frac{(2^n-1)^N}{(2^n-1)^{1+1/N} + 1/N} \right]$$

s.t. $K \in \mathbb{Z}$, $2 \leq K < L$ \hspace{1cm} (5.8)

This integer programming problem is NP-hard, so we relax the integer constraint for approximate solutions. We first solve the optimization problem without the integer constraint. For a solution $K'$, we compare the performance for the two nearest integers ([$K'$] and [K']), and choose the better for $K^\dagger$. If $K^\dagger \geq L$, we apply no reuse; if $K^\dagger < 2$, we set $K^\dagger = 2$. Note that $K^\dagger$ is a local optimum since the objective is continuous; however, from the results shown in Fig. 5.3, $K^\dagger$ is also a good approximation to the global optimum. The optimal reuse factor that maximizes the throughput satisfies

$$\beta^N \left[ \frac{1 + N\beta}{N} + 1 + \frac{\alpha(N + 1)}{1 + N\beta} \right] = \left( \frac{1 + N\beta}{N} \right)^{N+1}$$

where $\beta = \frac{2^n-1}{K^{1+\alpha}}$.

5.4 Incorporation of Cooperative Beamforming

In addition to “passively” avoiding interference using a large reuse factor, cooperative beamforming can be applied to “actively” mitigate interference by utilizing cooperation among multiple transmitters. Clearly, with spatial reuse, the interference results from the simultaneously active transmissions. In order to reduce the interference, every transmitting cluster applies ZFB (Zero-Forcing Beamforming) to steer nulls to the concurrently receiving cluster-heads. In the extreme case where $N$ is sufficiently large, the interference can be completely eliminated. With a limited number of cooperating nodes, the interference coming from nearby interferers can be canceled, and
hence the interference sensed at the receiver can be reduced significantly. Although ZFB is generally suboptimal, it is easy to implement and has near-optimal performance in practice [112]. Compared with Dis-STC, however, cooperative beamforming requires CSI at the transmitters, which usually incurs a large amount of overhead and hence might degrade the performance. The impact of overhead as well as the potential overhead-performance trade-off are studied in the following sections.

5.4.1 Impact of Cooperative Beamforming

For any transmitting cluster $C_i$, the intended receiver is node $n_{i+1}^+$ and the simultaneously active receivers are those nodes that are $mK$ hops away, $m \in \{1, 2, 3, \ldots \}$. Let us focus on one concurrently active receiver, $n_{i+1+K}^+$, whose reception would be interfered by the transmission from $C_i$. If cooperative beamforming is applied, denoting the beamforming vector used at $C_i$ as $w = [w_1, \ldots, w_N]$ and the channel coefficients from $N$ transmitters to $n_{i+1+K}^+$ as $h = [h_1, \ldots, h_N]$, then, the interfering signal received at $n_{i+1+K}^+$ from $C_i$ is

$$ y = hw^T s $$

(5.10)

where $s$ is the transmitted signal from $C_i$. Note that, in order to satisfy the power constraint, $w$ needs to be a normalized vector.

Assuming perfect CSI at the transmitters, by carefully tuning the beamforming vector, the transmitting cluster with $N$ nodes could transmit to the intended receiver without interfering $2M$ concurrently active receivers as long as $2M < N$. Specifically, for transmitting cluster $C_i$ and receiver $n_{i+1+K}^+$, $w$ can be chosen such that the interfering signal $y = hw^T s = 0$. The remaining $N - 2M$ spatial degrees of freedom are utilized to enhance the transmission to the intended receiver [113], for example, using MRT (Maximum Ratio Transmission) [12].

If every transmitting cluster applies the same transmission mode to place nulls at concurrently active receivers that are up to $MK$ hops away, then, for receiver $n_{i+1}^+$, the dominant interference comes from $(M + 1)K$ hops away, which is dramatically
where we apply the approximations $b_1 \approx \frac{(2^\alpha - 1)}{(N - 2M)(MK + K - 1)^\alpha}$ and $b_2 \approx \frac{(2^\alpha - 1)}{(N - 2M)(MK + K + 1)^\alpha} = b$ in step (i) (these approximations are reasonable, especially when $M$ and/or $K$ is smaller than what would have come from $K$ hops away. Therefore, with reuse factor $K$ and interference mitigation index $M$, similar to (5.4), the per-hop outage is

$$p_{out}^{BF} = \Pr \left\{ \log \left( 1 + \frac{(N - 2M)P_r|g_1|^2d^{-\alpha}}{\eta + (|g_2|^2d^{-\alpha} + |g_3|^2d^{-\alpha}) + P_N} \right) < \eta \right\}$$

$$\approx \Pr \left\{ |g_1|^2 < \frac{(2^\alpha - 1)}{N - 2M} \left[ \frac{|g_2|^2}{MK + K - 1}\alpha \right] + \frac{|g_3|^2}{MK + K + 1}\alpha \right\} \quad \text{(5.11)}$$

where $|g_1|^2$ is the gain of the VMISO link from $C_i$ to $n_i+1$, and $|g_2|^2$ and $|g_3|^2$ are the gains for the interfering links. Because CSI is available at the transmitters, compared with Dis-STC, ZFB provides an extra array gain of $N - 2M$ [3].

Since $N - 2M$ degrees of freedom are used for MRT, $|g_1|^2 \sim \text{Erlang}(N - 2M, 1)$; for the interfering links, no coordination is applied, so $|g_2|^2 \sim \text{Exp}(1/N)$ and $|g_3|^2 \sim \text{Exp}(1/N)$. Therefore,

$$p_{out}^{BF} = \int_{x < b_1x_2 + b_2x_3} \frac{x_1^{N-2M-1}e^{-x_1} e^{-x_2} e^{-x_3}}{(N - 2M - 1)! N^2} dx_1 dx_2 dx_3 \quad \text{(5.12)}$$

where $x_m = |g_m|^2$, $m = 1, 2, 3$, $b_1 = \frac{(2^\alpha - 1)}{(N - 2M)(MK + K - 1)^\alpha}$, and $b_2 = \frac{(2^\alpha - 1)}{(N - 2M)(MK + K + 1)^\alpha}$. To simplify the analysis, we apply a similar approximation as in [111]

$$p_{out}^{BF} = \Pr \{ x_1 < b_1x_2 + b_2x_3 \}
\approx (i) \Pr \left\{ x_1 < \frac{(2^\alpha - 1)}{(N - 2M)(MK + K - 1)^\alpha}(x_2 + x_3) \right\}
= \int_{x < b_2y} \frac{x^{N-2M-1}e^{-x} ye^{-\frac{y}{\gamma}}}{(N - 2M - 1)! N^2} dx dy
= \frac{1}{N^2(N - 2M - 1)!} \int_0^\infty \int_0^{by} x^{N-2M-1}e^{-x} ye^{-\frac{y}{\gamma}} dy
\approx (ii) \frac{1}{N^2(N - 2M - 1)!} \int_0^\infty \gamma(N - 2M, by)ye^{-\frac{y}{\gamma}} dy
\equiv (iii) \frac{1}{N^2(N - 2M - 1)!} b^{N-2M} \Gamma(N - 2M + 2) \sum_{j=0}^{\infty} \frac{N - 2M + 1 + j}{N - 2M + 1} \left[ \frac{b}{b + \frac{1}{N}} \right]^{j}
= (iv) \frac{b^{N-2M}(b + 1 - \frac{2M+1}{N})}{(b + \frac{1}{N})^{N-2M+1}} \quad \text{(5.13)}
large). Note that we substitute \( x = x_1 \) and \( y = x_2 + x_3 \sim \text{Erlang}(2, \frac{1}{N}) \) in step (ii). \( \gamma(\cdot, \cdot) \) is the incomplete gamma function and \( \Gamma(\cdot) \) is the gamma function. Steps (iii) and (iv) come from [114, Chapters 6 and 8]. Note that \( \frac{b}{b+\frac{1}{N}} \in (0, 1) \), so, we can easily obtain the result in step (v).

Clearly, keeping the spectral efficiency \( \eta \) fixed, \( b = \frac{2^{\eta-1}}{(N-2M)(MK+K)^\alpha} \) decreases when \( M \) and/or \( K \) increases. From (5.13), given the fact that \( N > 2M \geq 0 \) and \( b > 0 \), \( p_{\text{out}}^{\text{BF}} \) decreases when \( b \) decreases. This means that, as expected, when \( K \) increases (less interference) and/or \( M \) increases (better interference cancellation), the outage decreases. In addition, with fixed \( M \) and \( K \), the outage decreases when \( N \) increases, indicating the benefits of applying more cooperative nodes.

### 5.4.2 Optimal Reuse Scheduling

For each cluster, the \( N \) degrees of freedom can be utilized to mitigate interference using ZFB or to improve the transmission using MRT. Obviously, the allocation of degrees of freedom affects the optimal spatial reuse scheduling. Recalling the definition of network throughput, to jointly determine the optimal reuse factor and allocation of \( N \) degrees of freedom, we formulate the following problem

\[
\begin{align*}
\max_{K,M} & \quad r \left( 1 - \frac{p_{\text{out}}^{\text{BF}}}{2K} \right) \\
\text{s.t.} & \quad K \in \mathbb{Z}, \quad 2 \leq K < L \\
& \quad M \in \mathbb{Z}, \quad 0 \leq M < \left\lfloor \frac{N-1}{2} \right\rfloor 
\end{align*}
\]

(5.14)

Substituting (5.13) into (5.14), we have

\[
\begin{align*}
\max_{K,M} & \quad r \left\{ 1 - \frac{2^{\eta-1}}{(N-2M)(MK+K)^\alpha} \left[ \frac{2^{\eta-1}}{(N-2M)(MK+K)^\alpha} + 1 - \frac{2M-1}{N} \right] \right\} \\
\text{s.t.} & \quad K \in \mathbb{Z}, \quad 2 \leq K < L \\
& \quad M \in \mathbb{Z}, \quad 0 \leq M \leq \left\lfloor \frac{N-1}{2} \right\rfloor 
\end{align*}
\]

(5.15)
Note that, with fixed $M$, (5.15) degenerates into a similar problem as (5.8), which can be solved via the same procedure discussed in Section 5.3 and [111]. Therefore, the optimal $K$ for fixed $M$ satisfies

$$\left( \frac{\beta_1}{K^\alpha} \right)^{N-2M} \left\{ \beta_1^2 + K^\alpha \left[ \frac{K^\alpha \beta_2 (\alpha \beta_2 + \alpha + 1)}{N^2} - \frac{\beta_1 \beta_2 (\alpha - 1)}{N} + \frac{\beta_1 (\alpha \beta_2 + 1)}{N} \right] \right\}$$

$$= \left( \frac{\beta_1}{K^\alpha} + \frac{1}{N} \right)^{N-2M} \left[ K^\alpha \left( \frac{K^\alpha}{N^2} + \frac{2 \beta_1}{N} \right) + \beta_1^2 \right]$$

(5.16)

where $\beta_1 = \frac{2^{\eta-1}}{(N-2M)(M+1)^\eta}$ and $\beta_2 = N - 2M - 1$. After calculating the optimal $K$ without the integer constraint, we compare the performance for the two nearest integers and choose the better one. The edge cases are handled in the same manner as that in Section 5.3.3.

Since the number of possible values for $M$ is $\lfloor \frac{N-1}{2} \rfloor + 1$, which is usually relatively small, we can iterate through all possible values for $M$, and, for each fixed $M$, apply (5.16) to calculate the optimal $K$ for this particular allocation of degrees of freedom. Finally, the optimal $M$ and $K$ that maximize the network throughput can be computed.

### 5.4.3 Impact of Channel Estimation Overhead and Imperfect CSI

Different from applying Dis-STC, which does not require CSI at the transmitters, using ZFB requires the transmitters to have accurate CSI to the nodes at which nulls are to be placed. Although a larger $M$ results in less interference (more interference is canceled), it incurs a larger amount of overhead. This means that there is a trade-off between the interference level and the required overhead. In this section, we investigate the impact of overhead on determining the optimal $M$ that maximizes the effective network throughput.

In general cases, a training sequence is used to estimate the channel, and only the estimated channel coefficient, $\hat{h}$, is available. Assuming $h$ and $\hat{h}$ are jointly ergodic and stationary Gaussian processes, we can write

$$h = \hat{h} + e$$

(5.17)
where \( h \) is the actual channel coefficient and \( e \) is the estimation error, usually modeled as complex Gaussian with zero mean and variance \( \sigma_e^2 \) \([80]–[99]\). In [80], the authors show that, assuming a linear minimum mean-square-error estimator, \( \sigma_e^2 = \frac{1}{1 + \bar{\gamma} N_t} \), where \( N_t \) is the number of training symbols and \( \bar{\gamma} \) is the average SNR for the estimated channel.

Note that the beamforming vector \( \mathbf{w} \) in (5.10) is calculated based on \( \hat{h} \). This means that the interference cannot be perfectly canceled, and the residual interference at node \( n_{i+1+K} \) from transmitting cluster \( C_i \) is

\[
\mathbf{hws} = (\hat{h} + e)\hat{w}s = e\hat{w}s
\]

where \( e \) is the estimation error vector, the elements of which are independently identically distributed. Because \( \mathbf{w} \) is a normalized vector and \( e \) is independent from \( \hat{w} \), \( e\hat{w} \sim CN(0, \sigma_e^2) \). Then, the power of the residual interference at node \( n_{i+1+K} \) from transmitting cluster \( C_i \) follows an exponential distribution with mean \( \sigma_e^2 \).

In order to obtain the required CSI at the transmitters, making use of the broadcast nature of wireless transmissions and assuming the channels are reciprocal, we let every cluster-head take turns broadcasting training sequences. When cluster-head \( n_i \) is broadcasting its training sequence, all the other nodes in this cooperative network can estimate the channel from itself to node \( n_i \) (channels are reciprocal). Since the concurrently active receivers are cluster-heads, after the training process, every transmitting cluster has access to all the required CSI to put nulls at the desired nodes, i.e., the cluster-heads that are \( mK, m \in \{1, 2, \cdots, M\} \) hops away. If this training process is performed every \( T \) seconds, the ratio of the effective transmission time (total transmission time minus the channel estimation time consumption) would be \( 1 - \frac{NT_i}{T} \), where \( T_i \) is the time consumed in transmitting one training sequence. We assume that all the cluster-heads broadcast training sequences with the same length, using the same transmitting power, \( P_T \).

Taking into account the impact of imperfect CSI and the overhead incurred in
channel estimation, the effective network throughput becomes

$$
\left(1 - \frac{NT_t}{T}\right) r \left(1 - \frac{\hat{p}^{BF}}{2K}\right)
$$

(5.19)

where \(\hat{p}^{BF}\), the outage with imperfect CSI, is

$$
\hat{p}^{BF} = \Pr\left\{ \log \left( 1 + \frac{(N - 2M)P_T|g_1|^2d^{-\alpha}}{\sum_{m=1}^{M} \frac{P_T|g_N|^2d^{-\alpha}}{(mK+1)d^\alpha} + \frac{P_T|g_R|^2d^{-\alpha}}{(M+1)Kd^\alpha} + \frac{P_T|g_3|^2d^{-\alpha}}{[(M+1)K+1]d^\alpha} + P_N} \right) < \eta \right\}
$$

(5.20)

Comparing (5.20) with (5.11), the new interference terms represent the sum of all the residual interferences. Note that there are two interferers that are \(mK, m \in \{1, 2, \cdots, M\}\) hops away: one is to the left and \(\frac{P_T|g_N|^2}{(mK+1)d^\alpha}\) is the power of the residual interference, where \(|g_N|^2 \sim \text{Exp} \left(1 + \frac{P_TN_t}{[(mK+1)d^\alpha]} \right)\); the other is to the right and \(\frac{P_T|g_R|^2}{(mK-1)d^\alpha}\) is the residual interference power, where \(|g_R|^2 \sim \text{Exp} \left(1 + \frac{P_TN_t}{[(mK-1)d^\alpha]} \right)\).

Intuitively, the value of \(T\), which indicates how often the training is required, depends on how rapid the channel changes, which is usually a known system parameter (on average). Assuming the same modulation and coding for data transmission and channel estimation, the ratio \(\frac{T_t}{T}\) is equivalent to \(\frac{N_t}{N_c}\), where \(N_c\) is the total transmission time (during which one training process is needed) measured in number of symbols.

Note that the relationship between these two ratios can be easily adjusted for the cases where different transmission modes are used for data transmission and channel estimation. On the one hand, a larger \(N_t\) means more overhead that degrades the effective throughput. On the other hand, with a larger \(N_t\), the CSI is more accurate and hence the residual interference is smaller, which has the potential to improve the throughput. Therefore, \(N_t\) is yet another decision variable we need to optimize in maximizing the effective network throughput.
Considering the impact of channel estimation and imperfect CSI, the optimization problem becomes

$$\max_{K,M,N_t} \left( 1 - N \frac{N_t}{N_c} \right) \frac{r}{2K} \left( 1 - \hat{p}_{\text{out}}^{BF} \right)$$

s.t. $K \in \mathbb{Z}$, $2 \leq K < L$

$$M \in \mathbb{Z}, \quad 0 \leq M < \left\lfloor \frac{N - 1}{2} \right\rfloor$$

$$N_t \in \mathbb{Z}, \quad 0 \leq N_t \leq \left\lfloor \frac{N_c}{N} \right\rfloor$$

We solve this problem numerically and present the results in Section 5.5.

5.5 Performance Evaluation and Discussion

Considering a multi-hop linear network with $L = 15$ hops, we assume the transmit power at each node is 20 dBm, the noise power spectral density is $-174$ dBm/Hz, the bandwidth $B = 10$ MHz, the per-hop distance $d = 500$ m, and the desired per-hop efficiency is 6 bits/s/Hz.

5.5.1 Impact of Cooperation

The impact of reuse on the network throughput with different degrees of cooperation is presented in Fig. 5.3. Even with the efficiency loss from the local message dissemination, cooperation significantly improves the throughput. The optimal throughput with $N = 12$ is more than 5 times that without cooperation ($N = 1$). Note that whether cooperation is beneficial depends on transmission power, transmitter-receiver separation, and targeted SNR, as studied in [50,111].

In addition, we observe that, given $N$, there is a reuse factor that maximizes the throughput; the optimal values calculated using (5.9) are denoted in Fig. 5.3 using stars. For example, if $N = 12$, the optimal throughput is achieved when $K = 5$. The optimal reuse factor balances the channel utilization and the interference level. We can also see that (5.9) provides an effective guideline for choosing the reuse factor.
In the previous discussions, we assumed that the per-hop separation is identical and the cooperating nodes are located close enough to the cluster-head that the distances between the transmitting nodes and the receiver are the same. Although this assumption has been shown to be reasonable in practical networks [43], it is instructive to investigate the validity of our conclusions when the nodes are randomly distributed within the cluster. Keeping the other system settings the same, we assume that every cluster covers a circle with radius $d_r$, and all the nodes belonging to the cluster are randomly distributed within the circle. Note that, under this assumption, the SISO route comprised with cluster-heads is no longer a linear network and the per-hop distance is not identical or fixed, but a random value in $[d - 2d_r, d + 2d_r]$.

The performance with randomly distributed nodes is provided in Fig. 5.4. We observe that the impact of randomly distributed relays is small when $N$ is large. For

Figure 5.3: Normalized network throughput as a function of spatial reuse factor (DisSTC is used)
example, when $N = 12$, the optimal spatial reuse factor stays the same even when $d_r = 250$ m, which is the maximum permitted $d_r$ without overlapping clusters. However, when $N = 1$, both the throughput and the optimal reuse change dramatically. Interestingly, the randomness in the locations of the nodes improves the highest achievable throughput when $N = 1$, but slightly degrades the performance when $N = 4$ and $N = 12$. The reason for this is that the SISO link is weak, that is, the per-hop outage when $N = 1$ is relatively high; hence the randomness in the per-hop distance increases the probability of successful transmissions. On the contrary, when $N$ is large, the outage is small because of the cooperative diversity; in this case, the random per-hop separation degrades the performance. In general, the size of the cluster is relatively small compared with the distance among clusters (it is true especially in dense networks). In this case, e.g., $d_r/d \leq 20\%$, the impact of randomly distributed nodes is
small, and the previously derived optimal reuse works well.

5.5.2 Impact of Interference Mitigation

From Fig. 5.3, although there are plenty of cooperative nodes when $N = 12$, the optimal reuse factor is still 5, which means that the network throughput is at most $1/5$ of the per-hop throughput. This suggests that a cooperative network applying Dis-STC is interference-limited. In order to further improve the throughput, as discussed in Section 5.4.1, we apply ZFB to actively mitigate interference. Recalling that $M$ is the interference mitigation index, indicating that the interferences coming from $mK, m \in \{1, 2, \cdots, M\}$ hops away are canceled, the performance of ZFB is presented in Fig. 5.5.

![Normalized network throughput as a function of spatial reuse factor](image)

**Figure 5.5:** Normalized network throughput as a function of spatial reuse factor

A significant performance gain can be achieved when ZFB is used, especially
when \( N \) is large, that is, when there are many cooperating nodes. Also, a much smaller spatial reuse factor is needed to maximize the network throughput. For example, using ZFB, when \( N = 12 \), the optimal reuse factor is 2, which is the smallest available. The reason for this is that, by allocating some degrees of freedom to cancel interference, the SIR (Signal-to-Interference Ratio) can be dramatically improved. As shown in Fig. 5.6, when \( N = 12 \), the optimal \( M \) is 2; this means that, for each transmitting cluster \( C_i \), the precoding vector is chosen such that the interference coming from \( C_i \) at \( n_{i-1} \) (the active receiver for the concurrently transmitting cluster that is 2 hops to the left), \( n_{i+3} \) (2 hops to the right), \( n_{i-3} \) (4 hops to the left), and \( n_{i+5} \) (4 hops to the right) is eliminated. The remaining degrees of freedom are then dedicated to improving the transmission to the intended receiver. Similar to the results in Fig. 5.3, the derived optimal \( K \) and \( M \) using the algorithm described in Section 5.4.2 match the simulation results well.

### 5.5.3 Impact of Imperfect CSI

We have shown that ZFB dramatically improves the throughput by mitigating interference. However, implementing ZFB requires accurate CSI at the transmitters, which incurs a significant amount of overhead. Using the channel estimation model described in Section 5.4.3, the impact of channel estimation and imperfect CSI is presented in Fig. 5.7. We consider a wireless system with moderate mobility such that the coherence time is on the order of 100 msec [83], and the end-to-end transmission duration is set as 10 msec. The symbol length is assumed to be 0.01 msec, which implies that \( N_c = 1000 \), that is, channel estimation is performed every 1000\( T_s \), where \( T_s \) is the transmission time for one symbol. We assume that \( N_t \in [1,10] \), that is, 1 to 10 training symbols are used per training sequence. We find that the optimal value of \( N_t \) that maximizes the network throughput is almost always 1, indicating that, the impact of channel estimation overhead is critical within the current simulation settings.

According to Fig. 5.7, imperfect CSI has a critical impact on the performance of
ZFB. Because of the degraded SNR caused by channel estimation error and, more importantly, the residual interference, the throughput decreases significantly. In addition, the optimal reuse factor increases. For example, with imperfect CSI, when $N = 12$, the optimal reuse factor increases to 3 from 2 with perfect CSI. This is because that the interference coming from nearby interferers cannot be canceled completely, so a more conservative spatial reuse is required to reduce the interference. However, even with channel estimation overhead and imperfect CSI, ZFB still outperforms STC; so, applying ZFB is usually a better approach for improving the network throughput, especially when $N$ is large, that is, the cooperating cluster has a large number of nodes.

**Figure 5.6:** Normalized network throughput as a function of $M$ ($N = 12$, ZFB is used)
Figure 5.7: Normalized network throughput as a function of spatial reuse factor

5.6 Conclusions

In this chapter, we investigated spatial reuse scheduling in multi-hop linear cooperative networks. The impact of enabling spatial reuse and incorporating cooperation on the network throughput was addressed through analyses and simulations. We showed that (i) appropriate reuse is essential for efficient networks, (ii) incorporating cooperation is effective in improving the network performance when the SISO links are suffering high outages, and (iii), by efficiently allocating degrees of freedom, beamforming performs better than Dis-STC in improving the throughput. In addition, the reuse factor that maximizes the network throughput was derived, the performance of which was illustrated in the simulation results. Good approximations were also proposed to serve as guidelines for determining appropriate spatial reuse.

Linear networks are good models to begin with, but more practical network
models (e.g., grid networks) will be considered in following research. Furthermore, the practical implementation of cooperation and spatial reuse in wireless networks is an interesting and important topic, in particular, the trade-off between the overhead incurred in practical implementations and the performance improvement needs to be quantified.
Chapter 6
CONTRIBUTIONS AND FUTURE WORK

In this chapter, we summarize the contributions of this dissertation, and then discuss potential research topics for future work.

6.1 Contributions

6.1.1 Overhead-Performance Trade-off

In order to evaluate the real benefits of implementing cooperative communications in wireless networks, we investigated the overhead-performance trade-off, focusing on cooperative relaying. In Chapters 2 and 3, we first identified the components of the overhead and then quantified the impact of overhead by evaluating the spectral/energy-efficiency of different cooperative relaying schemes. Through analyses and simulations, we demonstrated that overhead plays an important role in determining the achievable network performance, and, hence, understanding the overhead-performance trade-off is critical in designing efficient networks.

More importantly, we provided useful guidelines in determining the appropriate cooperative relaying scheme for specific applications. In general, Dis-STC is not practical due to the centralized implementation that incurs significant overhead; $M$-group is preferable when the network is large and/or the SNR margin is large, while, TBBS is better in small networks.

6.1.2 Cooperative Routing with Minimum Overhead

The existing cooperative routing schemes rely on an initial SISO route. This means that, in some circumstances, even though cooperative diversity might help overcome the “voids” (in the eyes of the SISO routing schemes), cooperative routing cannot
establish a reliable route due to the fact that no SISO route can be discovered in the first place. Clearly, this limits the benefits we can achieve in using cooperative communications. Therefore, we proposed novel cooperative routing algorithms that build the cooperative route on the fly, and, hence, significantly improve the probability of successfully discovering a route.

Compared with existing schemes, our proposed algorithms perform much better, providing a higher probability to discover a route and a smaller end-to-end delay as well as a higher throughput; these advantages are particularly apparent when the network is sparse. In addition, a smaller amount of overhead is required because (1) $M$-group is applied to reduce the amount of overhead incurred in cooperation, and (2) a one-step algorithm simplifies the process of discovering and establishing cooperative routes.

6.1.3 Optimal Spatial Reuse for Wireless Networks

Spatial reuse scheduling is critical for efficient wireless networks. Focusing on multi-hop cooperative networks, we investigated the impact of applying cooperative communications in multi-hop linear networks and derived the reuse scheduling that maximizes the network throughput. We demonstrated, through analyses and simulations, that cooperation with the appropriate reuse can dramatically improve the network throughput.

In addition, motivated by the observation that, even with an optimal spatial reuse factor, the network is still interference-limited, we applied advanced cooperative communication techniques to actively mitigate the interference. Specifically, cooperative beamforming was incorporated and the joint optimization of the spatial reuse factor and the allocation of the spatial degrees of freedom was formulated and solved. We provided simulation results to illustrate the performance improvement provided by applying cooperative beamforming.
6.2 Future Work

6.2.1 Overhead-Performance Trade-off Analysis: Information Theoretic Approach

In our analyses on overhead-performance trade-off, usually, an application-specific approach was taken. For example, in our work on the spectral efficiency of cooperative relaying, we quantified the overhead incurred in implementing cooperative relaying schemes and studied the impact of overhead on the system efficiency. This kind of analysis provides useful guidelines; however, these conclusions are generally limited to the particular systems under consideration. It is easy to see, however, that the overhead-performance trade-off has a much broader implication.

Therefore, there has been some work attempting to reveal the fundamental overhead-performance trade-off from an information-theoretic angle [120–127]. For example, rate-distortion theory [121] has been used to model the overhead-performance trade-off; this idea dates back to Gallager’s paper [120]. In [126], the authors apply rate-distortion theory to study how the knowledge of the traffic information impacts the network throughput.

This information-theoretic approach is promising since it has the potential to obtain fundamental insights on the general overhead-performance trade-off. However, it is challenging to develop a general framework that abstracts the essential fundamentals, without being too specific about the applications. Based on our recent results [128], we will continue working along this direction.

6.2.2 Efficient Heterogeneous Cooperative Networks

To accommodate the rapidly growing data traffic in cellular networks, heterogeneous cellular networks have been introduced as a key feature of the LTE-Advanced standard [129], providing the next performance and capacity leap. Alternatively, offloading some of the data traffic to WiFi networks can also partially relieve the burden on cellular networks [130, 131]. Deploying such a mix of traditional macrocells, small
cells (e.g., femtocells [132] and relay stations [133]), and WiFi hot-spots has the potential to provide a consistent broadband experience of higher data rates and lower latency to users anywhere in the network. However, the inherent complexity of heterogeneous wireless networks (HetNets) imposes a number of new challenges in network design and management.

In our recent papers [134, 135], we investigated energy efficient user pairing (cell association) for HetNets. Focusing on energy efficiency, we formulated the scheduling as an optimization problem, and then, proposed an efficient algorithm that combines cell association and cooperative communications. Through simulations, we demonstrated the benefits of integrating peer-to-peer networks, applying efficient cell association, and utilizing cooperation. There are a number of possible extensions to our preliminary work, which include taking into account diverse means of direct communications among mobile users (e.g., device-to-device communications) and more realistic system settings (e.g., constrained backhaul capacity). In addition, the evaluation and comparison of a centralized approach, as well as distributed algorithms, are also of great practical interest.
BIBLIOGRAPHY


107


