Fundamental Limits of Wireless Two-Way Full-Duplex Communication Networks

BY

ZHIYU CHENG
B.S. East China University of Science and Technology, 2009
M.S. University of Illinois at Chicago, 2012

THESIS
Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Chicago, 2014

Chicago, Illinois

Defense Committee:

Natasha Devroye, Chair and Advisor
Daniela Tuninetti
Rashid Ansari
Milos Zefran
Emina Soljanin, Bell Laboratories
To my parents,

Lin Fang and Mingyuan Cheng.
ACKNOWLEDGMENTS

I am indebted to my advisor, Professor Natasha Devroye, for her encouragement, invaluable advice, and support during my Ph.D. study. Natasha is a rising star of great scholar that would light up my career and set an example for me.

I am very grateful to my motherland, China, in particular, for her support during my Ph.D. study in the US.

I would like to thank Dr. Daniela Tuninetti, Dr. Rashid Ansari, Dr. Milos Zefran from UIC and Dr. Emina Soljanin from Bell Labs for serving on my thesis committee.

Thanks to all my friends both in Chicago and elsewhere who made the past 4.5 years unforgettable and full of fun.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Rethinking the Two-way Channel: Adaptation/Interaction</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Two-way Networks</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Full-duplex Wireless, Degrees of Freedom, and Relaying</td>
<td>11</td>
</tr>
<tr>
<td>1.4 Outline</td>
<td>17</td>
</tr>
<tr>
<td>2 WIRELESS DETERMINISTIC TWO-WAY NETWORKS</td>
<td>19</td>
</tr>
<tr>
<td>2.1 Models, Definitions and Notations</td>
<td>19</td>
</tr>
<tr>
<td>2.2 Two-way Multiple-Access Broadcast Channel</td>
<td>23</td>
</tr>
<tr>
<td>2.2.1 An Introductory Example: Modulo 2 Adder MAC/BC</td>
<td>23</td>
</tr>
<tr>
<td>2.2.2 A More General Model for Deterministic MAC/BC</td>
<td>25</td>
</tr>
<tr>
<td>2.2.3 Linear Deterministic MAC/BC</td>
<td>29</td>
</tr>
<tr>
<td>2.3 Two-way Z Channels</td>
<td>31</td>
</tr>
<tr>
<td>2.3.1 An Introductory Example: Modulo 2 Adder Two-way Z Channel</td>
<td>31</td>
</tr>
<tr>
<td>2.3.2 Comments on a More General Model for the Two-way Z Channel</td>
<td>32</td>
</tr>
<tr>
<td>2.3.3 Linear Deterministic Two-way Z Channel</td>
<td>33</td>
</tr>
<tr>
<td>2.4 Two-way Interference Channels</td>
<td>34</td>
</tr>
<tr>
<td>2.4.1 An Introductory Example: Modulo 2 Adder Two-way IC</td>
<td>35</td>
</tr>
<tr>
<td>2.4.2 A More General Class of Two-way Deterministic ICs</td>
<td>36</td>
</tr>
<tr>
<td>2.4.3 Linear Deterministic Two-way IC</td>
<td>39</td>
</tr>
<tr>
<td>2.4.4 Symmetric Rate Comparison with Other IC Models</td>
<td>47</td>
</tr>
<tr>
<td>2.5 Summary</td>
<td>48</td>
</tr>
<tr>
<td>3 WIRELESS GAUSSIAN TWO-WAY NETWORKS</td>
<td>50</td>
</tr>
<tr>
<td>3.1 Gaussian Two-way Multiple-Access Broadcast Channel</td>
<td>50</td>
</tr>
<tr>
<td>3.1.1 Channel Model</td>
<td>50</td>
</tr>
<tr>
<td>3.1.2 The Limited Utility of Adaptation in the Gaussian Two-way MAC/BC</td>
<td>51</td>
</tr>
<tr>
<td>3.2 Gaussian Two-way Interference Channel</td>
<td>57</td>
</tr>
<tr>
<td>3.2.1 Channel Model, Definitions, and Partial Adaptation Lemma</td>
<td>58</td>
</tr>
<tr>
<td>3.2.2 Outer bounds</td>
<td>61</td>
</tr>
<tr>
<td>3.2.3 Very Strong Interference: INR ≥ SNR(1 + SNR)</td>
<td>69</td>
</tr>
<tr>
<td>3.2.4 Strong Interference: SNR ≤ INR ≤ SNR(1 + SNR)</td>
<td>70</td>
</tr>
<tr>
<td>3.2.5 Weak Interference: INR ≤ SNR</td>
<td>72</td>
</tr>
<tr>
<td>3.2.5.1 INR ≥ 1</td>
<td>73</td>
</tr>
<tr>
<td>3.2.5.2 INR &lt; 1</td>
<td>75</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.6 Final Comments on Adaptation versus No-adaptation, and</td>
<td>76</td>
</tr>
<tr>
<td>versus Perfect Output Feedback</td>
<td></td>
</tr>
<tr>
<td>3.3 K-pair-user Gaussian Two-way Interference Channel</td>
<td>79</td>
</tr>
<tr>
<td>3.3.1 System Model</td>
<td>79</td>
</tr>
<tr>
<td>3.3.2 Outer bounds and capacity/gap results</td>
<td>82</td>
</tr>
<tr>
<td>3.3.2.1 Capacity result for the linear deterministic model</td>
<td>82</td>
</tr>
<tr>
<td>3.3.2.2 Constant gap result for Gaussian model</td>
<td>85</td>
</tr>
<tr>
<td>3.4 Summary</td>
<td>89</td>
</tr>
<tr>
<td>4 DEGREES OF FREEDOM OF WIRELESS TWO-WAY FULL-DUPLEX NETWORKS</td>
<td>91</td>
</tr>
<tr>
<td>4.1 System Model</td>
<td>91</td>
</tr>
<tr>
<td>4.1.1 K-pair-user Full-duplex Two-Way IC</td>
<td>91</td>
</tr>
<tr>
<td>4.1.2 K-pair-user Full-duplex Two-way IC with a MIMO Relay</td>
<td>94</td>
</tr>
<tr>
<td>4.1.3 Types of Signals</td>
<td>97</td>
</tr>
<tr>
<td>4.2 DoF of K-pair-user Full-duplex Two-way IC</td>
<td>98</td>
</tr>
<tr>
<td>4.3 DoF of K-pair-user Full-duplex Two-way IC with an Instantaneous MIMO</td>
<td>104</td>
</tr>
<tr>
<td>4.3.1 Achieving the Maximal DoF</td>
<td>104</td>
</tr>
<tr>
<td>4.3.2 Comments on Reducing the Number of Antennas at the Instantaneous</td>
<td>109</td>
</tr>
<tr>
<td>4.4 Summary</td>
<td>113</td>
</tr>
<tr>
<td>4.5 Summary</td>
<td>121</td>
</tr>
<tr>
<td>5 CONCLUSION AND FUTURE WORK</td>
<td>123</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>126</td>
</tr>
<tr>
<td>Appendix A</td>
<td>127</td>
</tr>
<tr>
<td>Appendix B</td>
<td>129</td>
</tr>
<tr>
<td>Appendix C</td>
<td>130</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>132</td>
</tr>
<tr>
<td>VITA</td>
<td>139</td>
</tr>
<tr>
<td>TABLE</td>
<td>CONSTANT GAPS BETWEEN NON-ADAPTIVE SYMMETRIC HAN AND KOBAYASHI SCHEMES IN EACH DIRECTION AND PARTIALLY OR FULLY ADAPTIVE OUTER BOUNDS FOR THE SUM-RATE OF THE SYMMETRIC TWO-WAY GAUSSIAN IC [1].</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>FIGURE</td>
<td>PAGE</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
</tr>
<tr>
<td>8</td>
<td>95</td>
</tr>
<tr>
<td>9</td>
<td>115</td>
</tr>
<tr>
<td>10</td>
<td>128</td>
</tr>
</tbody>
</table>

Examples of point-to-point two-way channels breaking up into two parallel one-way channels [1].

Three multi-user two-way channel models and two of the classes under consideration [1].

Time-sharing based achievability for the proof of Theorem 2 [1].

$C_{\text{sym}}$ for various linear deterministic ICs as a function of $\alpha := \frac{q}{p}$; $q$ interfering link strength, $p$ direct link strength [1].

Example of a channel in which adaptation yields unbounded gain over non-adaptation. The non-adaptive scheme would not be able to achieve any non-trivial rates for $R_{12}$ and $R_{34}$, while the adaptive scheme would be able to achieve strictly positive rates [1].

Example of a channel in which perfect output feedback (denoted by the dashed arrows in the left figure) would be able to achieve strictly positive rates for $R_{12}$ and $R_{34}$, but the adaptive scheme for the same channel on the right would not be able to achieve non-zero rates for any of the links [1].

K-pair-user two-way interference channel. $M_{jk}$ denotes the message known at node $j$ and desired at node $k$; $\hat{M}_{jk}$ denotes that $k$ would like to decode the message $M_{jk}$ from node $j$ [2].

K-pair-user two-way interference channel with a MIMO relay. $M_{ij}$ denotes the message known at node $i$ and desired at node $j$; $\hat{M}_{ij}$ denotes that $j$ would like to decode the message $M_{ij}$ from node $i$ [3].

Transformation of the K-pair-user full-duplex two-way interference channel with a causal MIMO relay [3].

The Markov chain used in the outer bound proof of Theorem 4 [1].
SUMMARY

This thesis includes previous publications in IEEE [1-8]. This thesis studies wireless two-way full-duplex, interactive communication networks from an information theoretic perspective. Most wireless communication networks are two-way, where nodes act as both sources and destinations of messages. This allows for “adaptation” at or “interaction” between the nodes – a node’s channel inputs may be functions of its message(s) and previously received signals allowing for potentially larger rates than those achievable in feedback-free one-way channels where inputs are functions of messages only. However, examples exist of channels where adaptation is not beneficial from a capacity perspective; we ask whether analogous results hold for several multi-user two-way networks. We first consider deterministic two-way channel models: the binary modulo-2 addition channel and a generalization of this, and the linear deterministic channel which models Gaussian channels at high SNR. For these deterministic models we obtain the capacity region for the two-way multiple access/broadcast channel, the two-way Z channel and the two-way interference channel (under certain “partial” adaptation constraints in some regimes). We permit all nodes to adapt their channel inputs to past outputs (except for portions of the linear high-SNR two-way interference channel where we only permit 2 of the 4 nodes to fully adapt). However, we show that the two-way fully or partially adaptive capacity region consists of two parallel “one-way” regions operating simultaneously in opposite directions, i.e. adaptation is useless.
SUMMARY (Continued)

We next consider two noisy channel models: first, the Gaussian two-way MAC/BC, where we show that adaptation can at most increase the sum-rate by $\frac{1}{2}$ bit in each direction. Next, for the two-way interference channel, partial adaptation is shown to be useless when the interference is very strong. In the strong and weak interference regimes, we show that the non-adaptive Han and Kobayashi scheme utilized in parallel in both directions achieves to within a constant gap for the symmetric rate of the fully (for some regimes) or partially (for the remaining regimes) adaptive models. Then we generalize the two-way interference channel to the K-pair-user two-way interference channel (TWIC) and show that for symmetric scenarios and certain interference regimes, non-interactive schemes again achieve to within a constant gap for the fully adaptive Gaussian model.

Furthermore, we investigate the degrees of freedom (DoF, also known as the multiplexing gain) of the K-pair-user TWIC with and without a MIMO relay, where we emphasize all nodes operate in full-duplex mode. We first derive a new outer bound (allows interaction) to demonstrate that the optimal DoF of the K-pair-user TWIC is K: full-duplex operation doubles the DoF, but interaction does not further increase the DoF. We next employ a MIMO relay in the K-pair-user TWIC. If the relay is non-causal/instantaneous (at time k forwards a function of its received signals up to time k) and has 2K antennas, we demonstrate a one-shot scheme where the relay mitigates all interference to achieve the interference-free 2K DoF. In contrast, if the relay is causal (at time k forwards a function of its received signals up to time k – 1), we show that a full-duplex MIMO relay cannot increase the DoF of the K-pair-user TWIC beyond K, as
SUMMARY (Continued)

if no relay or interaction is present. We comment on reducing the number of antennas at the instantaneous relay.
CHAPTER 1

INTRODUCTION

Two-way communication, where users A and B wish to exchange a stream of information, is a natural form of communication of relevance in present and future wireless networks. Applications include two-way high data-rate tele-medicine over wireless broadband links, mobile video conferencing over next generation cellular networks, the synchronization of data among terminals, and communication between a base station and clients. Indeed, much of our current wireless communication is already two-way in nature, but it is not treated as such in practice. Rather, current channel coding schemes orthogonalize the two directions, rendering the two-way channel equivalent to two one-way communication links. While this is simple to implement, whether such non-adaptive two-way coding schemes are optimal from a capacity perspective remains an open question.

To assess the potential gains of two-way communications we will use information theory, which quantifies the fundamental limits of data compression and transmission, and was pioneered by Claude Shannon in Bell Labs decades ago. Shannon’s fundamental theorems on point-to-point communication [9] initiated the first golden age of information theory from 1948 to 1960s. After that, information theory somehow faded, since the theorems developed were far too complicated to implement at that time. However, with the birth of seminal papers on multi-user channels [10–13], the second golden age of information theory (known as network information theory) emerged from the mid 1970s to the early 1980s with many new results and
techniques [14]. Unfortunately, many problems, including Shannon’s two-way channel [15] (the first paper on network information theory, i.e. communication between more than one sender and receiver), remained open and there was little interest in these problems from communication theorists or practitioners. In the mid 1990s, with the increasing importance and dependence on the internet and wireless communications, network information theory which aims to establish the fundamental limits on information flow in networks and the optimal coding schemes that achieve these limits, again drew much attention. However, although the study of network information theory was revived, little work has since been done on the two-way communication problem, which seems to be one sticking point towards making progress towards the capacity of networks.

Two-way communications are often practically motivated by new-found possibilities in full-duplex wireless in which users can transmit and receive signals at the same time and same frequency band. Full-duplex communications are considered by some to be one of the key candidate technologies for increasing spectral efficiency in next generation wireless communications, and has drawn much recent interest among researchers in both academia and industry [16].

Full-duplex operation enables true two-way communications over the practically relevant Gaussian noise channels, and current work on such bidirectional communications has been focused on point-to-point channels. We will study the fundamental limits of wireless two-way full-duplex communications from a multi-user or network setting. In particular, we are interested in finding the capacity (region), i.e. the technology independent highest data rate that can be reliably transmitted at in a communication channel (network). This capacity
(region) or theoretical upper bound on the reliable (arbitrarily small error probability) data rate at which one may communicate may provide guidance in the design of communication systems.

1.1 Rethinking the Two-way Channel: Adaptation/Interaction

What makes two-way communications, in which two (or more) users exchange messages over the same shared channel, challenging are the possibilities that stem from having nodes be both sources and destinations of messages. This permits them to adapt their channel inputs to their past received signals. Such two-way adaptation was first considered in the point-to-point two-way channel by Shannon [15]. Shannon’s inner and outer bounds [15] are not tight in general, and a general computable\(^1\) formula for the capacity region of the point-to-point two-way channel remains open.

However, encouragingly, the capacity region – the fundamental information theoretic limit we will be interested in – is known for several point-to-point two-way channel models where the interaction between one’s own signal and that of the other user may be resolved. For example, in the two-way modulo 2 binary adder channel where channel outputs \(Y_1 = Y_2 = X_1 \oplus X_2\) for binary inputs \(X_1, X_2\) and \(\oplus\) modulo 2 addition, the capacity region is one bit per user per channel use. Each user is able to “undo” the effect of the other as shown in Figure 1 (a), something that is not possible in one channel use for the elusive binary multiplier channel with

\(^1\)By computable we mean single-letter expression without the use of unbounded cardinality auxiliary random variables. Multi-letter formulas for the capacity of two-way channels exist, see the expressions involving directed information over code-trees of [17].
Point-to-point two-way channel

\[ Y_1 = aX_1 + bX_2 + N_1 \]
\[ Y_2 = cX_1 + dX_2 + N_2 \]

Channel

\[ \begin{array}{c}
X_1(M_{12}) \\
\downarrow
\end{array} \quad \begin{array}{c}
Y_1 \\
\downarrow
\end{array} \quad \begin{array}{c}
Y_2 \\
\uparrow
\end{array} \quad \begin{array}{c}
X_2(M_{21})
\end{array} \]

\( a, b, c, d \in \mathbb{R} \)
\( N_1, N_2 \sim \mathcal{N}(0, 1) \)

\[ E[|X_1|^2] \leq P_1, E[|X_2|^2] \leq P_2 \]

Capacity region

\( R_1 \leq 1 \)
\( R_2 \leq 1 \)

(a)

Modulo 2 adder two-way channel

\[ X_1, X_2, Y_1, Y_2 \in \{0, 1\} \]
\[ Y_1 = Y_2 = X_1 \oplus X_2 \]
equivalent to
\[ Y_1 = X_1 \oplus X_2 \oplus X_1 = X_2 \]
\[ Y_2 = X_1 \oplus X_2 \oplus X_2 = X_1 \]

Gaussian two-way channel

\[ Y_1 = aX_1 + bX_2 + N_1 \]
\[ Y_2 = cX_1 + dX_2 + N_2 \]
\( a, b, c, d \in \mathbb{R} \)
\( N_1, N_2 \sim \mathcal{N}(0, 1) \)
\[ E[|X_1|^2] \leq P_1, E[|X_2|^2] \leq P_2 \]

Capacity region

\( R_1 \leq \log(1 + c^2 P_1) \)
\( R_2 \leq \log(1 + b^2 P_2) \)

(b)

Two parallel one-way channels

\[ \begin{array}{c}
X_1(M_{12}) \\
\downarrow
\end{array} \quad \begin{array}{c}
Y_1 \\
\downarrow
\end{array} \quad \begin{array}{c}
X_2(M_{21})
\end{array} \]

in parallel with

\[ \begin{array}{c}
\bar{Y}_2 \\
\uparrow
\end{array} \quad \begin{array}{c}
\bar{Y}_1 \\
\downarrow
\end{array} \quad \begin{array}{c}
\bar{Y}_2 \\
\uparrow
\end{array} \]

Figure 1. Examples of point-to-point two-way channels breaking up into two parallel one-way channels [1].

\[ Y_1 = Y_2 = X_1X_2 \]. In the binary modulo 2 adder channel, information independently flows in the \( \rightarrow \) and the \( \leftarrow \) “directions” and nodes need not interact, or adapt their current inputs to past outputs, to achieve capacity.

In fact, in his original paper [15] Shannon demonstrated a class of two-way channels for which his inner bound and outer bound match. That is, if the channel has sufficiently symmetric structure such that the optimal input distribution \( P(x_1, x_2) \) is obtained with independent \( x_1 \)
and \( x_2 \) (no interaction, or adaptation is useless), the capacity region can be determined. In a similar fashion, the capacity of a two-way Gaussian point-to-point channel is equal to two parallel Gaussian channels as shown in Figure 1 (b), which may be achieved without the use of adaptation at the nodes [18]. Similar results are true for two-way additive exponential noise family channels [19].

In the spirit of Shannon’s original findings that there exists two-way channels where capacity regions are known, we study two-way networks, which is introduced in the next section.

**A note on terminology.** In this work, “adaptation” or “interaction” is said to take place when the next channel input is a non-trivial function of the node’s past received signals. One may alternatively use the terms “feedback” or “cooperation” instead of adaptation or interaction. However, we feel that “adaptation” and “interaction” better highlights the nature of two-way communications where there is no real notion of feedback (which suggests backwards links which serve to aid communication in the forward direction) as all links may carry information for both directions simultaneously. “Cooperation” reflects the fact that nodes may help each other in multi-user two-way channels, but has been used in many existing one-way communication scenarios; “interaction” avoids notions of directionality.

### 1.2 Two-way Networks

We seek examples of multi-user two-way channels rather than point-to-point two-way channels where, even though nodes may adapt current inputs to past outputs, this is not beneficial from a capacity region perspective. In two-way networks, one may expect adaptation to, in general, be useful and enlarge the capacity region. For example, in multi-user Gaussian channels
one may intuitively expect adaptation to allow for correlation between channel inputs which may translate to coherent gains, or allow for routing messages along different paths. However, as we will see, there exist multi-user channels for which adaptation is useless. In particular, we introduce three two-way network models and a natural extension of the third model:

1. the **two-way Multiple Access / Broadcast channel (MAC/BC)** in which there are 4 messages and 3 terminals forming a MAC channel in the → direction (2 messages) and a BC channel in the opposite ← direction (2 messages);

2. the **two-way Z channel** in which there are 6 messages and 4 terminals forming a Z channel in the → direction (3 messages) and another Z channel in the opposite ← direction (3 messages);

3. the **two-way interference channel (IC)** with 4 messages and 4 terminals forming an IC in the → direction (2 messages) and another IC in the ← direction (2 messages).

4. the **K-pair-user two-way IC** with 2K messages and 2K terminals forming a K-user IC (K messages) in the → direction (K messages) and another K-user IC in the ← direction (K messages).

We emphasize that channel inputs at node $j$ at time $i$ may be functions of the received signals at node $j$ from times 1 to $i - 1$, and that data and “feedback” share the same links, i.e. there are no orthogonal feedback links. Our central contributions are the derivation of the exact, computable, or approximate (to within a constant gap) capacity region of several two-way networks in which adaptation is useless (or leads to bounded gaps) from a capacity perspective.
Typically two-way problems/networks result in multi-letter expressions or auxiliary random variables; our results do not. For the first three models:

- We consider **deterministic binary modulo 2 adder channels**. These are the simplest examples of multi-user two-way channels where one might intuitively expect adaptation to be useless. For these channel models, and slight generalizations thereof, we obtain outer bounds, and demonstrate that non-adaptive time-sharing schemes between nodes transmitting in the same direction achieves capacity. Nodes transmitting data in opposite directions simultaneously transmit.

- We next consider **linear deterministic models** which model Gaussian channels at high SNR [20] and again ask whether adaptation may increase the capacity regions beyond that of two parallel one-way multi-user channels in the $\rightarrow$ and $\leftarrow$ directions. We will show that it does not for the first two channel models by obtaining their capacity regions. For the two-way interference channel, we show that **partial adaptation** where only two of the four nodes may adapt, can “block” the two-way information flow and destroy the ability to relay / cooperate, resulting in a capacity region equal to two non-adaptive ICs. In addition, in some regimes of the relative link strengths, we obtain the capacity region for the symmetric model with **full adaptation** where all four nodes are permitted to adapt.

- We next consider two noisy Gaussian networks. First, for the **Gaussian two-way MAC/BC** we demonstrate that adaptation may only increase the sum-rate in each direction by up to $\frac{1}{2}$ bit. Next, we consider the **symmetric two-way Gaussian IC** where all “direct” channel gains are equal and all “cross-over” channel gains are equal. We derive new, computable outer bounds
for the symmetric sum-rates for this Gaussian channel model and show that: a) adaptation is useless in very strong interference for the partially adaptive model, b) in strong but not very strong interference, non-adaptive schemes perform to within 1 bit per user per direction of the fully adaptive capacity region, and c) the particular non-adaptive Han and Kobayashi scheme of [21] employed in each direction, achieves to within a constant gap (2 bits per user per direction maximally) of fully or partially adaptive outer bounds in all other regimes. We provide examples of non-symmetric Gaussian two-way ICs where adaptation may provide unbounded gain over non-adaptation, and where perfect output feedback may provide unbounded gain over adaptation.

For the K-pair-user two-way IC, compared to the 2-pair-user IC, each user in the K-pair-user two-way IC suffers interference from the K−1 users on the opposite side, as well as possibly from the K−1 users on the same side due to the interaction between users. Thus, received signals may in general be combinations of all 2^K messages. We derive new outer bounds for the symmetric K-pair-user two-way linear deterministic and Gaussian ICs. For the linear deterministic channel model, we show the sum-capacity in the moderately weak and strong interference regimes, and this corresponds to the sum-capacity when no adaptation is permitted (i.e. the “W” curve for two one-way K-user ICs). Achievability follows from the existing non-interactive scheme for the one-way K-user linear deterministic IC [22]. For the Gaussian model, again for the moderately weak and strong interference regimes, we show that the symmetric sum-capacity is to within a constant gap of two non-interactive outer bounds of two simultaneous one-way K-user Gaussian ICs, which in turn is to within a constant gap of non-interactive achievability scheme for most
channel gains (i.e. outside an outage set), as shown in [23]. The technique used for deriving
outer bounds bears semblance to the proof of the outer bound for the one-way K-user Gaussian
IC with feedback [24]. However, due to the additional messages, additional noise and adaptation
in our channel model, the construction of some terms in the proof is non-trivial and novel.

The emphasis of this work is on demonstrating when adaptive schemes are useless, and
when, even if adaptation is permitted, it does not significantly increase the capacity region.

**Related work.** This work builds on: point-to-point two-way channels, one-way multi-user
deterministic channels, and one-way multi-user channels with feedback. Little work exists thus
far on two-way multi-user channels.

The capacity region of the general point-to-point discrete memoryless two-way channel may
be written in terms as a limit of multi-letter expressions as in [15, Section 15], or [17, Theorem
4.1]. Given the complexity in computing this capacity region, it is not entirely satisfying and
the capacity region of the two-way channel is generally considered to be open. The binary
multiplier channel (BMC) [25–29] is a nice example of a deterministic, binary, common output
two-way channel where capacity is not exactly known, though its capacity may be expressed in
terms of directed information as in [17, Corollary 4.1].

The first of our three channel models is a two-way MAC/BC channel. The capacity regions
of the linear-deterministic one-way MAC and BC channels were obtained in [30]. An achiev-
able rate region and an outer bound of a similar two-way and adaptive multi-user half-duplex
two-way channel is derived in [31, 32] for Gaussian and discrete memoryless channels (DMC),
respectively. In particular, the achievable rate region derived employs adaptation using Block
Markov encoding, and the outer bound contains both auxiliary random variables and messages in its expression and is thus difficult to compute. These works differ from our model in that we assume full-duplex operation, have 2 broadcast messages rather than a common one. Other than [31, 32], the two-way MAC/BC has not been considered, and bears most resemblance to a combined MAC channel with feedback and BC channel with feedback (see references in [33, Ch. 17, Bibliographic Notes], and in particular [34, 35]), though we note that in our two-way model there are no “free” feedback links – any feedback must travel over the same links as the data in the opposite direction, and hence the MAC and BC with feedback results are not directly applicable.

The second channel model we consider is the two-way Z channel, with 6 messages. The one-way Z channel (with 3 messages, rather than the Z Interference channel with 2 messages) was first studied in [36], in which a general outer bound, and a matching inner bound for a special class of degraded Z channels are obtained. The capacity region of the one-way deterministic Z channel with invertibility constraints similar in flavor to those in [37], is found in [38], which will be of use here.

The last channel model considered is the two-way IC in which there are 4 messages and 4 terminals forming ICs in the → and ← directions. The capacity region of the one-way modulo 2 adder IC is known [33] and is a special example of a more general class of deterministic IC for which capacity is known [37], including the one-way linear deterministic IC [30]. The work here is also related to one-way ICs with perfect output feedback [39, 40], with rate-limited feedback [41], with generalized feedback [42], and interfering feedback [39, 43]. In all these
channel models only two messages are present and the “feedback” links, whether perfect, noisy, or interfering still serve only to further rates in the forward direction. The only other example of a 4-message two-way interference channel besides our prior work [4–7] is in Section VI of [43,44], where an example of a linear deterministic scheme is provided which shows that, at least for one particular asymmetric linear deterministic two-way IC in weak interference in the → and strong interference in the ← direction, that adaptation can significantly improve the capacity region over non-interaction. The general capacity region of the linear deterministic two-way IC (with 4 messages) remains open in general despite the example in [43,44] and the progress made here. One final word on terminology: we will refer to the 4 message two-way IC as the “two-way IC” and the 2 message channel of [43,44] – considered in all sections but Section VI – as the “two-way interference channel with interfering feedback” to emphasize that the rates are still flowing in one direction only. Further comparisons with ICs with/without feedback [21,39,40,43,44] will be made in Section 2.4 and 3.2. The model of two-way IC is also used for studying multi-cell cooperation from more practical perspective in wireless communications considering imperfect channel state information [45].

The K-pair-user two-way IC is a natural extension of two-way IC, and most related to the one-way K-user IC, whose related work is found in the next sub-section.

1.3 Full-duplex Wireless, Degrees of Freedom, and Relaying

In wireless communications, current two-way systems often employ either time or frequency division to achieve two-way or bidirectional communication. This restriction is due to a combination of hardware and implementation imperfections and effectively orthogonalizes the two di-
rections, rendering the bidirectional channel equivalent to two one-way communication systems. However, much progress has been made on the design of full-duplex wireless systems [46, 47], which show great promise for increasing data rates in future wireless technologies. In this work we seek to understand the potential of full-duplex systems in a two-way multi-user or network setting, and do so from a multi-user information theoretic perspective by obtaining the degrees of freedom of several full-duplex two-way networks with and without relays.

We currently understand the theoretical limits of a point-to-point, full-duplex Gaussian two-way channel where two users wish to exchange messages over in-band Gaussian channels in each direction: the capacity region is equal to two independent Gaussian noise channels operating in parallel [18]. Full duplex operation thus roughly doubles the capacity of this simple two-way network.

To extend our understanding of the impact of full-duplex operation to two-way networks with interference, the two-way interference channel (TWIC) has been studied in [1, 6, 7, 44]. Recall that all 4 nodes in this network act as both sources and destinations of messages. This allows for interaction between the nodes: a node’s channel inputs may be functions of its message and previously received signals. The capacity region of the point-to-point two-way channel is still open in general, though we know the capacity for the Gaussian channel, and is known to be a remarkably difficult problem. Similarly, the capacity region of the one-way IC is still open, though we know its capacity to within a constant gap for the Gaussian noise channel [21]. In general then, finding the full capacity region of the full-duplex TWIC is a difficult task, though progress has been made for several classes of deterministic channel models [1] (as shown in
this thesis) and capacity is known to within a constant gap in certain parameter regimes and adaptation constraints \([1,2,7]\) (again, as shown in this thesis).

The degrees of freedom (DoF) \([48]\) is an alternative (to the all out, challenging capacity region) approximate capacity characterization that intuitively corresponds to the number of independent interference-free signals that can be communicated in a network at high signal to noise ratios (SNR) which has been of significant recent interest in one-way interference networks \([49,50]\). Here, we seek to extend our understanding of the DoF to two-way networks, whose study is motivated by the fact that full-duplex operation is becoming practically realizable. Some progress has already been made: the DoF of the full-duplex TWIC has been shown to be 2 \([6,7]\) (shown here). This is interesting, because the capacity of any network with interaction at nodes is no smaller than that of the same network where interaction is not possible (interaction can mimic non-interaction). However, that the TWIC with interaction has DoF 2 demonstrates that interaction between users does not increase the DoF of the two-way IC beyond the doubling that full-duplex operation gives. We ask whether the same is true for K-pair user two-way, full-duplex interference channels with and without a MIMO relay node.

In this work, we consider the K-pair-user two-way interference channel, as introduced in the previous section, with and without the presence of a MIMO relay. Recall that compared to the 2-pair-user IC, the K-pair-user two-way IC experiences interference from significantly more users: not only may a receiver see signals from all other transmitters transmitting in the same direction, but due to the adaptation involved, these signals may also contain information about users transmitting in the opposite direction. Hence, any user may see a combination of
the signals of all other $2K - 1$ users in addition to seeing self-interference (SI) signals, which are transmitted by the user itself or received via other signals due to adaptation. Canceling SI is one of the main challenges in real full-duplex wireless systems. However, in this theoretical work for the Gaussian channels involved, the self-interference is known to the receiver and as such, theoretically, it can be subtracted off. We then explore the limits of communication under the assumption that this self-interference may be removed. Our main results are [3, 8, 51]:

1) We first show that the sum degrees of freedom of the $K$-pair-user TWIC is $K$, i.e. $K/2$ in each direction, for both time-varying and (almost all) constant channel coefficients. In other words, each user still gets half a DoF and interaction between users – even though our outer bound permits it – is again useless from a DoF perspective. Intuitively this is because all the links in the network have similar strengths in the DoF sense, so that a user cannot “route” other users’ desired signals through backward links since they are occupied by its own data signals. In addition, coherent power gains which may be the result of adaptation and the ability of nodes to correlate their channel inputs, do not affect the DoF (i.e. coherent power gains for Gaussian channels lead to additive power gains inside the logarithm rather pre-log, or DoF/multiplexing gains). Achievability follows from known results of the one-way $K$-user IC; the contribution lies in the novel outer bound. Full-duplex operation is thus seen to double the DoF, but interaction is not able to increase the DoF beyond this.

2) We next consider the $K$-pair-user TWIC with an additional, multi-antenna, full-duplex relay node which does not have a message of its own and only seeks to aid the communication of the $K$-pair users. We ask whether the presence of such a relay node may increase the DoF.
Interestingly, we show that while the DoF of the K-pair-user TWIC is K – indicating that interference is present and somewhat limiting rates in the K-user IC in each direction – that the presence of an instantaneous MIMO relay with 2K antennas may increase the DoF to the maximal value of 2K, i.e. each user in the network is able to communicate with its desired user in a completely interference-free (in the DoF sense) environment. The key assumption needed is for the relay to be non-causal or instantaneous – meaning that at time k it may forward a signal based on the received up to and including time k. We see that full-duplex operation combined with instantaneous / non-causal relaying with multiple antennas may in this case quadruple the DoF over the one-way K-user IC.

3) Finally, we show a result which is sharp contrast to the previous point: if the relay is now causal instead of non-causal, meaning that at time k it may only forward a signal which depends on the received signals up to and including time k − 1, then we derive a novel outer bound which shows that the DoF of the K-pair-user TWIC with a causal MIMO relay is K (regardless of the number of antennas at the relay). This is the same as that achieved without a relay, and without interaction. In summary, full-duplex operation again doubles the DoF, but a causal, full duplex relay is unable to increase the DoF beyond that.

**Related work.** The degrees of freedom of a variety of one-way communication networks have been characterized [50, 52–55]. However, much less is known about the DoF of two-way communications. Very recently, [56] considered a half-duplex two-pair two-way interference channel (where nodes other than the relay may not employ interaction and hence are much more restricted than the nodes here, i.e. transmit signals are functions of the messages only and not...
past outputs) with a 2-antenna relay and showed that $4/3$ DoF are achievable. No converse results were provided. In [57], the authors identified the DoF of the full-duplex 2-pair and 3-pair two-way multi-antenna relay MIMO interference channel, in which there is no interference between users who only communicate through the relay (no direct links). We consider direct links between all users (not in the same sides) in the two-way interference channels, as well as links between all users and the relay. We also note that the general results of [50], which state that relays, noisy cooperation, perfect feedback and full-duplex operation does not increase the DoF of one-way networks, do not apply, as we consider nodes which are both sources and destinations of messages (two-way rather than one-way).

The $K$-user interference channel, as an extension of the 2-user interference channel, information theoretically models wireless communications in networks involving more than two-pairs of users. Using the idea of interference alignment [49, 58, 59], the DoF of the $K$-user (one-way) IC for both time-varying channels and (almost all)\(^1\) constant channels has been shown to be $K/2$ in [49] and [60] respectively. The generalized DoF of the $K$-user IC without and with feedback have been characterized in [22] and [24] (full feedback from receiver $i$ to transmitter $i$) respectively. Authors in [61] showed that for almost all constant channel coefficients of fully connected two-hop wireless networks with $K$ sources, $K$ relays and $K$ destinations (source nodes are not destination nodes as they are here, i.e., the network is one-way), the DoF is $K$.

\(^1\)The precise definition of “almost all” may be found in [60].
We note that our work differs from prior work in that we consider an interactive, full-duplex Gaussian K-pair-user TWIC for the first time, with and without a relay (which may be either non-causal or causal), and obtain not only sum-rate achievability but also converse DoF results for all three general channel models considered. We emphasize that we seek information theoretic DoF results, which act as benchmarks / upper bound for the performance of realistic scenarios.

1.4 Outline

We first explore wireless deterministic two-way networks in Chapter 2. Specifically, we introduce our notation, definitions of adaptation and capacity regions, and the three general channel models considered in Section 2.1. We then proceed to study deterministic two-way MAC/BC, two-way Z channel, and two-way IC in Section 2.2, 2.3, and 2.4 respectively. Wireless Gaussian two-way networks is considered in Chapter 3. In particular, we first study Gaussian two-way MAC/BC and Gaussian two-way IC in Section 3.1 and 3.2 respectively. Then we propose the K-pair-user two-way IC and study its linear deterministic and Gaussian models in Section 3.3.

In Chapter 4 we characterize the degrees of freedom of wireless two-way full-duplex networks. Specifically, we describe our system model of K-pair-user full-duplex two-way IC with and without a MIMO relay in Section 4.1. We show that K DoF is optimal for the Gaussian K-pair-user TWIC in Section 4.2. Then we proceed to consider the K-pair-user TWIC with an instantaneous MIMO relay in Section 4.3, where we show that the maximum 2K DoF may be achieved with the help of an instantaneous MIMO relay with at least 2K antennas. In Section
4.4 we then show that if the relay is causal rather than non-causal, that, for the K-pair-user TWIC, the presence of a full-duplex, multi-antenna relay cannot increase the DoF beyond K. We conclude the thesis in Chapter 5 with some general observations and intuition of our findings. In addition, several directions for future research are provided.
CHAPTER 2

WIRELESS DETERMINISTIC TWO-WAY NETWORKS

In this chapter, we will explore three deterministic multi-user two-way channel models separately: 1) two-way MAC/BC; 2) two-way Z channel; and 3) two-way interference channel. The deterministic models are noiseless and easier to study compared to noisy channels, as we focus on whether adaptation can increase capacity rather than on whether it may effectively combat noise. Specifically, we study three deterministic models: 1) the binary modulo 2 adder channels, which perhaps the simplest model to begin with; 2) invertible and cardinality constrained channels, which are a slightly generalized version of 1); and 3) the linear deterministic channels, which model the Gaussian channels at high SNR and often yield good intuition for quasi-optimal schemes at finite SNR.

2.1 Models, Definitions and Notations

We consider three multi-user two-way channels, where all nodes act as both transmitters (encoders) and receivers (decoders), as shown in Figure 2, and described by:

- the two-way MAC/BC channel: transmitters 1 and 3 send independent messages $M_{12}$ and $M_{32}$ to receiver 2, respectively, forming a MAC in the $\rightarrow$ direction. Transmitter 2 sends independent messages $M_{21}$ and $M_{23}$ to receivers 1 and 3, respectively, forming a BC in the $\leftarrow$ direction.
• the two-way Z channel: transmitters 1 and 4 send messages \( M_{12} \) and \( M_{43} \) to receivers 2 and 3 respectively. Transmitters 2 and 3 send messages \( (M_{21}, M_{23}) \) and \( (M_{32}, M_{34}) \) to receivers 1,3 and 2,4 respectively.

• the two-way interference channel: transmitters 1 and 3 send messages \( M_{12} \) and \( M_{34} \) to receivers 2 and 4, respectively, forming an IC in the \( \rightarrow \) direction. Similarly, transmitters 2 and 4 send messages \( M_{21} \) and \( M_{43} \) to receivers 1 and 3 respectively, forming another IC in the \( \leftarrow \) direction.

For each of these models, let \( M_{jk} \) denote the message from node \( j \) to node \( k \); all messages are independent and uniformly distributed over \( M_{jk} := \{1, 2, \ldots, 2^{n_{jk}}\} \), where the ranges of \( j, k \) depend on the channel model (all subsets of \( \{1, 2, 3, 4\} \)) and \( R_{jk} \) is the rate of transmission from node \( j \) to node \( k \). For example, in the MAC/BC, \( R_{12} \) is the rate of message \( M_{12} \) but \( R_{13} \) and \( M_{13} \) do not exist.

All channels are assumed to be memoryless and at each channel use, described by the input/output relationships in Figure 2. Let \( X_j \) and \( Y_k \) denote the channel input of node \( j \) and output at node \( k \) used to describe the model (per channel use). Let \( X_{j,i} \) \( (Y_{j,i}) \) denote the channel input (output) at node \( j \) at channel use \( i \), and \( X^n_j := (X_{1,1}^n, X_{1,2}^n, \ldots, X_{1,n}^n) \). Let \( [x]^+ = \max(0, x) \). For the binary modulo 2 adder channels the input and output alphabets are \( \{0, 1\} \), and \( \oplus \) denotes modulo 2 addition. For the linear deterministic models, the channel inputs and outputs are binary vectors, and all addition will be bit-wise and modulo 2. We furthermore let \( S \) denote an \( N \times N \) lower shift matrix, where \( N = \max(n_{jk}) \) over the relevant \( j, k \) for each channel model, and \( n_{jk} \) defines the number of signal bit levels from transmitter \( j \) to receiver \( k \). We
will also consider another type of channel models: the “deterministic, invertible and cardinality
constrained” deterministic channel models, which will be defined in the appropriate sections.

A node $j$ is said to employ adaptation or interaction if the channel input at time $i$ is a
function of the previously received outputs,

\[ X_{j,i} = f_j(M_{jK}, Y^{i-1}_j), \]  

(2.1)

where $f_j$ ($j \in \{1, 2, 3, 4\}$) are deterministic functions, and $M_{jK} := \{M_{jk} | k \in K\}$ are the (sets of)
messages from node $j$ to all the nodes in $K \subset \{1, 2, 3, 4\}$ where $K$ depends on the channel model,
and may be obtained from Figure 2. If a node behaves in a non-adaptive or restricted fashion
then its inputs are functions of its messages only, i.e. $X_{j,i} = f_j(M_{jk})$. If some nodes adapt
while the others do not, we refer to this as partial adaptation, and will specify which nodes
adapt. Receiver $k$ uses a decoding function $g_k : \mathcal{Y}_k^n \times \mathcal{M}_{kI} \rightarrow \hat{\mathcal{M}}_{J_k}$ to obtain estimates of all
transmitted messages destined to received $k$, $\hat{\mathcal{M}}_{J_k} := \{\hat{M}_{jk} | j \in J, J \subset \{1, 2, 3, 4\}\}$ depending on
the model, given knowledge of its own message(s) $M_{kI}$ for suitable $I \subset \{1, 2, 3, 4\}$, which again
depends on model. The capacity region of each channel model is the closure of the set of rate
tuples for which there exist encoding and decoding functions (of the appropriate rates) which
simultaneously drive the probability that any of the estimated messages is not equal to the true
message, to zero as $n \rightarrow \infty$. 

Figure 2. Three multi-user two-way channel models and two of the classes under consideration [1].
2.2 Two-way Multiple-Access Broadcast Channel

We first consider the 3 user, full-duplex two-way MAC/BC network as shown in Figure 2(a). As an introductory example, we first show that adaptation cannot increase the capacity for the modulo 2 adder MAC/BC and a slight generalization thereof, and capacity may be achieved via time-sharing. Finally, we consider the linear deterministic two-way MAC/BC and show that, once again, adaptation is useless.

2.2.1 An Introductory Example: Modulo 2 Adder MAC/BC

In the two-way modulo 2 adder MAC/BC, we emphasize that all three users may employ full adaptation – i.e. all channel inputs at time $i$ may be a function of previously received channel outputs at that node. There are no additional orthogonal, or free, “feedback” links. The capacity region may be stated as follows.

**Theorem 1.** The capacity region of the two-way modulo 2 adder MAC/BC channel is the set of non-negative rate tuples $(R_{12}, R_{32}, R_{21}, R_{23})$ such that

\[
R_{12} + R_{32} \leq 1 \tag{2.2}
\]

\[
R_{23} + R_{21} \leq 1. \tag{2.3}
\]

**Proof.** The outer bound follows from the cut-set bound. The inner bound follows by time-sharing as in Figure 3 *without adaptation*: $\alpha$ time-shares between channel inputs $X_1$ and $X_3$ for the MAC channel in the $\rightarrow$ direction, while $\beta$ time-shares between the messages $M_{21}$ and
Remark 1. We note that the outer bound may alternatively be derived from either 1) Fano’s inequality and first principles, taking into account the ability of the nodes to adapt (we provide one such example for the linear deterministic MAC/BC channel in Theorem 3 for completeness), or 2) yet another alternative is to provide Tx 1 and 3 with both $M_{21}, M_{23}$ and perfect channel output feedback $Y_{2,i-1}$ (to derive a bound on $R_{12} + R_{32}$ in the direction) and to provide Tx 2 with channel output feedback $Y_{1,i-1}$ and $Y_{3,i-1}$ and messages $M_{12}, M_{32}$ (to derive a bound on $R_{21} + R_{23}$ in the direction) at time $i$. Then, one can mimic the outer bound for a class of MACs with FB for the direction as derived by Willems [34] and the physically degraded BC with feedback of [35] for the BC direction (which goes through without a problem for the modulo 2 adder and linear deterministic models). Willems’ class of channels is one for which (in our notation), at least one of $H(X_1 | Y_2, X_3, X_2)$ or $H(X_3 | Y_2, X_1, X_2)$ is zero for all input distributions. We note that we provide Tx 1 and Tx 3 with $M_{21}, M_{23}$ in addition to the output feedback in order to be able to construct the inputs $X_{2,i}$, so that $X_2$ would be placed in the conditioning of the bounds of [34]. Note also that while the capacity of Willems’ class of discrete memoryless channels with feedback is expressed in terms of an auxiliary random variable $U$ which is the result of the feedback and its ability to correlate channel inputs. In general, this would result in a larger region than the MAC channel without feedback. However, for our binary modulo 2 channel law, even with conditioning on $X_2$, and the fact that $X_1, X_2, X_3$ may all be correlated, these evaluate to the same region; adaptation is useless.
Remark 2. The capacity region of Theorem 1 is the same as that of a modulo 2 adder MAC and a modulo 2 adder BC channel in parallel, which do not interact. That is, the capacity of a one-way modulo 2 adder MAC is $R_{12} + R_{32} \leq 1$, while that of a one-way modulo 2 adder BC (which is actually just a BC with $Y_1 = X_2 = Y_3$) is $R_{23} + R_{21} \leq 1$. No adaptation is needed to achieve these regions: capacity is achieved by time-sharing amongst the data traveling in the same “direction” (i.e. between nodes 1 and 3, and between messages $M_{21}$ and $M_{23}$) but not between the two directions themselves.

2.2.2 A More General Model for Deterministic MAC/BC

Adaptation is useless for the simple modulo 2 adder MAC/BC channel and capacity is achieved using time-sharing in each direction. We ask whether there exists a larger class of channels for which this holds. We answer this positively by considering the “deterministic, invertible and alphabet restricted” class of two-way MAC/BCs with:

$$Y_1 = F_1(X_1, X_2)$$

$$Y_2 = F_2(X_1, X_2, X_3)$$

$$Y_3 = F_3(X_2, X_3)$$

where $F_m()$, $m \in \{1, 2, 3\}$ are deterministic functions which also satisfy

- **P1:** $|X_1| = |X_2| = |X_3| = |X| = |Y_1| = |Y_2| = |Y_3| = |Y| = \kappa$ for known $\kappa \in \mathbb{N}^+$. 

• **P2:** Given $X_1$, $Y_1$ is invertible, i.e. $\exists$ a function $G_1$ s.t. $X_2 = G_1(X_1, Y_1)$. Similarly, we assume $\exists G_{21}, G_{23}, G_3$: $X_1 = G_{21}(X_2, X_3, Y_2)$, $X_3 = G_{23}(X_1, X_3, Y_2)$, and $X_2 = G_3(X_3, Y_3)$. These conditions exclude two-way channels such as the binary multiplier channel.

• **P3:** $\exists x_3^*$ such that given $X_3 = x_3^*$, $X_1, X_2$ both uniform on their alphabets implies both $Y_1$ and $Y_2$ uniform on their alphabets. Similarly, $\exists x_1^*$ such that given $X_1 = x_1^*$, $X_3, X_2$ both uniform on their alphabets implies $Y_2, Y_3$ uniform on their alphabets. This ensures we can achieve the full $\log(\kappa)$, and is true only for channels with a high degree of symmetry.

Under these conditions (which we only claim are sufficient and not necessarily necessary), the capacity region of the deterministic MAC/BC is given in the following Theorem. These conditions somewhat resemble the symmetry structure seen in discrete one-way additive noise channels, where feedback has been shown to be useless in \cite{62}.

**Theorem 2.** The capacity region of the two-way “deterministic, invertible and alphabet restricted” MAC/BC satisfying the conditions P1, P2, P3 is the set of non-negative $(R_{12}, R_{32}, R_{21}, R_{43})$ satisfying:

$$R_{12} + R_{32} \leq \log \kappa$$

$$R_{21} + R_{23} \leq \log \kappa,$$

which may be achieved via time-sharing (in each direction).

**Proof.** The outer bound follows directly from the cut set, or may be directly derived from Fano’s inequality or as a result of MAC and BC channels with feedback, as in Remark 1. The
restriction on the alphabet sizes condition \( P1 \) prohibits “coherent gain” - like phenomena in the outer bound, where correlation between user inputs may be beneficial, as in for example the Gaussian MAC channel with feedback.

Our achievability scheme time-shares between user 1 and user 3 in the → (MAC) direction, while simultaneously time-sharing between sending data to user 1 and 3 in the ← (BC) direction, as in Figure 3. There, we see two time-sharing coefficients \( 0 \leq \alpha, \beta \leq 1 \), where \( \alpha \) time-shares in the → direction and \( \beta \) in the ← direction. Let us consider the rates achieved in time slot (1), of duration \( \alpha \) (WLOG we have taken \( \alpha < \beta \)). Node 1 encodes \( M_{12} \) into \( X_1 \) uniformly distributed over the \( \kappa \) input symbols; node 2 encodes \( M_{21} \) into \( X_2 \) uniformly distributed over \( \kappa \) input symbols and node 3 fixes \( X_3 = x_3^* \) (rate 0). We claim this scheme achieves the rates \( R_{12} = R_{21} = \alpha \log(\kappa) \), \( R_{23} = R_{32} = 0 \). Consider \( R_{12} \): node 2 receives \( Y_2 = F_2(X_1, X_2, x_3^*) \). Since node 2 knows \( X_2 \) and knows that \( X_3 = x_3^* \), by \( P2 \), it may construct \( X_1 = G_{21}(X_2, x_3^*, Y_2) \) to decode \( M_{12} \). By \( P3 \) this may be done at full rate \( \alpha \log(\kappa) \). Similar arguments for time slots (2) and (3) demonstrate that the rates in (Equation 2.4)-(Equation 2.5) are achievable.

One example, besides the binary modulo 2 adder channel, is the channel with input alphabets \{0, 1, \cdots, \kappa - 1\} for some \( \kappa \) and \( Y_1 = X_1 + X_2 \mod \kappa \), \( Y_2 = X_1 + X_2 + X_3 \mod \kappa \), and \( Y_3 = X_2 + X_3 \mod \kappa \).

**Remark 3.** The restriction on the cardinality (sufficient, but may not be necessary) was brought about by simply considering the two-way MAC/BC binary adder channel (not modulo), with inputs \( X_1 = X_3 = \{0, 1\} \) and outputs \( Y_2 = X_1 + X_3 \) with alphabet \{0, 1, 2\} in the MAC direction,
for which it is easy to derive inner and outer bounds both of the form $R_{12} + R_{32} \leq H(X_1 + X_3)$.

In general, one would hope, like Shannon did for the point-to-point two-way channel [15], to derive multi-user inner and outer bounds of the same form. However, even if one is able to do so (which may be too much to hope for in general, but may be reasonable for certain classes of deterministic models), we are left with the distributions over which these bounds are taken. That is, back to our example, to claim that adaptation is useless, the inner bound should be taken over independent input distributions $p(x_1)p(x_3)$, while the outer bound, in general allowing for adaptation, is taken over joint distributions $p(x_1, x_3)$ (unless one restricts the set of input distributions perhaps via dependence-balance-bound-like techniques [29], an open problem). Inner and outer bounds taken over these different sets of distributions do not match.
for the binary adder channel with ternary output. As such, we restricted the channels to those for which a form of cooperation (or adaptation) between users cannot possibly help – which is the case for the modulo adder channels, and as we will see, the similar, in terms of properties, linear deterministic channels.

2.2.3 Linear Deterministic MAC/BC

The two-way linear deterministic MAC/BC channel is defined by the input/output equations as in Figure 2(a). We recall that all nodes are permitted to adapt, so that at channel use $i$, $X_{1,i} = f_1(M_{12}, Y_{1}^{i-1})$, $X_{2,i} = f_2(M_{21}, M_{23}, Y_{2}^{i-1})$, and $X_{3,i} = f_3(M_{32}, Y_{3}^{i-1})$. The capacity region may be stated as follows:

**Theorem 3.** The capacity region of the two-way linear deterministic MAC/BC is the set of non-negative rate tuples $(R_{12}, R_{32}, R_{21}, R_{23})$ such that

\[
MAC \rightarrow \begin{cases} 
R_{12} \leq n_{12}, & R_{32} \leq n_{32}, \\
R_{12} + R_{32} \leq \max(n_{12}, n_{32}) \end{cases} 
\]  

\[
BC \leftarrow \begin{cases} 
R_{21} \leq n_{21}, & R_{23} \leq n_{23} \\
R_{21} + R_{23} \leq \max(n_{21}, n_{23}) \end{cases} 
\]

**Proof.** Achievability may be argued via [20] by mimicking a one-way MAC and one-way BC channel in opposite directions and noting that each user may subtract off its own transmitted signal from its received signal. The outer bounds may be obtained by the cut-set, or via an
alternative direct proof. We include an example of this alternative (to the cut-set) outer bound proof below, out of interest and to demonstrate how adaptation may be taken into account.

\[ n(R_{12} - \epsilon) \leq I(M_{12}; Y^n_2|M_{21}, M_{23}, M_{32}) \leq \sum_{i=1}^{n} [H(Y_{2,i}|Y^{-1}_2, M_{21}, M_{23}, M_{32}, X^n_2)] \]

\[ = \sum_{i=1}^{n} [H(Y_{2,i}|Y^{-1}_2, M_{21}, M_{23}, X^n_2, X^n_3)] \leq \sum_{i=1}^{n} [H(S^{-n_{12}}X_{1,i})] \leq n(n_{12}), \]

where (a) follows since given \((Y_2^{-2}, M_{21}, M_{23})\), we may construct \(X^n_2\), which cancels out the “self-interference” term \(X_{2,i}\) in \(Y_2,i\). We note that the self-interference term can be always cancelled out in this way in the converse of additive models. Step (b) follows from the fact that given \(M_{32}, X^n_2\), we may construct \(X^n_3\). The other single rate bounds follow similarly.

\[ n(R_{12} + R_{32} - \epsilon) \leq I(M_{12}, M_{32}; Y^n_2|M_{21}, M_{23}) \]

\[ \leq \sum_{i=1}^{n} [H(Y_{2,i}|Y^{-1}_2, M_{21}, M_{23}, X^n_2)] \]

\[ \leq \sum_{i=1}^{n} [H(S^{-n_{12}}X_{1,i} + S^{-n_{32}}X_{3,i})] \leq n(\max(n_{12}, n_{32})). \]

We may analogously obtain the other sum-rate bound.

Remark 4. Without adaptation, the channel would correspond to a MAC channel simultaneously transmitting with a BC channel with restricted nodes. This coincides with our outer bound with adaptation, which may furthermore be achieved in one channel use: adaptation is useless.
2.3 Two-way Z Channels

We now consider the 4 user, full-duplex network shown in Figure 2(b). The 6 message network, resembles a cascade of three two-way channels, in the shape of a Z (in each direction).

2.3.1 An Introductory Example: Modulo 2 Adder Two-way Z Channel

The two-way modulo 2 adder Z channel is discrete and memoryless, and all four users may employ full adaptation. The capacity region of this channel is stated as follows:

\textbf{Theorem 4.} The capacity region of the two-way modulo 2 adder Z channel is the set of non-negative rate tuples \((R_{12}, R_{21}, R_{23}, R_{32}, R_{34}, R_{43})\) such that

\begin{align*}
R_{12} + R_{32} + R_{34} &\leq 1 \quad (2.8) \\
R_{21} + R_{23} + R_{43} &\leq 1 \quad (2.9)
\end{align*}

The proof is found in Appendix A and is not a direct consequence of the cut-set outer bound.

\textbf{Remark 5.} We note that the proof of the sum-rate outer bound of the Z channel in Theorems 4, 5, and the sum-rate bounds of the two-way IC in Theorems 6, 7, 9, 13 all follow the same general idea of giving an asymmetric genie to one receiver, as initially done in [37] for the one-way IC, and quite similar to the Z channel outer bound in [38], and in particular [39, 40] for the one-way IC with feedback (IC with FB). That is, in the \(\rightarrow\) direction, we provide one of the receivers with the message of the non-desired message in the \(\rightarrow\) direction (as in the IC with FB) as well as all messages of the \(\leftarrow\) direction (particular to the two-way channels, as no \(\leftarrow\)
messages in one-way channels), and the desired signal received at the other receiver of the \(\rightarrow\) direction (similar to the genies given in the IC with FB). The additional messages and receiver output (relative to one-way models) are needed to create various inputs, as may be done with less side-information in one-way models.

Remark 6. We again notice that since time-sharing achieves the above region, adaptation does not enlarge the capacity region. We again see that the messages in the \(\rightarrow\) and the \(\leftarrow\) directions may be simultaneously communicated, but that the messages within one direction must be time-shared.

2.3.2 Comments on a More General Model for the Two-way Z Channel

Similar to the more general “deterministic, invertible and restricted” class of two-way MAC/BC channels where it was shown that non-adaptive time-sharing achieves capacity, we may extend the two-way Z modulo 2 adder model to a more general class of two-way Z channels. The converse follows along similar lines as for the modulo 2 adder channel. In terms of achieving the outer bounds \(R_{12} + R_{32} + R_{34} \leq \log \kappa\) and \(R_{21} + R_{23} + R_{43} \leq \log \kappa\), one sufficient condition involves restricting the input and output alphabet sizes to be equal (eliminating some of the potential benefits of adaptation via user cooperation), as well as several symmetry constraints akin to extensions of \(P2\) and \(P3\). Again, one example of such a channel model is the modulo \(\kappa\) channel. We omit the full statement as it follows in a straightforward and analogous fashion to Theorems 2 and 4.
2.3.3 Linear Deterministic Two-way Z Channel

The linear deterministic two-way Z channel is defined by the input / output equations in Figure 2(b). The capacity region is again that of two parallel Z channels in opposite directions; adaptation is useless.

**Theorem 5.** The capacity region of the two-way linear deterministic Z channel is the set of all rate-tuples \( (R_{12}, R_{21}, R_{23}, R_{32}, R_{34}, R_{43}) \) which satisfy the following:

\[
\begin{align*}
Z \rightarrow & \quad \begin{cases} 
R_{12} \leq n_{12}, & R_{32} \leq n_{32}, & R_{34} \leq n_{34} \\
R_{12} + R_{32} \leq \max(n_{12}, n_{32}) \\
R_{32} + R_{34} \leq \max(n_{32}, n_{34}) \\
R_{12} + R_{32} + R_{34} \leq \max(n_{12}, n_{32}) + [n_{34} - n_{32}]^+ 
\end{cases} \\
Z \leftarrow & \quad \begin{cases} 
R_{43} \leq n_{43}, & R_{23} \leq n_{23}, & R_{21} \leq n_{21} \\
R_{43} + R_{23} \leq \max(n_{43}, n_{23}) \\
R_{23} + R_{21} \leq \max(n_{23}, n_{21}) \\
R_{43} + R_{23} + R_{21} \leq \max(n_{43}, n_{23}) + [n_{21} - n_{23}]^+.
\end{cases}
\]

**Proof.** We first note that the capacity of a class of deterministic Z channels is shown in [38, Th. 3.1]. To show achievability of the above, we use the achievability scheme of [38, Th. 3.1] in each \( \rightarrow \) and \( \leftarrow \) direction with non-adaptive nodes. Due to the additive nature of the channel, each receiver may cancel or subtract out its own “self-interference” term \( S^{N-n_j}X_j \) from its received signal. By making the appropriate correspondences, the above is achievable and equivalent to two one-way Z channels.
For the converse, note that all but the triple-rate bounds may be obtained by the cut-set bound, or independently by giving the appropriate side-information or genie to the receivers (as illustrated in previous models). The non-cut-set triple rate bound may be obtained as follows:

\[
\begin{align*}
    n(R_{12} + R_{32} + R_{34} - \varepsilon) & \leq I(M_{12}; Y^n_2 | M_{21}, M_{23}, M_{43}) + I(M_{32}, M_{34}; Y^n_2, Y^n_4 | M_{43}, M_{12}, M_{21}, M_{23}) \\
    & \leq H(Y^n_2 | M_{21}, M_{23}, M_{43}) + H(Y^n_4 | M_{43}, M_{12}, M_{21}, M_{23}, Y^n_2) \\
    & \leq \sum_{i=1}^{n} [H(Y_{2,i}^{i-1}, M_{21}, M_{23}, X^n_2) + H(Y_{4,i} | M_{12}, M_{21}, M_{23}, M_{43}, Y_{4,i}^{i-1}, X^n_2, Y^n_2, X^n_2)] \\
    & \leq \sum_{i=1}^{n} [H(S^{n-n_{12}}X_{1,i} + S^{n-n_{12}}X_{3,i}) + H(S^{n-n_{34}}X_{3,i} | M_{12}, M_{21}, M_{23}, M_{43}, Y_{4,i}^{i-1}, X^n_2, X^n_2)] \\
    & \leq \sum_{i=1}^{n} [H(S^{n-n_{12}}X_{1,i} + S^{n-n_{12}}X_{3,i}) + H(S^{n-n_{34}}X_{3,i} | S^{n-n_{32}}X_{3,i}, X^n_2, X^n_2)] \\
    & \leq n(max(n_{12}, n_{32}) + [n_{34} - n_{32}^+] + n_{12} + n_{34} - n_{32}).
\end{align*}
\]

In (a), \(X^n_2\) in the second entropy term follows since given, \(M_{12}\) and \(X^n_2\), we may construct \(X^n_1\). □

Remark 7. Again, we are always able to achieve the desired rates in Theorem 5 in only one channel use, therefore adaptation is useless. The capacity region of this channel, a 6 dimensional region, is exactly equivalent to the capacity region of the two one-way linear deterministic Z channels.

2.4 Two-way Interference Channels

The last deterministic multi-user two-way network we consider is a 4 user, 4 message, full-duplex network, shown in Figure 2(c). This channel model merges elements of two-way, feed-
back, and interference, and forms two parallel interference channels in the → and ← directions.

Again, we first introduce the modulo 2 adder model of this channel and show that adaptation is useless, generalizing this to a slightly larger class of symmetric channels. This generalization is not as straightforward as for the MAC/BC and Z channels, and hence is discussed in somewhat more depth. Finally, for the symmetric linear deterministic two-way interference channel, we show that full adaptation is useless when the interference is very strong, strong, and in some of the weak regimes, while in all other regimes we show that partial adaptation is useless (i.e. if only 2 of the nodes adapt, might as well have none of the nodes adapt).

2.4.1 An Introductory Example: Modulo 2 Adder Two-way IC

We are again motivated by the two-way, modulo 2 adder IC, perhaps the simplest example of a two-way IC in which adaptation is useless, and capacity is achieved through time-sharing.

Theorem 6. The capacity region of the two-way modulo 2 adder interference channel is the set of non-negative rate tuples \((R_{12}, R_{21}, R_{34}, R_{43})\) such that

\[
R_{12} + R_{34} \leq 1 \quad (2.10)
\]

\[
R_{21} + R_{43} \leq 1. \quad (2.11)
\]

Proof. We may achieve this region using two time-sharing random variables; one between nodes 1 and 3, and a second between nodes 2 and 4. The converse follows by the cut-set bound, or may alternatively be derived as done in the next subsection for a more general class of channels. \(\square\)
2.4.2 A More General Class of Two-way Deterministic ICs

We ask whether the above two-way modulo 2 IC results may be extended to a more general class of deterministic ICs in which adaptation is useless and capacity is achieved through time-sharing. In both the MAC/BC and Z channel models we were able to accomplish this by imposing certain cardinality, invertibility and symmetry constraints. One example of a channel in this class is the modulo-$\kappa$ (for some $\kappa$) channel. To extend results to the two-way IC we make two additional restrictions: 1) we do not consider “self-interference”, and 2) we impose symmetry of the outputs (common output in each direction). Both of these conditions are sufficient for obtaining sum-rate outer bounds equal to $\log \kappa$ in each direction (where $\kappa$ is the input/output alphabet size); whether they are necessary remains open.

Consider a class of deterministic two-way interference channels without self-interference, described by:

\[
Y_1 = F_{\rightarrow}(X_2, X_4) = Y_3 \quad \text{(there is no self-interference, symmetric channel)}
\]
\[
Y_2 = F_{\leftarrow}(X_1, X_3) = Y_4 \quad \text{(there is no self-interference, symmetric channel)}
\]

where $F_{\rightarrow}, F_{\leftarrow}$ are deterministic functions. Further restrict the class of channels to those with:

- **P1IC**: $|X_1| = |X_2| = |X_3| = |X_4| = |Y_1| = |Y_2| = |Y_3| = |Y_4| = \kappa$ for known $\kappa \in \mathbb{N}^+$.

- **P2IC**: “Invertibility” constraints reminiscent of Costa and El Gamal [37]. In the notation of [37], we assume $f_1 = f_2 = F_{\rightarrow}$ (and similarly, in the reverse direction we have $f_1 = f_2 = F_{\leftarrow}$), and that $g_1 = g_2$ are the identity functions, i.e. $g_1(X_1) = X_1$ and $g_2(X_3) = X_3$ (and
similarly for the reverse direction). Then we require that, given $X_1$, $Y_2$ is invertible, i.e. 
$\exists$ a function $G_2$ s.t. $X_3 = G_2(X_1, Y_2)$. Similarly, we assume $\exists G_1, G_3, G_4$: $X_4 = G_1(X_2, Y_1)$, $X_2 = G_3(X_4, Y_3)$, and $X_1 = G_4(X_3, Y_4)$.

- **P3IC**: to ensure the outer bound is achievable through time-sharing, we impose that $F_{\rightarrow}$ is a function such that $\exists x_3^*$ such that $X_1$ and $Y_2 = Y_4 = F_{\rightarrow}(X_1, X_3 = x_3^*)$ are in 1-to-1 correspondence, and $\exists x_1^*$ such that $X_3$ and $Y_2 = Y_4 = F_{\rightarrow}(X_1 = x_1^*, X_3)$ are in 1-to-1 correspondence, (and similarly for $F_{\leftarrow}$).

For this class of channels, the capacity is given by the following:

**Theorem 7.** The capacity region of the two-way “deterministic, invertible and alphabet restricted” IC satisfying the conditions **P1IC**, **P2IC**, **P3IC** is the set of non-negative rates $(R_{12}, R_{34}, R_{21}, R_{43})$ satisfying:

\[
R_{12} + R_{34} \leq \log \kappa 
\]

(2.12)

\[
R_{21} + R_{43} \leq \log \kappa, 
\]

(2.13)

which may be achieved via time-sharing (in each direction).

**Proof.** Consider the $\rightarrow$ direction. Under the above restrictions, the capacity region of the class of deterministic (one-way) interference channels in [37] may be simplified to

\[
R_{12} + R_{34} \leq \log \kappa
\]

(2.14)
which may be achieved by time-sharing between the inputs $X_1$ uniform over the $\kappa$ input symbols, while $X_3 = x_3^*$ and vice versa. That the rates (Equation 2.14) are achievable may alternatively be directly verified.

We find the matching outer bound:

$$n(R_{12} + R_{34} - \epsilon)$$

$$\leq I(M_{12}; Y_2^n | M_{21}, M_{43}) + I(M_{34}; Y_4^n | M_{12}, M_{21}, M_{43})$$

$$\leq \sum_{i=1}^{n} [H(Y_{2,i} | Y_2^{i-1}, M_{21}, M_{43}) - H(Y_{2,i} | Y_2^{i-1}, M_{12}, M_{21}, M_{43}) + H(Y_{2,i} | Y_2^{i-1}, M_{12}, M_{21}, M_{43})]$$

$$\leq \sum_{i=1}^{n} [H(Y_{2,i})] \leq n \log \kappa,$$

where in (a) we dropped two negative entropy terms, and were able to replace $Y_{4,i}$ by $Y_{2,i}$, allowing us to cancel the 2nd and 3rd terms. This is the central reason why we have restricted $Y_2 = Y_4$ and $Y_1 = Y_3$, whether one may somehow cancel these terms when this is not the case is open. Restricting the alphabet size as in $\text{P1IC}$ yields the final inequality.

$\square$

**Remark 8.** We have proposed a slightly more general model for deterministic two-way ICs in which adaptation is useless. Our conditions are sufficient but not necessary. For instance, in the binary multiplier two-way interference channel described by $Y_1 = Y_3 = X_2X_4$ and $Y_2 = Y_4 = X_1X_3$, with all inputs and outputs binary. It is not difficult to show that adaptation is useless for this model and the capacity of this channel is equivalent to the capacity of two one-way binary multiplier interference channels in parallel, the same capacity region as in Theorem 6.
In addition, we will show in Section 3.2 that adaptation is also useless for the Gaussian two-way interference channel with partial adaptation when the two-way interference is very strong; this channel is not in the class of channels considered above either.

2.4.3 Linear Deterministic Two-way IC

The two-way linear deterministic interference channel is defined by the input / output equations in Figure 2(c). In this section we will be considering the general linear deterministic IC, as well as the “symmetric” linear deterministic IC for which $p := n_{12} = n_{21} = n_{34} = n_{43}$, $q := n_{14} = n_{41} = n_{23} = n_{32}$, and $\alpha := q/p$. This will allow us to compare the symmetric, normalized sum capacity of various one and two-way interference channels, defined as $C_{\text{sym}}(\alpha) := \frac{R_{12} + R_{34}}{2}$.

Recall our definition of partial adaptation (nodes 1 and 3 are fixed or “restricted”) of Section 2.1:

\begin{align*}
X_{1,i} &= f_1(M_{12}) , \quad X_{2,i} = f_2(M_{21}, Y_{2}^{i-1}) \quad (2.15) \\
X_{3,i} &= f_3(M_{34}) , \quad X_{4,i} = f_4(M_{43}, Y_{4}^{i-1}) \quad (2.16)
\end{align*}

We first prove a Lemma, which is key in showing that partial adaptation is useless, and that the inability of certain nodes to adapt essentially “blocks” the ability of adaptation to help at all.
Lemma 8. Under partial adaptation conditions (Equation 2.15) – (Equation 2.16), for some deterministic functions \( f_5 \) and \( f_6 \),

\[
X_{2,i} = f_5(M_{12}, M_{21}, M_{34}) \perp M_{43}, \quad \forall i
\quad (2.17)
\]

\[
X_{4,i} = f_6(M_{43}, M_{34}, M_{12}) \perp M_{21}, \quad \forall i
\quad (2.18)
\]

where \( \perp \) denotes independence.

Proof. Note that \( X_{2,i} = f_2(M_{21}, Y_i^{i-1}) \) and \( Y_i^{i-1} = S^{N-n_{12}}X_i^{i-1} + S^{N-n_{21}}X_i^{i-1} + S^{N-n_{32}}X_i^{i-1} \). Since \( X_i^{i-1} \) and \( X_i^{j-1} \) are functions only of \( M_{12} \) and \( M_{34} \) respectively, we may conclude that there exists a function \( f^* \) such that \( X_{2,i} = f^*(M_{21}, M_{12}, M_{34}, X_i^{i-1}) \). Iterating this argument, and noting that \( X_{2,1} \) is only a function of \( M_{21} \), we obtain the lemma. The result for \( X_{4,i} \) follows by a similar argument. \( \square \)

Theorem 9. The capacity region of the two-way linear deterministic interference channel under partial adaptation constraints is the set of \((R_{12}, R_{21}, R_{34}, R_{43})\) which satisfy the following:

\[
R_{12} \leq n_{12}, \quad R_{34} \leq n_{34} \quad \text{(IC→ a)}
\]

\[
R_{12} + R_{34} \leq \max(n_{12}, n_{32}) + [n_{34} - n_{32}]^+ \quad \text{(IC→ b)}
\]

\[
R_{12} + R_{34} \leq \max(n_{34}, n_{14}) + [n_{12} - n_{14}]^+ \quad \text{(IC→ c)}
\]

\[
R_{12} + R_{34} \leq \max([n_{12} - n_{14}]^+, n_{32}) + \max([n_{34} - n_{32}]^+, n_{14}) \quad \text{(IC→ d)}
\]
2R_{12} + R_{34} \leq \max(n_{12}, n_{32}) + [n_{12} - n_{14}]^+ + \max([n_{34} - n_{32}]^+, n_{14}) \quad (IC \rightarrow e)

R_{12} + 2R_{34} \leq \max(n_{34}, n_{14}) + [n_{34} - n_{32}]^+ + \max([n_{12} - n_{14}]^+, n_{32}) \quad (IC \rightarrow f)

R_{21} \leq n_{21}, \quad R_{43} \leq n_{43} \quad (IC \leftarrow a)

R_{21} + R_{43} \leq \max(n_{21}, n_{41}) + [n_{43} - n_{41}]^+ \quad (IC \leftarrow b)

R_{21} + R_{43} \leq \max(n_{43}, n_{23}) + [n_{21} - n_{23}]^+ \quad (IC \leftarrow c)

R_{21} + R_{43} \leq \max([n_{21} - n_{23}]^+, n_{41}) + \max([n_{43} - n_{41}]^+, n_{23}) \quad (IC \leftarrow d)

2R_{21} + R_{43} \leq \max(n_{21}, n_{41}) + [n_{21} - n_{23}]^+ + \max([n_{43} - n_{41}]^+, n_{23}) \quad (IC \leftarrow e)

R_{21} + 2R_{43} \leq \max(n_{43}, n_{23}) + [n_{43} - n_{41}]^+ + \max([n_{21} - n_{23}]^+, n_{41}) \quad (IC \leftarrow f)

**Proof.** For achievability, note that self-interference may be removed at each receiver due to this channel model’s linearity, in which case the physical channel model reduces to two one-way IC in opposite directions. We may thus apply the well-known Han-Kobayashi scheme [63] in each direction, ignoring the ability of the nodes to adapt, achieving the expression in (IC$\rightarrow$) and (IC$\leftarrow$).

Now we prove the converse. Single-rates follow as in (Equation 2.6), and using Lemma 8 (where we use partial adaptation). For the sum-rate (IC$\rightarrow$ b):

\[
\begin{align*}
n(R_{12} + R_{34} - \epsilon) \\
\quad \overset{(a)}{\leq} I(M_{12}; Y_2^n | M_{21}, M_{43}) + I(M_{34}; Y_4^n, Y_2^n | M_{12}, M_{21}, M_{43})
\end{align*}
\]
\[ \begin{align*}
\leq & \ I(M_{12}; Y^n_2 | M_{21}, M_{43}) + I(M_{34}; Y^n_2 | M_{21}, M_{12}, M_{43}) + H(Y^n_4 | M_{21}, M_{12}, M_{43}, Y^n_2) \\
\overset{(b)}{=} & \ I(M_{12}; Y^n_2 | M_{21}, M_{43}) + I(M_{34}; Y^n_2 | M_{21}, M_{12}, M_{43}) \\
& + \sum_{i=1}^{n} [H(S^{N-n_{34}}X^n_3, | M_{21}, M_{12}, M_{43}, Y^{i-1}_4, X^n_4, Y^n_2, X^n_2, X^n_1)] \\
\leq & \sum_{i=1}^{n} [H(Y^{i-1}_2, | M_{21}, X^n_2) - H(Y^{i-1}_2, | Y^{i-1}_2, M_{12}, M_{21}, M_{43}) + H(Y^{i-1}_2, | Y^{i-1}_2, M_{12}, M_{21}, M_{43})] \\
& + H(S^{N-n_{34}}X^n_3, | M_{21}, M_{12}, M_{43}, Y^{i-1}_4, X^n_4, S^{N-n_{12}}X^n_1 + S^{N-n_{32}}X^n_2, X^n_2, X^n_1)] \\
\leq & \sum_{i=1}^{n} [H(S^{N-n_{12}}X^n_1, + S^{N-n_{32}}X^n_3, + S^{N-n_{32}}X^n_3, | S^{N-n_{32}}X^n_3)] \\
\leq & \ n(\max(n_{12}, n_{32}) + [n_{34} - n_{32}]^+) \\
\overset{(c)}{=} & \ n(\max(p, q) + [p - q]^+) .
\end{align*} \]

We introduce the genie \( Y^n_2 \) in the second mutual information term in (a), i.e. we provide asymmetric side information to only one receiver. In (b), we add \( X^n_1 \) in the entropy term because of the iterated argument that, given \( M_{12}, X^n_2, X^n_4 \), we can construct \( X^n_1 \). For (c), we assumed a symmetric channel.

**Remark 9.** Note that we do not need partial adaptation in this bound, and so these conclusions actually hold for full adaptation. This implies that for the symmetric channel, full adaptation is useless when two-way interference is strong \((1 \leq \alpha \leq 2, \alpha = q/p)\) and weak in some interval \((2/3 \leq \alpha \leq 1, \alpha = q/p)\). Interestingly, when \(2/3 \leq \alpha \leq 2\), the “V” curve is also the capacity for the linear deterministic symmetric interference channel with feedback [40]. If we add another
asymmetric genie $Y_4^n$ in the first term in (a), then we obtain the second sum-rate bound ($IC \rightarrow c$).

It may further be shown that for symmetric channels, adaptation is also useless when two-way interference is very strong ($\alpha > 2, \alpha = q/p$). To show this, we re-derive the single-rate bounds this time not assuming partial adaptation (allowing for full adaptation), and using symmetry in the last step:

$$n(R_{12} - \epsilon) \leq I(M_{12}; Y_2^n, Y_3^n | M_{21}, M_{34}) \leq H(Y_2^n, Y_3^n | M_{21}, M_{34})$$

$$= \sum_{i=1}^{n} [H(Y_{2,i}, Y_{3,i} | Y_{2}^{i-1}, Y_{3}^{i-1}, M_{21}, M_{34}, X_2^{i}, X_3^{i})]$$

$$= \sum_{i=1}^{n} [H(S_{N-n12} X_{1,i}, S_{N-n43} X_{4,i} | Y_{2}^{i-1}, Y_{3}^{i-1}, M_{21}, M_{34}, X_2^{i}, X_3^{i})]$$

$$\leq \sum_{i=1}^{n} [H(S_{N-n12} X_{1,i}, S_{N-n43} X_{4,i})] = n \max(n_{12}, n_{43}) = np$$

Under very strong interference constraints, this is also known to be achievable. Thus, we have obtained the capacity for the symmetric linear deterministic two-way IC when $\alpha \geq 2/3$, where we see that full adaptation is useless. We will comment more on this in subsection 2.4.4, and in Figure 4.

We now continue with the sum-rate outer bound ($IC \rightarrow d$), which uses a similar genie to that in Costa and El Gamal’s [37] capacity result for a class of deterministic ICs, i.e. gives to one receiver the interference created at the other receiver by the desired message. The same type of genie (though this time noisy) is used in the new outer bound for the Gaussian one-
way interference channel by Etkin, Tse and Wang [21]. The main difference is that we also provide the transmitters in the $\rightarrow$ direction the messages in the $\leftarrow$ direction, or $M_{21}$ and $M_{43}$, in order to be able to create $X_2$ and $X_4$ and remove these from the entropy terms, obtaining only entropies of combinations of the variables in the $\rightarrow$ direction.

$$n(R_{12} + R_{34} - \epsilon) \leq I(M_{12}; Y_2^n, S^{N-n_{14}}X_1^n, M_{21}, M_{43}) + I(M_{34}; Y_4^n, S^{N-n_{32}}X_3^n, M_{21}, M_{43})$$

$$= H(Y_2^n | S^{N-n_{14}}X_1^n, M_{43}, M_{21}) + H(S^{N-n_{14}}X_1^n | M_{43}, M_{21}) - H(Y_2^n, S^{N-n_{14}}X_1^n | M_{12}, M_{21}, M_{43})$$

$$+ H(Y_4^n | S^{N-n_{32}}X_3^n, M_{43}, M_{21}) + H(S^{N-n_{32}}X_3^n | M_{43}, M_{21}) - H(Y_4^n, S^{N-n_{32}}X_3^n | M_{34}, M_{21}, M_{43})$$

$$= H(Y_2^n | S^{N-n_{14}}X_1^n, M_{43}, M_{21}) + H(Y_4^n | S^{N-n_{32}}X_3^n, M_{43}, M_{21})$$

$$+ \sum_{i=1}^{n} H(S^{N-n_{14}}X_{1,i} | S^{N-n_{14}}X_{1,i}^{i-1}, M_{43}, M_{21}, M_{34}) - H(Y_{2,i} | Y_2^{i-1}, M_{12}, M_{21}, M_{43}, X_2^i, X_1^i)$$

$$+ \sum_{i=1}^{n} H(S^{N-n_{32}}X_{3,i} | S^{N-n_{32}}X_{3,i}^{i-1}, M_{43}, M_{21}, M_{12}) - H(Y_{4,i} | Y_4^{i-1}, M_{34}, M_{21}, M_{43}, X_4^i, X_3^i)$$

$$= H(Y_2^n | S^{N-n_{14}}X_1^n, M_{43}, M_{21}) + H(Y_4^n | S^{N-n_{32}}X_3^n, M_{43}, M_{21})$$

$$+ \sum_{i=1}^{n} H(S^{N-n_{14}}X_{1,i} | S^{N-n_{14}}X_{1,i}^{i-1}, M_{43}, M_{21}, M_{34}, X_1^i, X_3^i)$$

$$- H(S^{N-n_{32}}X_{3,i} | S^{N-n_{32}}X_{3,i}^{i-1}, M_{43}, M_{21}, M_{43}, X_2^i, X_3^i)$$

$$+ \sum_{i=1}^{n} H(S^{N-n_{32}}X_{3,i} | S^{N-n_{32}}X_{3,i}^{i-1}, M_{43}, M_{21}, M_{12}, X_1^i, X_2^i)$$

$$- H(S^{N-n_{14}}X_{1,i} | S^{N-n_{14}}X_{1,i}^{i-1}, M_{34}, M_{21}, M_{43}, X_3^i, X_1^i)$$

$$\leq \sum_{i=1}^{n} [H(S^{N-n_{12}}X_{1,i} + S^{N-n_{32}}X_{3,i} | S^{N-n_{14}}X_{1,i}) + H(S^{N-n_{14}}X_{1,i} + S^{N-n_{34}}X_{3,i} | S^{N-n_{32}}X_{3,i})]$$
\[ \leq n(\max([n_{12} - n_{14}]^+, n_{32}) + \max([n_{34} - n_{32}]^+, n_{14})), \]

where (d) follows from the independence of the messages, (e) by the chain rule of entropy and the fact that we can create \( X_i^1 \) given \( M_{12} \), and we can create \( X_i^2 \) given \( M_{21}, Y_{i-1}^2 \) (similarly for \( X_3^1, X_4^1 \)). We have also added the independent messages \( M_{34} \) and \( M_{12} \) to the 3rd and 5th terms' conditioning. For (f), we have expanded the \( Y_{2,i} \) and \( Y_{4,i} \) in the entropy terms of the 4th and 6th terms and removed the contributions from the conditioning. In the 3rd term, we can create \( X_3^1 \) from \( M_{34} \) and \( X_4^1 \) from \( S_{N-n_{14}}X_{i-1}^i, X_3^1 \) and \( M_{43} \) (similarly for the 5th term creating \( X_1^1 \) and \( X_2^1 \)).

The sum-rate bound in the opposite direction may be derived in a similar fashion, despite the asymmetry caused by the partial adaptation condition.

**Remark 10.** We needed partial adaptation at nodes 1 and 3 (Lemma 8) to show bounds \( (IC \rightarrow d) \) and \( (IC \leftarrow d) \). By symmetry, we may obtain the same result if nodes 2 and 4 were restricted.

Finally,

\[ n(2R_{12} + R_{34} - \epsilon) \]

\[ \leq I(M_{12}; Y_{2}^n | M_{21}, M_{43}) + I(M_{12}; Y_{4}^n | M_{21}, M_{43}, M_{43}) + I(M_{34}; Y_{4}^n, S_{N-n_{12}}X_{3}^n | M_{21}, M_{43}) \]

\[ = H(Y_{2}^n | M_{21}, M_{43}) - H(Y_{2}^n | M_{21}, M_{43}, M_{12}) + H(Y_{4}^n | M_{21}, M_{43}, M_{43}) \]

\[ + H(Y_{2}^n | M_{21}, M_{43}, M_{34}, Y_{4}^n) + H(Y_{4}^n, S_{N-n_{12}}X_{3}^n | M_{21}, M_{43}) - H(Y_{4}^n, S_{N-n_{12}}X_{3}^n | M_{34}, M_{21}, M_{43}) \]

\[ \overset{(h)}{=} H(Y_{2}^n | M_{21}, M_{43}) - H(Y_{2}^n | M_{21}, M_{43}, M_{12}) + H(S_{N-n_{12}}X_{3}^n | M_{43}, M_{21}, M_{12}) \]
\[ + H(Y^n_4|S^{N-n_{12}}X^n_3, M_{43}, M_{21}) + H(Y^n_4|M_{21}, M_{43}, M_{34}) \]

\[- H(Y^n_4, S^{N-n_{12}}X^n_3|M_{34}, M_{21}, M_{43}) + H(Y^n_4|M_{21}, M_{43}, M_{34}, Y^n_4) \]

\[ \leq \sum_{i=1}^{n} \left[ H(S^{N-n_{12}}X_{1,i} + S^{N-n_{32}}X_{3,i}) + H(S^{N-n_{14}}X_{1,i} + S^{N-n_{34}}X_{3,i} | S^{N-n_{12}}X_{3,i}) \right] \]

\[ + H(S^{N-n_{12}}X_{1,i} | S^{N-n_{14}}X_{1,i}) \]

\[ = n(\max(n_{12}, n_{32}) + \max([n_{34} - n_{32}]^+, n_{14}) + [n_{12} - n_{14}]^+), \]

where (h) follows from the definition of partial adaptation and Lemma 8 (skipping a transition to multi-letter for brevity), and (i) by canceling the 2nd and 3rd terms, as well as the 5th and the 6th terms. We may similarly prove the other bounds of this form (IC→ f), (IC← e) and (IC← f).

We again see that, under partial adaptation constraints, adaptation is useless and we obtain the capacity region of two one-way ICs. Essentially, partial adaptation prevented messages being relayed by other messages (which was also impossible in the MAC/BC and Z channels). For example, under full adaptation, message \( M_{12} \) may be relayed from Tx1 to Rx 2 through nodes 3 and 4. This path is “blocked” by the partial adaptation assumption, as node 3 could not adapt to carry \( M_{12} \). However, it should be pointed out that this is not necessary in general: full adaptation in the two-way modulo 2 adder IC is useless as we showed in the previous subsection, but the path is not blocked.
2.4.4 Symmetric Rate Comparison with Other IC Models

For symmetric linear deterministic ICs, we may compare the symmetric sum-capacity $C_{sym}$ of various one-way and two-way models. Recalling $\alpha := q/p$, we plot $C_{sym}$ as a function of $\alpha$ for the IC [21], the IC with noiseless output feedback [40], the IC with rate-limited feedback [41] (for a fixed value of $\beta = 0.125$ in the notation of [41]), and the two-way IC with full adaptation considered here (for $\alpha \geq 2/3$ only). Several observations may be made: the two-way IC with partial adaptation behaves like two one-way interference channels operating in parallel over the forward and backwards link. This tells us that allowing partial adaptation is useless – i.e. may as well not adapt. Interestingly, the same holds true even for full adaptation for $\alpha > 2/3$. This was also concluded for the linear deterministic one-way interference channel with interfering feedback links in [39]; what is interesting is that we can just as well squeeze in extra information messages in the feedback link (in the two-way interference channel model) rather than use the backwards links for feedback. The symmetric sum-capacity for the fully adaptive two-way IC remains open for $\alpha < 2/3$; it is solved for partial adaptation.

Recently, the work in [43,44] has considered a one-way interference channel with interfering feedback links (again forming an interference channel), a generalization of some of the deterministic interference channels with feedback considered in [39], where the feedback link spends fraction $\lambda$ of its time sending feedback, and uses the remaining $(1-\lambda)$ for other things (such as for example sending independent backwards messages, though adaptation as in (Equation 2.1) is not considered). This differs from our model which integrates sending feedback and messages over all links, allows for adaptation, and does not force this separation. Because of these dif-
ferences, and since the symmetric sum-capacity in [43,44] in our notation for $\alpha \geq 2/3$, it is a function of this parameter $\lambda$, it is not plotted here.

2.5 Summary

In this chapter, we have demonstrated a few examples of wireless deterministic two-way networks for which adaptation, or the ability of nodes to adapt their current channel inputs based on previously received channel outputs, is useless from a capacity region perspective, i.e. non-adaptive schemes achieve outer bounds derived for the fully adaptive models. Specifically,
we obtained the capacity regions of the two-way MAC/BC channel, the two-way Z channel, and the two-way IC of binary modulo-2 addition model, the “deterministic, invertible and cardinality constrained” model, and the linear deterministic model. Interestingly, adaptation (full or partial) is not needed to attain the capacity regions even though it is permitted. In other words, the capacity regions of considered deterministic two-way networks can be decomposed, i.e. they are equivalent to the capacity regions of corresponding one-way networks in parallel.
CHAPTER 3

WIRELESS GAUSSIAN TWO-WAY NETWORKS

All channel models considered in the previous chapter were deterministic. We now ask
whether we may obtain insight into whether adaptation is useless / useful in certain noisy
channels. We do so by considering the Gaussian two-way MAC/BC, the Gaussian two-way IC,
and an extension of that, the K-pair-user Gaussian two-way IC in this chapter.

3.1 Gaussian Two-way Multiple-Access Broadcast Channel

We demonstrate that adaptation in the real Gaussian two-way MAC/BC with independent
noises can only improve the sum-capacity up to $1/2$ bit per direction. We show this by com-
paring non-adaptive inner bounds for this channel to outer bounds to the two-way Gaussian
MAC/BC. Our outer bound for the $\rightarrow$ MAC direction is derived directly; the outer bound for
the $\leftarrow$ BC direction follows by enhancing the BC channel by giving Tx 2 perfect output feedback
and rendering the channel degraded, at which point the converse of [35, Thm.2] follows.

3.1.1 Channel Model

At each channel use, the Gaussian two-way MAC/BC is described by the input/output
relationships

\[ Y_1 = X_2 + Z_1 \]
\[ Y_2 = X_1 + X_3 + Z_2 \]
\[ Y_3 = X_2 + Z_3, \]

subject to power constraints \( \mathbb{E}[|X_j|^2] \leq P_j, j \in \{1, 2, 3\} \), and independent, identically distributed complex Gaussian noise \( Z_j \sim CN(0, N_j) \) at all nodes \( j \in \{1, 2, 3\} \); WLOG assume that \( N_3 \geq N_1 \).

Note that we have removed the “self-interference” terms such as \( X_1 \) in the expression of \( Y_1 \) (for example) in contrast to the deterministic models considered. This is for ease of exposition, to make the parallels with the MAC and BC channels more direct. Note that in a Gaussian (additive) model, these “self-interference” terms can always be subtracted at a given node in any case. In contrast, in the Gaussian two-way interference channel, which is studied in next section, we will NOT eliminate the self-interference terms from the channel model, to demonstrate how they may be handled directly. Recall that the inputs of the Gaussian two-way MAC/BC are fully adaptive, i.e.

\[
X_{1,i} = f_1(M_{12}, Y_i^{(1)}), \quad X_{2,i} = f_2(M_{21}, M_{23}, Y_i^{(2)}), \quad X_{3,i} = f_3(M_{32}, Y_i^{(3)}). 
\]

(3.1)

### 3.1.2 The Limited Utility of Adaptation in the Gaussian Two-way MAC/BC

We now have the following theorem.

**Theorem 10.** *Adaptation in the Gaussian MAC/BC channel may only improve the sum-rate in the \( \rightarrow \) and \( \leftarrow \) directions by up to \( 1/2 \) bit per direction.*
Proof. First let us consider the → direction. For achievability, let the → direction use the capacity achieving scheme for the non-adaptive Gaussian MAC, whose sum-rate is dominated by
\[ R_{12} + R_{32} \leq \frac{1}{2} \log \left( 1 + \frac{P_1 + P_3}{N_2} \right). \] (3.2)

For the converse, first consider the MAC direction, and follow steps along the lines of a MAC with feedback as in [34,64]:

\[ n(R_{12} + R_{32}) = H(M_{12}, M_{32}) = H(M_{12}, M_{32}, M_{21}, M_{23}) \]
\[ = H(M_{12}, M_{32}, M_{21}, M_{23}, Y_2^n) + I(M_{12}, M_{32}; Y_2^n | M_{21}, M_{23}) \] (3.3)
\[ \leq n \epsilon_n + \sum_{i=1}^{n} H(Y_{2,i}|Y_{2,i}^{i-1}, M_{21}, M_{23}) - H(Y_{2,i}|Y_{2,i}^{i-1}, M_{12}, M_{32}, M_{21}, M_{23}, Y_1^n, Y_3^n) \] (3.4)
\[ \leq n \epsilon_n + \sum_{i=1}^{n} H(Y_{2,i}|X_{1,i}, X_{2,i}, X_{3,i}) \] (3.5)
\[ \leq n \epsilon_n + \sum_{i=1}^{n} H(X_{1,i} + X_{3,i} + Z_{2,i}) - H(Z_{2,i}) \] (3.6)
\[ \leq n \epsilon_n + \frac{n}{2} \log \left( 1 + \frac{P_1 + P_3 + 2\sqrt{P_1P_3}}{N_2} \right) \] (3.7)

where (a) follows by Fano’s inequality for the first term, by the chain rule of entropy for the 2nd and 3rd terms, and by conditioning reduces entropy in adding \( Y_1^n \) and \( Y_3^n \) to the 3rd term, (b) since for the 2nd term, given \( M_{21}, M_{23} \) and \( Y_2^{i-1} \) one can construct \( X_{2,i} \) and then conditioning reduces entropy, and for the 3rd term since given all the terms in the conditioning we may create \( X_{1,i}, X_{2,i}, \) and \( X_{3,i} \) and then use the memoryless property of the channel model, (c) follows by
conditioning reduces entropy and by the memoryless channel, (d) since it suffices to consider $X_1, X_3$ to be jointly Gaussian and is outer bounded when they are maximally correlated, as adaptation may permit joint $p(x_1, x_3)$.

Now, taking the difference between the outer bound to the adaptive two-way MAC/BC in the MAC direction in (Equation 3.8) and the non-adaptive inner bound of (Equation 3.2) yields

$$(\text{Equation 3.8}) - (\text{Equation 3.2}) = \frac{1}{2} \log \left( 1 + \frac{2\sqrt{P_1P_3}}{N_2 + P_1 + P_3} \right)$$

$$\leq \frac{1}{2} \log \left( 1 + \frac{P_1 + P_3}{P_1 + P_3} \right) = \frac{1}{2},$$

where the inequality follows as $2\sqrt{P_1P_3} \leq P_1 + P_3$, and we have decreased the denominator.

For the $\leftarrow$ direction use the capacity achieving scheme for the non-adaptive single-antenna Gaussian broadcast channel, which yields the rates, for $0 \leq \alpha \leq 1$

$$R_{21} \leq \frac{1}{2} \log \left( 1 + \frac{\alpha P_2}{N_1} \right), \quad R_{23} \leq \frac{1}{2} \log \left( 1 + \frac{(1 - \alpha)P_2}{N_3 + \alpha P_2} \right), \quad \text{(3.9)}$$

For the converse, for the BC $\leftarrow$ direction we enhance the channel as follows:

- Give Tx 2 perfect output feedback, i.e. access to $Y_{1,i-1}, Y_{3,i-1}$ at time $i$ as well as access to $M_{12}, M_{32}$. Together with feedback, this allows it to create $X_{1,i}^1, X_{3,i}^1$.

- Render the channel physically degraded by providing Rx 1 with $Y_{3}^{\alpha}$. Then Rx 3’s output is trivially a physically degraded version of Rx 1’s output. This is where we use the fact
that, WLOG $N_3 \geq N_1$ (if the reverse had been true we would have given $Y_1^n$ to Rx 3 instead). This is crucial in ensuring a constant gap to a non-adaptive scheme.

The converse of [35, Thm. 2], which shows that feedback does not change the capacity region of the physically degraded BC, then follows along all the same steps with the notation correspondences

( [35] ↔ this paper) as follows:

$$W_1 \leftrightarrow M_{21}, \; W_2 \leftrightarrow M_{23}, \; Y^n \leftrightarrow (Y^n_1, Y^n_3), \; Z^n \leftrightarrow Y^n_3, \; X^n \leftrightarrow X^n_2$$

The key point in proving the converse is [35, Lemma 3], which follows in a straightforward manner even given the added adaptation constraint (i.e. $X_{2,i}$ is also a function of $Y_{2}^{i-1}$ which is not present in the original [35, Thm. 2]), but we re-state and prove it here in our notation for clarity and completeness.

**Lemma 11.** Analogous to Lemma 3 of [35]. For all $\lambda \geq 0$,

$$n(R_{21} + R_{23}) \leq I(M_{23}; Y^n_3) + \lambda I(M_{21}; Y^n_1, Y^n_3 | M_{23})$$

$$\leq \sum_{i=1}^{n} I(U_i; Y_{3,i}) + \lambda I(X_{2,i}; Y_{1,i}, Y_{3,i} | U_i)$$

where $U_i := (M_{23}, Y_{1}^{i-1}, Y_{3}^{i-1})$. 
Proof. For the first term,

\[ I(M_{23}; Y_3^n) = \sum_{i=1}^{n} I(M_{23}; Y_{3,i}^{i-1}) = \sum_{i=1}^{n} H(Y_{3,i}^{i-1}) - H(Y_{3,i}^{i-1}|M_{23}, Y_3^{i-1}) \]
\[ \leq \sum_{i=1}^{n} H(Y_{3,i}) - H(Y_{3,i}|M_{23}, Y_3^{i-1}, Y_1^{i-1}) = \sum_{i=1}^{n} I(Y_{3,i}; U_i) \]

by definition of \( U_i := (M_{23}, Y_3^{i-1}, Y_1^{i-1}) \). For the second term,

\[ I(M_{21}; Y_1^n, Y_3^n|Y_{23}) = \sum_{i=1}^{n} I(M_{21}; Y_{1,i}, Y_{3,i}|M_{23}, Y_1^{i-1}, Y_3^{i-1}) = \sum_{i=1}^{n} I(M_{21}; Y_{1,i}, Y_{3,i}|U_i) \]
\[ \leq \sum_{i=1}^{n} I(M_{21}, X_{2,i}; Y_{1,i}, Y_{3,i}|U_i) = \sum_{i=1}^{n} H(Y_{1,i}, Y_{3,i}|U_i) - H(Y_{1,i}, Y_{3,i}|U_i, M_{21}, X_{2,i}) \]
\[ = \sum_{i=1}^{n} I(Y_{1,i}, Y_{3,i}; X_{2,i}|U_i) \]

where several steps in the proof of [35, Lemma 3] are not needed as our channel is trivially degraded.

Following the same arguments as in [35, Thm. 2], the above Lemma yields an outer bound equivalent to the region in (Equation 3.10), where we note that in addition to \( U_i = (M_{23}, Y_1^{i-1}, Y_3^{i-1}) \) to construct

\[ X_{2,i} = f(M_{21}, M_{23}, Y_2^{i-1}, M_{12}, M_{32}, Y_1^{i-1}, Y_3^{i-1}) = f(U_i, M_{21}, Z_2^{i-1}, M_{12}, M_{32}) \]

we also need \( M_{12}, M_{32}, M_{21}, Z_2^n \), but that, given the above definition of the random variable \( U_i \), the factorization of the inputs as \( p(u)p(x_2|u) \) still holds. Note that with some abuse of
notation we have left the channel distribution as \( p(y_1, y_3|x_2)p(y_3|y_1, y_3) \) to emphasize that Rx 1 has access to both \( Y_1^n, Y_3^n \) (we have forced the channel to be degraded) and thus that Rx 3, with access to \( Y_3^n \) only is trivially a degraded version of this. The outer bound for the \( \leftarrow \) BC direction is thus given by the set of all non-negative \( R_{21}, R_{23} \) such that

\[
R_{21} \leq I(X_2; Y_1, Y_3|U), \quad R_{23} \leq I(U; Y_3)
\]

(3.10)

over all distributions of the form \( p(u)p(x_2|u)p(y_1, y_3|x_2)p(y_3|y_1, y_3) \). Evaluation for the Gaussian channel, as done in [65], yields an outer bound of

\[
R_{21} \leq \frac{1}{2} \log \left( 1 + \frac{\alpha P_2 N_1 + N_3}{N_1} \right), \quad R_{23} \leq \frac{1}{2} \log \left( 1 + \frac{(1-\alpha)P_2}{N_3 + \alpha P_2} \right)
\]

(3.11)

for \( 0 \leq \alpha \leq 1 \).

Taking the difference between the sum of the outer bound to the adaptive two-way MAC/BC in the BC direction in (Equation 3.11) and the sum of the non-adaptive inner bounds of (Equation 3.9) yields

\[
(Equation \ 3.11) - (Equation \ 3.9) = \frac{1}{2} \log \left( 1 + \frac{\alpha P_2 N_1 + N_3}{N_1} \right) - \frac{1}{2} \log \left( 1 + \frac{\alpha P_2}{N_1} \right)
\]

\[
\leq \frac{1}{2} \log \left( 1 + \frac{2\alpha P_2}{N_1} \right) - \frac{1}{2} \log \left( 1 + \frac{\alpha P_2}{N_1} \right) \leq \frac{1}{2}.
\]

where (a) follows as \( \frac{N_1 + N_3}{N_3} = 1 + \frac{N_1}{N_3} \leq 2 \) since \( N_1 \leq N_3 \). \( \square \)
Remark 11. We note that this result also implies that for the one-way Gaussian MAC with FB and the one-way BC with FB, feedback and adaptation of the nodes can only increase capacity by up to 1/2 bit (sum-rate) per direction. This fact has been partially noted in [40].

Remark 12. We also note that, in a similar vein to Theorem 1, in the Gaussian channel, in the MAC $\rightarrow$ direction, the two-way sum-rate $R_{12} + R_{32}$ is outer bounded by the capacity of the one-way MAC channel with perfect output feedback (once we provide the messages of the opposite direction, and output feedback as genies). Similarly, in the BC $\leftarrow$ direction, we first render the channel degraded (and provide messages in the opposite direction as genies) and then outer bound this direction by an outer bound to the one-way degraded BC with perfect output feedback. This raises the interesting question of whether each direction of two-way channels is in fact outer bounded by the one-way counterpart with perfect output feedback. We suspect so, at least in channels in which one’s own “self-interference” or signal may be cancelled, or in which some invertibility conditions hold.

3.2 Gaussian Two-way Interference Channel

We now consider the Gaussian two-way interference channel, and ask when non-adaptive schemes such as the celebrated Han and Kobayashi [63] perform as well, or nearly as well, as adaptive schemes. We do not construct any inner bounds which employ adaptation; our focus is on showing when non-adaptive schemes perform “well”. Rather, we derive an outer bound for the Gaussian two-way IC under full adaptation (all 4 nodes may adapt) and several under partial adaptation (only 2 of the 4 may adapt) constraints. We then show that non-adaptive schemes sometimes achieve the capacity, or at least to within a constant gap of either the fully or
partially adaptive schemes. We note that while the converses and the steps are new and exploit carefully chosen genies, when we evaluate these by further outer-bounding our outer-bounds, interestingly, we sometimes re-obtain some of the outer bounds of the interference channel [21] or the interference channel with feedback [40]. This in turn is sufficient to achieve capacity to within a constant gap, which we emphasize, sometimes is limited to partial adaptation and will be made explicit.

3.2.1 Channel Model, Definitions, and Partial Adaptation Lemma

At each channel use, the Gaussian two-way IC is described by the input/output relationships

\[ Y_1 = g_{11}X_1 + g_{21}X_2 + g_{41}X_4 + Z_1 \]
\[ Y_2 = g_{12}X_1 + g_{22}X_2 + g_{32}X_3 + Z_2 \]
\[ Y_3 = g_{23}X_2 + g_{33}X_3 + g_{43}X_4 + Z_3 \]
\[ Y_4 = g_{14}X_1 + g_{34}X_3 + g_{44}X_4 + Z_4, \]

where \( g_{jk}, \) for \( j, k \in \{1, 2, 3, 4\} \) are the complex channel gains. We assume the power constraints \( \mathbb{E}[|X_j|^2] \leq P_j = 1, j \in \{1, 2, 3, 4\}, \) and independent, identically distributed complex Gaussian noise \( Z_j \sim \mathcal{CN}(0, 1) \) at all nodes \( j \in \{1, 2, 3, 4\}. \) Define \( \text{SNR}_{12} = |g_{12}|^2, \text{SNR}_{21} = |g_{21}|^2, \text{SNR}_{34} = |g_{34}|^2, \text{SNR}_{43} = |g_{43}|^2, \) and \( \text{INR}_{14} = |g_{14}|^2, \text{INR}_{41} = |g_{41}|^2, \text{INR}_{23} = |g_{23}|^2, \text{INR}_{32} = |g_{32}|^2. \) Note that “self-interference” terms such as \( g_{11}X_1 \) are included in the expression of \( Y_1 \) (for example). In this Gaussian model, it is clear that since node 1 knows \( X_1 \) we may remove this self-interference term due to the additive nature of the channel. However, we leave it in our expressions to
emphasize precisely this fact. In other channels such as the two-way binary multiplier channel, where \( Y = X_1 X_2 \) one cannot “undo” ones’ own channel, which is one source of difficulty for this elusive two-way channel. In all converses, the fact that we can cancel or subtract out a node’s “self-interference” is shown explicitly.

We say that the Gaussian two-way interference channel operates under “full adaptation” if we allow

\[
X_{1,i} = f_1(M_{12}, Y_i^{1-1}), \quad X_{2,i} = f_2(M_{21}, Y_2^{1-1}) \tag{3.12}
\]

\[
X_{3,i} = f_3(M_{34}, Y_3^{1-1}), \quad X_{4,i} = f_4(M_{43}, Y_4^{1-1}) \tag{3.13}
\]

Similarly, it operates under “partial adaptation” if we only allow the following:

\[
X_{1,i} = f_1(M_{12}), \quad X_{2,i} = f_2(M_{21}, Y_2^{1-1}) \tag{3.14}
\]

\[
X_{3,i} = f_3(M_{34}), \quad X_{4,i} = f_4(M_{43}, Y_4^{1-1}) \tag{3.15}
\]

i.e. nodes 1 and 3 are “restricted” [15]. By symmetry, we may alternatively allow nodes 2 and 4 to be restricted and 1, 3 to be fully adaptive; whether allowing 1, 2 or 1, 4 to be restricted and the complement fully adaptive remains an open problem.

We are interested in the symmetric capacity (or sum-rate), when all the SNRs equal a given SNR, and all the INRs equal a given INR. For full adaptation, due to the symmetry, we consider the per-user rates \( R_{sym} = \frac{R_{12} + R_{34}}{2} = \frac{R_{21} + R_{43}}{2} \). In partial adaptation, there is only partial symmetry (nodes 1 and 3 are fixed, while 2 and 4 are not), and hence we will consider
the per user rates $R_{\text{sym} \rightarrow} = \frac{R_{12} + R_{43}}{2}$ and $R_{\text{sym} \leftarrow} = \frac{R_{21} + R_{34}}{2}$ for the forward and reverse directions respectively. We will derive outer bounds for $R_{\text{sym}}$ under full adaptation and $R_{\text{sym} \rightarrow}$, $R_{\text{sym} \leftarrow}$ under partial adaptation, and show these to be achievable to within constant gaps by non-adaptive schemes.

We first prove a modified version of Lemma 8 relevant in partial adaptation for the Gaussian channel.

**Lemma 12.** Under partial adaptation (Equation 3.14) – (Equation 3.15), for some deterministic functions $f_5$ and $f_6$,

\begin{align*}
X_{2,i} &= f_5(M_{12}, M_{21}, M_{34}, Z_{i-1}^2) \perp M_{43}, \quad \forall i \tag{3.16} \\
X_{4,i} &= f_6(M_{43}, M_{34}, M_{12}, Z_{i-1}^4) \perp M_{21}, \quad \forall i \tag{3.17}
\end{align*}

where $\perp$ denotes independence.

**Proof.** Note that $X_{2,i} = f_2(M_{21}, Y_{i-1}^2)$ and $Y_{i-1}^2 = g_{12}X_{i-1}^1 + g_{22}X_{i-1}^2 + g_{32}X_{i-1}^3 + Z_{i-1}^2$. Since $X_{i-1}^1$ and $X_{i-1}^3$ are functions only of $M_{12}$ and $M_{34}$ respectively, we may conclude that there exists a function $f^*$ such that $X_{2,i} = f^*(M_{21}, M_{12}, M_{34}, X_{i-1}^2, Z_{i-1}^2)$. Iterating this argument, and noting that $X_{2,1}$ is only a function of $M_{21}$, we obtain the lemma. The result for $X_{4,i}$ follows similarly. That $X_{2,i}$ is independent of $M_{43}$ follows since $M_{43}$ is independent of all the arguments inside $f^*$. \qed
3.2.2 Outer bounds

We now present two outer bounds for the Gaussian two-way IC under full and partial adaptation respectively. We derive general outer bounds, imposing symmetry only in the final step.

**Theorem 13.** Outer bound: full adaptation. For the Gaussian two-way symmetric IC under full adaptation, any achievable symmetric rate \( R_{sym} = \frac{R_{12} + R_{34}}{2} = \frac{R_{21} + R_{43}}{2} \), achievable by each user, satisfies,

\[
R_{sym} \leq \frac{1}{2} \log \left( 1 + \text{SNR} + \text{INR} + 2\sqrt{\text{SNR} \times \text{INR}} \right) + \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right). \tag{3.18}
\]

**Proof.** It is sufficient to consider \( R_{12} + R_{34} \) due to symmetry. This bound is inspired by the corresponding sum-rate bound in the linear deterministic model, i.e., we add asymmetric genie \( Y_2^n \) at node 4 and this resembles the bounding technique used by Suh and Tse for the interference channel with feedback [40]. We also note that we could have equivalently provided node 4 with the genie \( g_{32}X_3^n + Z_2^n \) instead of \( Y_2^n \) (in addition to \( M_{12}, M_{21}, M_{43} \) and \( Z_1^n \)) which more resembles the type of genie seen in ICs and ICs with feedback. We have given \( Y_2^n \) as it is then easier to see how node 4 may create \( X_2^n \) based on \( M_{21} \) and \( Y_2^n \), and the bounds work out to the same. Notice the genie \( Z_1^n \) in the conditioning of both terms as well which is not seen in the feedback...
bounds [40]; this is needed in order to, together with the genie \( M_{12}, M_{21}, M_{43}, Y^n_2 \), be able to create \( X^n_1 \) at node 4 (essentially, to create \( Y^n_1 \) to create \( X^n_1 \)).

\[
n(R_{12} + R_{34} - \epsilon) \leq I(M_{12}; Y^n_2 | M_{21}, M_{43}, Z^n_1) + I(M_{34}; Y^n_2 | M_{12}, M_{21}, M_{43}, Z^n_1) \\
= H(Y^n_2 | M_{21}, M_{43}, Z^n_1) - H(Y^n_2 | M_{21}, M_{43}, Z^n_1, M_{12}) \\
+ H(Y^n_2, Y^n_4 | M_{21}, M_{12}, M_{43}, Z^n_1) - H(Y^n_2, Y^n_4 | M_{21}, M_{12}, M_{43}, Z^n_1, M_{34}) \\
\leq H(Y^n_2 | M_{21}, M_{43}, Z^n_1) + H(Y^n_4 | M_{21}, M_{12}, M_{43}, Z^n_1, Y^n_2) - H(Z^n_2, Z^n_4) \\
\overset{(a)}{=} H(Y^n_2 | M_{21}, M_{43}, Z^n_1) - H(Z^n_2) \\
+ \sum_{i=1}^{n} [H(g_{34} X_{3,i} + Z_{4,i} | M_{21}, M_{12}, M_{43}, Y^{i-1}_4, X^i_1, Y^n_2, X^n_2, Z^n_1, X^i_1)] - H(Z^n_4) \\
\overset{(b)}{\leq} \sum_{i=1}^{n} [H(Y_{2,i} | Y^{i-1}_2, M_{21}, X_{2,i}) - H(Z_{2,i}) + H(g_{34} X_{3,i} + Z_{4,i} | X_{4,i}, g_{32} X_{3,i} + Z_{2,i}, X^i_1, X^n_2)] - H(Z_{4,i}) \\
\overset{(c)}{\leq} \sum_{i=1}^{n} H(g_{12} X_{1,i} + g_{32} X_{3,i} + Z_{2,i} | X_{2,i}) - H(Z_{2,i}) + H(g_{34} X_{3,i} + Z_{4,i} | X_{4,i}, g_{32} X_{3,i} + Z_{2,i}) - H(Z_{4,i}) \\
(3.19)
\]

In step (a), \( X^n_1 \) in the conditioning of the third term is constructed from \( (M_{12}, X^n_2, X^n_1, Z^n_1) \). In step (b), we used conditioning reduces entropy, and \( g_{32} X_{3,i} + Z_{2,i} \) in the conditioning of the third term is decoded from \( Y^n_2 \). In step (c), we only keep the self-interference \( X_{4,i} \) and drop the terms \( X^n_1, X^n_2 \) in the conditioning of the third term. We could leave these and express the outer bound in terms of correlation coefficients between the inputs (which in general may be correlated due to full adaptation). However, in subsequent steps we will seek to maximize,
or outer bound this outer bound to obtain a simple analytical expression, which amounts to setting certain correlation coefficients to 0, or equivalently, dropping the terms $X_i^1, X_i^2$ in the conditioning. Further evaluation yields (Equation 3.18), for details please refer to Appendix B.

**Remark 13.** Sum-rate bound: Note that the final, evaluated symmetric, normalized sum-rate bound in (Equation 3.18) has the same form as the IC with perfect output feedback outer bound [40, upper bound on (7)], though they are arrived at using different genies (though similar in many senses as mentioned above). In both channel models, inputs may be arbitrarily correlated as no additional arguments for restricting the input distributions have been made, leading to similar bounding techniques using correlation coefficients.

**Theorem 14.** Outer bound: partial adaptation. For the Gaussian two-way IC under partial adaptation (Equation 3.14) – (Equation 3.15), in addition to the bounds in Theorem 13, we may also conclude that any achievable rates ($R_{12}, R_{21}, R_{34}, R_{43}$), and $R_{\text{sym}+} = \frac{R_{12} + R_{34}}{2}$ and $R_{\text{sym}-} = \frac{R_{21} + R_{43}}{2}$ must satisfy,

\[
\begin{align*}
R_{12} &\leq \log(1 + \text{SNR}_{12}) \\
R_{21} &\leq \log(1 + \text{SNR}_{21}) \\
R_{34} &\leq \log(1 + \text{SNR}_{34}) \\
R_{43} &\leq \log(1 + \text{SNR}_{43})
\end{align*}
\]
\[
R_{\text{sym} \rightarrow} \leq \log \left( 1 + \text{INR} + \text{SNR} - \frac{\text{INR} \times \text{SNR}}{1 + \text{INR}} \right) \quad (3.24)
\]

\[
R_{\text{sym} \leftarrow} \leq \begin{cases} 
\log \left( 1 + \text{INR} + \frac{\text{SNR}}{\text{INR}} \right), & \text{if } \text{SNR} \leq \text{INR}^3 \\
\log \left( 1 + \frac{\sqrt{\text{SNR}} + \sqrt{\text{INR}}}{1 + \text{INR}} \right), & \text{if } \text{SNR} > \text{INR}^3
\end{cases} \quad (3.25)
\]

**Proof.** For the single-rate bounds, it is sufficient to show the first two due to symmetry. Notice that we must treat the \( \rightarrow \) and \( \leftarrow \) directions separately however due to the asymmetry of the partial adaptation.

\[
n(R_{12} - \epsilon) \leq I(M_{12}; Y^n_2|M_{21}, M_{34})
\]

\[
\leq H(Y^n_2|M_{21}, M_{34}) - H(Y^n_2|M_{21}, M_{34}, M_{12}, X^n_1, X^n_2, X^n_3)
\]

\[
\leq \sum_{i=1}^n [H(Y^n_2|Y^n_2^{i-1}, M_{21}, X^n_2, M_{34}, X^n_3, X^n_1) - H(Z^n_2, i)]
\]

\[
\leq \sum_{i=1}^n [H(g_{12}X^n_1 + Z^n_2, i) - H(Z^n_2, i)] \leq \sum_{i=1}^n \log(1 + \text{SNR}_{12})
\]

\[
n(R_{21} - \epsilon) \leq I(M_{21}; Y^n_1|M_{12}, M_{43}, M_{34}, Z^n_{4,1}^{-1})
\]

\[
\leq H(Y^n_1|M_{12}, M_{34}, M_{43}, Z^n_{4,1}^{-1}) - H(Y^n_1|M_{12}, M_{34}, M_{43}, Z^n_{4,1}^{-1}, M_{21}, X^n_1, X^n_2, X^n_3)
\]

\[
\leq \sum_{i=1}^n [H(Y^n_1|M_{12}, M_{34}, M_{43}, Z^n_{4,1}^{-1}, Y^n_1^{i-1}, X^n_1, X^n_2, X^n_3) - H(Z^n_1, i)]
\]

\[
\leq \sum_{i=1}^n [H(g_{21}X^n_2, i + Z^n_1, i) - H(Z^n_1, i)] \leq \sum_{i=1}^n \log(1 + \text{SNR}_{21})
\]

where (a) and (b) follows from the definition of partial adaptation and Lemma 12.
Next, we consider the sum-rate bounds (Equation 3.24) and (Equation 3.25), which are inspired by the techniques used by Etkin, Tse and Wang for the interference channel [21]. For the $\rightarrow$ direction of the symmetric rate,

\[
n(R_{12} + R_{34} - \epsilon) \leq I(M_{12}; Y_2^n, g_{14}X_1^n + Z_4^n, M_{21}, M_{43}) + I(M_{34}; Y_4^n, g_{32}X_3^n + Z_2^n, M_{21}, M_{43})
\]

\[
\begin{align*}
&[a]\ H(Y_2^n | g_{14}X_1^n + Z_4^n, M_{43}, M_{21}) + H(g_{14}X_1^n + Z_4^n | M_{43}, M_{21}) - H(Y_2^n, g_{14}X_1^n + Z_4^n | M_{12}, M_{21}, M_{43}) \\
&+ H(Y_4^n | g_{32}X_3^n + Z_2^n, M_{43}, M_{21}) + H(g_{32}X_3^n + Z_2^n | M_{43}, M_{21}) - H(Y_4^n, g_{32}X_3^n + Z_2^n | M_{34}, M_{21}, M_{43}) \\
&[b]\ H(Y_2^n | g_{14}X_1^n + Z_4^n, M_{43}, M_{21}) + \sum_{i=1}^{n} [H(g_{14}X_{1,i} + Z_{4,i} | g_{14}X_{i}^{i-1} + Z_{4,i}^{i-1}, M_{12}, M_{21}, M_{43}, X_{2,i}^{i-1}, X_1^{i-1})] \\
&+ H(Y_4^n | g_{32}X_3^n + Z_2^n, M_{43}, M_{21}) + \sum_{i=1}^{n} [H(g_{32}X_{3,i} + Z_{2,i} | g_{32}X_{3}^{i-1} + Z_{2,i}^{i-1}, M_{43}, M_{21}, M_{12})] \\
&- H(Y_{2,i}, g_{14}X_{1,i} + Z_{4,i} | Y_{2}^{i-1}, g_{14}X_{1}^{i-1} + Z_{4}^{i-1}, M_{12}, M_{21}, M_{43}, X_{2,i}^{i-1}, X_1^{i-1}) \\
&[c]\ H(Y_2^n | g_{14}X_1^n + Z_4^n, M_{43}, M_{21}) \\
+ \sum_{i=1}^{n} [H(g_{14}X_{1,i} + Z_{4,i} | g_{14}X_{i}^{i-1} + Z_{4,i}^{i-1}, M_{43}, M_{21}, M_{34}, X_{3,i}^{i-1}, Y_{4,i}^{i-1}, X_{i}^{i-1}, g_{32}X_{3}^{i-1} + Z_{2,i}^{i-1})] \\
&- H(g_{32}X_{3,i} + Z_{2,i}, Z_{4,i} | Y_{2}^{i-1}, g_{14}X_{1}^{i-1} + Z_{4}^{i-1}, M_{12}, M_{21}, M_{43}, X_{2,i}^{i-1}, X_1^{i-1}, g_{32}X_{3}^{i-1} + Z_{2,i}^{i-1})] \\
&+ H(Y_4^n | g_{32}X_3^n + Z_2^n, M_{43}, M_{21}) \\
+ \sum_{i=1}^{n} [H(g_{32}X_{3,i} + Z_{2,i} | g_{32}X_{3}^{i-1} + Z_{2,i}^{i-1}, M_{43}, M_{21}, M_{12}, X_{1}^{i-1}, Y_{2}^{i-1}, X_{2,i}^{i-1}, g_{14}X_{1}^{i-1} + Z_{4}^{i-1})] \\
&- H(g_{14}X_{1,i} + Z_{4,i}, Z_{2,i} | Y_{4}^{i-1}, g_{32}X_{3}^{i-1} + Z_{2}^{i-1}, M_{34}, M_{21}, M_{43}, X_{4,i}^{i-1}, X_{3,i}^{i-1}, g_{14}X_{1}^{i-1} + Z_{4}^{i-1})] \\
&[d]\ \sum_{i=1}^{n} [H(Y_{2,i} | Y_{2}^{i-1}, g_{14}X_{1}^{i-1} + Z_{4,i}^{i-1}, M_{43}, M_{21}) - H(Z_{2,i})] 
\end{align*}
\]
In the first step, we have given $(g_{14}X^n_{1} + Z^n_4)$ and $(g_{32}X^n_{3} + Z^n_2)$ as side information. Step (a) follows from the independence of the messages. In step (b), the 2nd and 5th terms follow since $g_{14}X_{1,i}$ and $g_{32}X_{3,i}$ are functions only of $M_{12}$ and $M_{34}$, and the 3rd and 6th terms follow from the definition of partial adaptation. For (c), in the conditioning of the 2nd term, we are able to add $(X_{3}^{i-1}, g_{32}X^{i-1}_{3} + Z^{i-1}_2)$ due to partial adaptation constraints, and $(Y_{4}^{i-1}, X^{i}_{4})$ are constructed from $(g_{14}X_{i-1}^{i-1} + M_{43}, X^{i}_{3})$. The 5th term follows similarly. In step (d), $-H(Z_{2,i})$ and $-H(Z_{4,i})$ are obtained from a portion of the 6th and 3rd terms in (c) respectively using the chain rule (noises are independent from other terms), and the remainder (chain rule) of the 6th and 3rd terms are cancelled by the 2nd and 5th terms respectively.

To obtain (Equation 3.24) we continue to outer bound (Equation 3.26) in terms of SNR and INR, using the fact that Gaussians maximize entropy subject to variance constraints. Specifically, one may intuitively see that, if one defines $\lambda_{jk} = E[X_j X^*_k]$, that one may express (Equation 3.26) in terms of $\lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{34}, \lambda_{23}$. One also notices from the conditional entropy expression in (Equation 3.26) that taking $\lambda_{14} = \lambda_{23} = \lambda_{12} = \lambda_{34} = 0$, and since $\lambda_{13} = 0$ (naturally, by partial adaptation) will maximize the outer bound. This may alternatively be
worked out by calculating the conditional covariance matrices directly (as we will show for the
next bound on $R_{\leftarrow}$). In this case then, for each $i$, we may bound

$$H(g_{12}X_1 + g_{32}X_3 + Z_2|g_{14}X_1 + Z_4) - H(Z_2) \leq H(g_{12}X_1 + g_{32}X_3 + Z_2|g_{14}X_1 + Z_4) - H(Z_2)$$

$\leq \log 2\pi e (\text{Var}(g_{12}X_1 + g_{32}X_3 + Z_2|g_{14}X_1 + Z_4)) - \log 2\pi e (\text{Var}(Z_2))$

$\leq \log \left( 1 + \text{SNR} + \text{INR} - \frac{\text{SNR} \times \text{INR}}{1 + \text{INR}} \right),$

which together with the symmetric expressions for the third and fourth terms in (Equation 3.26)
yield (Equation 3.24).

For the $\leftarrow$ direction, we are similarly able to obtain the sum-rate bound, despite some
asymmetry due to partial adaptation, where we leave out some steps:

$$n(R_{21} + R_{43} - \epsilon) \leq I(M_{21};Y^n_1, g_{23}X^n_2 + Z^n_3, M_{12}, M_{34}) + I(M_{43};Y^n_3, g_{41}X^n_4 + Z^n_1, M_{12}, M_{34})$$

$\leq ...$

$\leq \sum_{i=1}^{n} [H(g_{21}X_{2,i} + g_{41}X_{4,i} + Z_{1,i}|g_{23}X_{2,i} + Z_{3,i}, X_{1,i}) - H(Z_{1,i})$

$+ H(g_{43}X_{4,i} + g_{23}X_{2,i} + Z_{3,i}|g_{41}X_{4,i} + Z_{1,i}, X_{3,i}) - H(Z_{3,i})] \quad (3.27)$

We again proceed to outer bound (Equation 3.27) to obtain (Equation 3.25). It is sufficient
to evaluate the first two terms in (Equation 3.27) due to symmetry. Once again, we could outer
bound (Equation 3.27) in terms of the conditional covariance matrices and then proceed to
select values of the correlation coefficients (complex) $\lambda_{jk} := E[X_j X_k^*]$ which maximize this outer
A more intuitive method is to note that the conditional entropies in (Equation 3.27) will be maximized if $\lambda_{14} = \lambda_{32} = 0$, and $\lambda_{12} = \lambda_{34} = 0$ (similar to (Equation B.2)), which may also be obtained by dropping $X_{1,i}, X_{3,i}$ in the conditioning terms. At that point, we are only left with the coefficient $\lambda_{24} = E[X_2X_4^*]$, (which in contrast to the $\rightarrow$ bound is not automatically 0 due to the possible adaptation in the $\leftarrow$ direction) yielding the following bound for $R_{\text{sym} \leftarrow} = \frac{R_{21} + R_{43}}{2}$ by symmetry:

$$
R_{\text{sym} \leftarrow} \leq H(g_{21}X_2 + g_{41}X_4 + Z_1|g_{23}X_2 + Z_3, X_1) - H(Z_1)
$$

$$
\leq H(g_{21}X_2 + g_{41}X_4 + Z_1|g_{23}X_2 + Z_3) - H(Z_1)
$$

$$
\leq \log 2\pi e (\text{Var}(g_{21}X_2 + g_{41}X_4 + Z_1|g_{23}X_2 + Z_3)) - \log 2\pi e (\text{Var}(Z_1))
$$

$$
\leq \log \left(1 + \text{INR} + \text{SNR} + 2|\lambda_{24}| \cos \theta \sqrt{\text{SNR} \times \text{INR}}
\right.

\left. - \frac{\text{SNR} \times \text{INR} + \text{INR}^2|\lambda_{24}|^2 + 2\sqrt{\text{SNR} \times \text{INR}^3/2}|\lambda_{24}| \cos \theta}{1 + \text{INR}} \right).
$$

(3.28)

where $\theta$ is the angle of $g_{21}g_{41}^*\lambda_{24}$. To maximize (Equation 3.28), we take the partials of the expression with respect to $|\lambda_{24}|$ and $\theta$ and set these to 0. For these to equal 0 for all $\text{SNR}$ and $\text{INR}$ we must have $\theta = 0$ and $|\lambda_{24}| = \frac{\sqrt{\text{SNR} \times \text{INR}}}{\text{INR}^2}$ (discussed next). Note that we must constrain $|\lambda_{24}| \in [0, 1]$. In the interval $|\lambda_{24}| \in \left[0, \frac{\sqrt{\text{SNR} \times \text{INR}}}{\text{INR}^2} \right]$ one may verify that the function is increasing in $|\lambda_{24}|$. Thus, if $\frac{\sqrt{\text{SNR} \times \text{INR}}}{\text{INR}^2} \leq 1$, $(|\lambda_{24}| = \frac{\sqrt{\text{SNR} \times \text{INR}}}{\text{INR}^2}, \theta = 0)$ maximizes (Equation 3.28); this happens if $\text{SNR} \leq \text{INR}^3$, and yields the first bound in (Equation 3.25). Otherwise, for
\[ \text{SNR} > \text{INR} \geq \text{SNR}^{3}, \quad (\lambda_{24} = 1, \theta = 0) \] maximizes (Equation 3.28), yielding the second equation in (Equation 3.25).

\[ \square \]

**Remark 14.** The sum-rate bound for \( R_{\text{sym} \rightarrow} \) of (Equation 3.24) has the same form as Etkin, Tse and Wang’s outer bound for one-way Gaussian interference channel [21, (12)] which is useful in weak interference. The sum-rate bound for \( R_{\text{sym} \leftarrow} \) is quite different, and we note that it may be verified that (Equation 3.25) is always at least as large as (Equation 3.24), as one might expect given the partial adaptation constraints on nodes in the \( \rightarrow \) direction, but none on the nodes in the \( \leftarrow \) direction.

We next show that these outer bounds, derived for the fully adaptive or partially adaptive models, may be achieved to within a constant gap or capacity by non-adaptive schemes – i.e. the simultaneous decoding or the Han and Kobayashi scheme operating in the two directions independently. We break our analysis into three sub-sections: 1) very strong interference, 2) strong interference, and 3) weak interference. The overall finite gap results are summarized in Table I.

### 3.2.3 Very Strong Interference: \( \text{INR} \geq \text{SNR}(1 + \text{SNR}) \)

We first show that a non-adaptive scheme may achieve the capacity for the two-way Gaussian IC under a partially adaptive model in very strong interference. For the symmetric two-way Gaussian IC, define “very strong interference” as the class of channels for which \( \text{INR} \geq \text{SNR}(1 + \text{SNR}) \), as in [21, below equation (21)]. It is well known that the capacity region of the one-
way Gaussian IC in very strong interference is that of two parallel Gaussian point-to-point channels [66], which may be achieved by having each receiver first decode the interfering signal, treating its own as noise, subtracting off the decoded interference, and decoding its own message. Given that the interference is so strong, this may be done without a rate penalty. We ask whether the same is true for the two-way Gaussian IC with partial adaptation. The answer is affirmative and the capacity region is given by the following theorem:

**Theorem 15.** The capacity region for the two-way Gaussian interference channel with partial adaptation in very strong interference is the set of rate pairs \((R_{12}, R_{21}, R_{34}, R_{43})\), such that

\[
\text{(Equation 3.20)} \quad \text{and} \quad \text{(Equation 3.23)}
\]

are satisfied.

**Proof.** Each node may ignore its ability to adapt, and rather transmit using a \(CN(0,1)\) Gaussian random code. Each receiver may cancel its own self-interference, and then proceed to first decode the single interfering term before decoding its own message. This standard non-adaptive scheme may achieve the outer bound in (Equation 3.20)–(Equation 3.23) in Theorem 14.

3.2.4 **Strong Interference:** \(\text{SNR} \leq \text{INR} \leq \text{SNR}(1 + \text{SNR})\)

In this regime, we are able to show that a non-adaptive scheme may achieve capacity to within a constant gap of any fully adaptive scheme (in contrast to any partially adaptive scheme in the last subsection). A symmetric two-way Gaussian IC, as in [21], is said to be in “strong interference” when \(\text{INR} \geq \text{SNR}\).
The capacity region of one-way Gaussian interference channel in strong interference is given by [67], and for symmetric channels, the sum-capacity when the interference is strong but not very strong, i.e. $\text{SNR} \leq \text{INR} \leq \text{SNR}(1 + \text{SNR})$, may be written as

$$R_{\text{sym}} = \frac{R_{12} + R_{34}}{2} \leq \frac{1}{2} \log (1 + \text{SNR} + \text{INR}).$$

(3.29)

We note that this rate is achievable for the two-way Gaussian IC by using the simultaneous non-unique decoding scheme for the interference channel in strong interference [13,33,67]) in the $\rightarrow$ and $\leftarrow$ directions, and noting that any self-interference may be canceled. This non-adaptive scheme which achieves (Equation 3.29) in each direction also achieves to within 1 bit (per user, per direction) of our fully adaptive outer bound (Equation 3.18) in strong but not very strong interference.

**Theorem 16.** The sum-capacity for the two-way symmetric Gaussian interference channel with full adaptation in strong (but not very strong) interference is within 1 bit to (Equation 3.29) (per user, per direction)

Proof.

$$(\text{Equation 3.18}) - (\text{Equation 3.29})$$

$$= \frac{1}{2} \log (1 + \text{SNR} + \text{INR} + 2\sqrt{\text{SNR} \times \text{INR}}) + \frac{1}{2} \log \left(1 + \frac{\text{SNR}}{1 + \text{INR}}\right) - \frac{1}{2} \log (1 + \text{SNR} + \text{INR})$$

$$\leq \frac{1}{2} \log 2(1 + \text{SNR} + \text{INR}) + \frac{1}{2} \log \left(1 + \frac{\text{SNR}}{1 + \text{INR}}\right) - \frac{1}{2} \log (1 + \text{SNR} + \text{INR})$$

$$(a) \leq \frac{1}{2} \log \left(1 + \frac{\text{INR}}{\text{INR}}\right) = 1.$$
In step (a), we use the fact that $1 + \text{SNR} + \text{INR} + 2\sqrt{\text{SNR} \times \text{INR}} \leq 2(1 + \text{SNR} + \text{INR})$. Step (b) follows from the condition of strong interference $\text{INR} \geq \text{SNR}$. Since our bound (Equation 3.18) is valid for the symmetric assumptions of full adaptation, we conclude that the non-adaptive schemes’ gap to the fully adaptive outer bound for each user, for each direction is at most 1 bit.

**Remark 15.** Note that if we were to evaluate the fully adaptive outer bound of (Equation 3.18) under partial adaptation constraints instead, i.e., $X_1$ and $X_3$ are only functions of $M_{12}$ and $M_{34}$ respectively, then we would be able to set $\lambda_{13}$ in (Equation B.1) equal to 0, yielding a new outer bound $R_{\text{sym} \rightarrow} \leq \frac{1}{2} \log(1 + \text{SNR} + \text{INR}) + \frac{1}{2} \log (1 + \frac{\text{SNR}}{1 + \text{INR}})$. In this case a gap of $\frac{1}{2}$ bit instead of 1 bit may be shown for $R_{\text{sym} \rightarrow}$. However, due to the asymmetry of partial adaptation ($\lambda_{24}$ in general not equal to 0), in the opposite direction, we would still have a 1 bit gap for $R_{\text{sym} \leftarrow}$.

### 3.2.5 Weak Interference: $\text{INR} \leq \text{SNR}$

We now show that the well known Han and Kobayashi scheme employed in the $\rightarrow$ and $\leftarrow$ directions may achieve to within a constant number of bits of the fully or partially adaptive (depends on the channel regimes, or relative SNR and INR values) sum-capacity for the two-way Gaussian IC.

**Theorem 17.** A non-adaptive scheme may achieve to within a 2-bit per user per direction of the partially adaptive sum-capacity for the two-way symmetric Gaussian IC in weak interference. In some channel regimes, this non-adaptive scheme also achieves to within a constant gap of any fully adaptive scheme.
Proof. As for the one-way IC [21], we break our proof into two regimes: \( \text{INR} \geq 1 \) or \( \text{INR} < 1 \).

### 3.2.5.1 \( \text{INR} \geq 1 \)

Outer bounds have already been derived. Consider now using the specific choice of the Han and Kobayashi (HK) strategy utilized for the symmetric one-way IC as in [21, (4)] in each direction. That is, view nodes 1,2 as transmitters and 3,4 as receivers in the \( \rightarrow \) direction and employ the particular choice of the HK scheme where private messages are encoded at the level of the noise, and similarly for the \( \leftarrow \) direction consider nodes 3,4 as transmitters and 1,2 as receivers. Due to the additive nature of the channel and each node’s ability to first cancel out their self-interference, one may achieve the following rates per user, per node for each direction when \( \text{INR} \geq 1 \) for the symmetric two-way Gaussian IC:

\[
R_{\text{HK}} = \min \left\{ \frac{1}{2} \log \left( 1 + \text{INR} + \text{SNR} \right) + \frac{1}{2} \log \left( 2 + \frac{\text{SNR}}{\text{INR}} \right) - 1, \log \left( 1 + \text{INR} + \frac{\text{SNR}}{\text{INR}} \right) - 1 \right\}
\]

\((3.30)\)

\[
= \min(R_{\text{HK}1}, R_{\text{HK}2}).
\]

\((3.31)\)

**If the first term in** (Equation 3.30) **is active** we show a constant gap to the outer bound (Equation 3.18),

\[
(Equation \ 3.18) - R_{\text{HK}1}
\]

\[
\leq \frac{1}{2} \log 2(1 + \text{SNR} + \text{INR}) - \frac{1}{2} \log(1 + \text{INR} + \text{SNR}) + \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{\text{INR}} \right) - \frac{1}{2} \log \left( 2 + \frac{\text{SNR}}{\text{INR}} \right) + 1
\]

\[
\leq \frac{1}{2} \log(2) + \frac{1}{2} \log(1) + 1 = 1.5
\]
Remark 16. Since our bound (Equation 3.18) is derived assuming full adaptation, we may conclude that this gap holds for both $R_{\text{sym} \rightarrow}$ and $R_{\text{sym} \leftarrow}$ (i.e. holds for $R_{\text{sym}}$). If we were to consider partial adaptation ($\lambda_{13} = 0$), this gap could be reduced to 1 bit instead of 1.5 bits for $R_{\text{sym} \rightarrow}$, but would remain 1.5 bits for $R_{\text{sym} \leftarrow}$ as $\lambda_{24} \neq 0$ in general for partial adaptation.

If the second term in (Equation 3.30) is active, we use (Equation 3.24) to bound the gap for $R_{\text{sym} \rightarrow}$ as

\[
(Equation 3.24) - R_{HK2} = \log \left(1 + \frac{\text{INR} \times \text{SNR}}{1 + \text{INR}}\right) - \log \left(1 + \frac{\text{SNR}}{\text{INR}}\right) + 1 = 1
\]

Since our bound (Equation 3.24) has the same form as the ETW bound [21], the sum-capacity of the Gaussian two-way interference channel with partial adaptation in the forward direction is also to within 1 bit of the specific HK rate (Equation 3.30), (Equation 3.31) when $\text{INR} \geq 1$.

We use outer bound (Equation 3.25) for the backward direction, to bound the gap for $R_{\text{sym} \leftarrow}$, noting that we need to consider both cases separately. If the first term in (Equation 3.25) is relevant ($\text{SNR} \leq \text{INR}^2$):

\[
(Equation 3.25) - R_{HK2} = \log \left(1 + \frac{\text{SNR}}{\text{INR}}\right) - \log \left(1 + \frac{\text{SNR}}{\text{INR}}\right) + 1 = 1
\]
If the second term in (Equation 3.25) is relevant (SNR ≥ INR):

\[
\text{(Equation 3.25)} - R_{HK2} = \log \left( \frac{(1 + 2\text{INR} + \text{SNR} + 2\sqrt{\text{SNR} \times \text{INR}})\text{INR}}{(1 + \text{INR})(1 + \text{INR})\text{INR} + \text{SNR}} \right) + 1
\]

\[
\overset{[a]}{\leq} \log \left( \frac{(2(1 + \text{INR} + \text{SNR} + \text{INR})\text{INR}}{(1 + \text{INR})(1 + \text{INR})\text{INR} + \text{SNR}} \right) + 1
\]

\[
\leq \log \left( \frac{2(\text{INR} + \text{SNR} \times \text{INR} + 2\text{INR}^2 + \text{SNR} + \text{INR}^3)}{\text{INR} + \text{SNR} \times \text{INR} + 2\text{INR}^2 + \text{SNR} + \text{INR}^3} \right) + 1 = 2
\]

where (a) follows the fact that \(1 + \text{SNR} + \text{INR} + 2\sqrt{\text{SNR} \times \text{INR}} \leq 2(1 + \text{SNR} + \text{INR})\).

\textbf{Remark 17.} We have shown that the capacity region of the Gaussian two-way interference channel with partial adaptation (fix \(X_1\) and \(X_3\)) is within at most 2 bits per user per direction to the sum-rate achieved by two simultaneous HK schemes in opposite directions when \(\text{INR} \geq 1\). Again, we may conclude that partial adaptation cannot significantly increase the capacity for Gaussian two-way IC with weak interference.

\subsection{3.2.5.2 \(\text{INR} < 1\)}

In this case, a symmetric version of the HK scheme may be obtained from [21, (69)], for which each of the four users may achieve the following rate:

\[
R_{\text{INR} < 1} \leq \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right)
\]  \quad (3.32)

We show that this achieves to within 1 bit of the outer bound (Equation 3.18)

\[
\text{(Equation 3.18)} - R_{\text{INR} < 1} \leq \frac{1}{2} \log \left( 1 + \text{SNR} + \text{INR} + 2\sqrt{\text{SNR} \times \text{INR}} \right) - \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right)
\]
<table>
<thead>
<tr>
<th>Two-way Interference</th>
<th>Constant Gaps per user per direction, in bits (to outer bound)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Strong</td>
<td>0 (partial)</td>
</tr>
<tr>
<td>Strong</td>
<td>1 (full)</td>
</tr>
<tr>
<td>Weak</td>
<td>1 (full)</td>
</tr>
<tr>
<td>HK1 is active</td>
<td>1.5 (full)</td>
</tr>
<tr>
<td>HK2 is active</td>
<td>1 (partial)</td>
</tr>
<tr>
<td>$\nu \geq 1$</td>
<td>$\rightarrow$ direction</td>
</tr>
<tr>
<td>$\nu &lt; 1$</td>
<td>$\rightarrow$ direction</td>
</tr>
</tbody>
</table>

TABLE I

CONSTANT GAPS BETWEEN NON-ADAPTIVE SYMMETRIC HAN AND KOBAYASHI SCHEMES IN EACH DIRECTION AND PARTIALLY OR FULLY ADAPTIVE OUTER BOUNDS FOR THE SUM-RATE OF THE SYMMETRIC TWO-WAY GAUSSIAN IC [1].

$$\leq \frac{1}{2} \log \left( \frac{2(1 + \text{SNR} + \text{INR})(1 + \text{INR})}{1 + \text{SNR} + \text{INR}} \right)$$

$$\leq \frac{1}{2} \log(4) = 1$$

where (a) uses the condition INR < 1. Since (Equation 3.18) was obtained for full adaptation, we conclude that the capacity of the Gaussian two-way IC is to within 1 bit to the HK region when INR < 1 for both directions.

3.2.6 Final Comments on Adaptation versus No-adaptation, and versus Perfect Output Feedback

In the above, we have highlighted classes of two-way interference channels for which adaptation is useless. For the Gaussian channel, only highly symmetric scenarios were considered. The conclusions made for such symmetric scenarios, while insightful, do not tell the whole story.
That is, it must be noted that adaptation can provide unbounded gains over non-adaptation for certain channels. A simple example of a channel in which we see this is shown in Figure 5. Here the forward channel has no direct links, while the reverse channel has no cross-over links. In this scenario, a non-adaptive scheme would not be able to achieve any positive rate for $R_{12}$ and $R_{34}$. However, the adaptive scheme would be able to achieve positive rates for $R_{12}$ and $R_{34}$ by “routing” the messages. For example, message $M_{12}$ may exploit adaptation to take the path: Tx 1 → cross-over link to Rx 4 → direct reverse link from Tx 4 to Rx 3 → cross-over link Tx 3 to Rx 2. An adaptive network may thus provide unbounded gain over a non-adaptive network for at least some of the rates. Note however that if the reverse direction is “routing” messages for the forward direction, its own message rates will decrease.

One can also find examples of networks where perfect output feedback provides unbounded gain over adaptation. Note that in most scenarios where adaptation is useless, perfect output feedback is known to be useless as well. For example, for the symmetric linear deterministic
two-way IC shown in Figure 4, the symmetric sum-rate of the linear deterministic one-way IC and one-way IC with feedback are identical for $\frac{2}{3} \leq \alpha \leq 2$, indicating that feedback is useless. In this regime, adaptation was also shown to be useless. In the two-way MAC/BC, for all deterministic models, perfect output feedback may be shown to be useless, and adaptation was also useless. One might ask whether feedback and adaptation being useless always go hand in hand.

The following example shows the intuitive fact that adaptation being useless does not imply that feedback is useless. To show that feedback may provide an unbounded gain over pure adaptation, consider the two-way IC in Figure 6. For message $M_{12}$ to travel from Tx 1 to Rx 2, using feedback it may do so by taking the path: Tx 1 $\rightarrow$ cross-over link to Rx 4 $\rightarrow$ feedbacks to Tx 3 $\rightarrow$ cross-over link to Rx 2. However, clearly if we employ only adaptation over these forward and reverse links, $M_{12}$ is only able to be decoded by Rx 4 and even with adaptation has no possible way to reach Rx 2. So, feedback may improve the capacity in an unbounded way over adaptation, at least when feedback is “free”, or perfect (not over other interfering links in the reverse direction). An alternative example of when feedback may outperform adaptation is in the symmetric linear deterministic IC: for $\alpha > 2$, as seen in Figure 4, feedback outperforms adaptation. This is intuitive, as feedback is provided over perfect, infinite capacity links, whereas adaptation must take place over the same links over which the data travels.

Whether feedback being useless necessarily implies that adaptation is also useless is an interesting open question; all known examples for additive, memoryless channels seem to suggest this but it has not been rigorously shown. In a similar vein, whether each direction of a two-way
channel is outer bounded by its one-way counterpart with perfect output feedback is another open question, related to Remarks 1 and 12.

3.3 K-pair-user Gaussian Two-way Interference Channel

In this section, we propose and study a natural extension of the (2-pair-user) two-way interference channel: the K-pair-user two-way interference channel. We derive new outer bounds for both linear deterministic and Gaussian symmetric K-pair-user two-way IC, and show either the sum-capacity or constant gap results in certain interference regimes.

3.3.1 System Model

We describe the K-pair-user two-way interference channel, and in particular, the Gaussian and linear deterministic channel models.

We consider a K-pair-user two-way interference channel as shown in Figure 7, where there are $2K$ messages and $2K$ terminals forming a K-user IC in the $\rightarrow$ direction ($K$ messages) and
another K-user IC in the ← direction (K messages). All nodes are able to operate in full-duplex mode, i.e. they can transmit and receive signals simultaneously.

The channel inputs and outputs of user \( j \in \{1, 2, ..., 2K\} \) at discrete time \( i \) are \( X_{j,i} \) and \( Y_{j,i} \) that lie in alphabets \( \mathcal{X}_j \) and \( \mathcal{Y}_j \) respectively. The messages \( M_{jk} \) of rate \( R_{jk} \) from transmitter \( j \) to receiver \( k \) are uniformly distributed in \( \{1, 2, \cdots 2^{nR_{jk}}\} \) for \( j, k \in \{1, 2, ..., 2K\} \) and blocklength \( n \).

Let \( A_j^i = (A_{j,1}, A_{j,2}, ..., A_{j,i}) \), for any given time \( i \). A node is said to employ interaction if the channel input at time \( i \) is a function of the previously received outputs, \( X_{j,i} = f_j(M_{jk}, Y_{j,i-1}) \), where \( f_j \) (\( j \in \{1, 2, ..., 2K\} \)) are deterministic functions. Receiver \( k \) uses a decoding function \( d_k : \mathcal{Y}_k^n \rightarrow \tilde{M}_{jk} \) to obtain an estimate \( \tilde{M}_{jk} \) of the transmitted message \( M_{jk} \). The capacity
region is the closure of all rate tuples which simultaneously drive the probability that any of
the estimated messages is not equal to the true message, to zero as $n \to \infty$.

**Gaussian model.** The K-pair-user Gaussian two-way interference channel at each channel
use, is described by (with subscripts “o” for odd and “e” for even)

\[
Y_o = \sum_{m=1}^{K} g_{2m,o} X_{2m} + Z_o, \quad o = 1, 3, ..., 2K - 1
\]

\[
Y_e = \sum_{m=1}^{K} g_{2m-1,e} X_{2m-1} + Z_e, \quad e = 2, 4, ..., 2K.
\] (3.33) (3.34)

where $g_{jk}, j, k \in \{1, 2, ..., 2K\}$ is the channel coefficient from node j to node k, and the network is
subject to complex Gaussian noise $Z_l \sim \mathcal{CN}(0, 1), l \in \{1, 2, ..., 2K\}$. Let $P$ be the transmit power
constraint at each user: $\mathbb{E}[|X_l|^2] \leq P, l \in \{1, 2, ..., 2K\}$, and let $P = 1$ without loss of generality.
Then define $\text{SNR}_{l,l+1} = |g_{l,l+1}|^2, \text{SNR}_{l+1,l} = |g_{l+1,l}|^2, l \in \{1, 2, ..., 2K\},$ and $\text{INR}_{jk} = |g_{jk}|^2,$ for $j, k$
in the appropriate sets that denote cross links between users. Note that we have removed the
“self-interference” terms such as $g_{11}X_1$ in the expression of $Y_1$ (for example) since they can be
easily subtracted off due to the additive nature of the channel.

**Symmetric capacity.** We are interested in the symmetric capacity when all the SNRs equal
a given $\text{SNR}$, and all the INRs equal a given $\text{INR}$. We consider the per-user rates $R_{\text{sym}} = \frac{R_{12} + R_{34}}{2} = \frac{R_{21} + R_{43}}{2}$.

**Linear deterministic model.** For the linear deterministic model which models the Gaussian
channel at high SNR, the channel inputs and outputs are binary vectors, and all addition is
bit-wise and modulo 2. We define $n_{jk} = \lfloor \log g_{jk}^2 P \rfloor$ to indicate the number of signal bit levels
from transmitter \( j \) to receiver \( k \). Let \( S \) denote an \( N \times N \) lower shift matrix, where \( N = \max(n_{jk}) \).

Now the channel inputs/outputs relationship can be described as

\[
Y_o = \sum_{m=1}^{K} S^{N-n_{2m,o}} X_{2m}, \quad o = 1, 3, ..., 2K - 1 \tag{3.35}
\]

\[
Y_e = \sum_{m=1}^{K} S^{N-n_{2m-1,e}} X_{2m-1}, \quad e = 2, 4, ..., 2K. \tag{3.36}
\]

### 3.3.2 Outer bounds and capacity/gap results

For the symmetric \( K \)-pair-user two-way interference channel with interaction in two “medium” interference regimes (to be specified in the theorem statements), we derive new outer bounds and demonstrate a capacity result for the linear deterministic and a constant gap result for the Gaussian model.

#### 3.3.2.1 Capacity result for the linear deterministic model

We consider the symmetric case in which all direct links have the same number of signal bit levels \( p \), and all cross links have the same number of signal bit levels \( q \), and define \( \alpha = q/p \).

**Theorem 18.** The symmetric sum-capacity of linear deterministic \( K \)-pair-user two-way interference channel with interaction when \( 2/3 < \alpha < 2, \alpha \neq 1 \) is the rate which satisfies:

\[
R_{\text{sym}} \leq \frac{1}{2} (\max(p, q) + [p - q]^+) \tag{3.37}
\]
Proof. Achievability follows from the known non-adaptive scheme as in [22] (used in each direction separately). For the converse, valid for interactive, two-way channel models, let $M_\Lambda$ denote all the messages except $M_{12}, M_{34}$. Then,

$$n(R_{12} + R_{34})$$

$$\leq H(M_{12}) + H(M_{34})$$

$$= H(M_{12}, M_{34}|M_\Lambda)$$

$$\leq H(M_{12}, M_{34}, Y_2^n, Y_4^n|M_\Lambda)$$

$$= H(Y_2^n|M_\Lambda) + H(M_{34}|M_\Lambda, Y_2^n)$$

$$+ H(Y_2^n|Y_3^n, M_\Lambda, Y_4^n) + H(M_{12}|M_{34}, Y_2^n, Y_4^n, M_\Lambda)$$

$$\leq H(Y_2^n) + H(M_{34}|Y_2^n) + H(Y_2^n|M_{34}, M_\Lambda, Y_4^n) + H(M_{12}|Y_2^n)$$

$$\leq n(\max(p, q) + 2\epsilon) + H(Y_2^n|M_{34}, M_\Lambda, Y_4^n).$$

We proceed to bound the remaining entropy term:

$$H(Y_2^n|M_{34}, M_\Lambda, Y_4^n)$$

$$\leq H(Y_2^n, Y_3^n, Y_5^n, ..., Y_{2K-1}^n|M_{34}, M_\Lambda, Y_4^n)$$

$$= \sum_{i=1}^{n} [H(Y_{2,i}, Y_{3,i}, Y_{5,i}, ..., Y_{2K-1,i}|Y_{2,i}^{i-1}, Y_3^{i-1}, Y_5^{i-1}, ..., Y_{2K-1}^{i-1}), Y_{2K-1,i}, Y_1^n, M_{34}, M_\Lambda, X_4^n, X_2^i, X_3^i, X_5^i, ..., X_{2K-1}^i)]$$

$$= \sum_{i=1}^{n} [H(Y_{2,i}|Y_{2,i}^{i-1}, Y_3^{i-1}, Y_5^{i-1}, ..., Y_{2K-1}^{i-1}, Y_4^n, M_{34}, M_\Lambda, X_4^n, X_2^i, X_3^i, X_5^i, ..., X_{2K-1}^i)]$$. 

where in (a) we use a Markov chain that given \((M_4^n, Y_4^n, X_3^i, X_5^i, \ldots, X_{2K-1}^i, X_6^i, X_8^i, \ldots, X_{2K}^i)\) and the symmetric nature (channel gains in the cross links between transmitter 1 and the non-desired receivers are the same) of the model, we can construct \(X_b^i\). Similarly, \(X_8^i, \ldots, X_{2K}^i\) are constructed. Combining everything and considering the symmetric rate, completes the proof. \(\square\)

**Remark 18.** We note that in the regime \(2/3 < \alpha < 2\) (which contains the “moderately weak and strong interference regimes in the terminology of [22]), this capacity result is the same as that of two \(K\)-user ICs operating simultaneously in both directions, i.e. the same as two \(K\)-user ICs as in [22]. The technique is standard and similar to the proof in [24], but with more messages and care to be taken because of the interaction.

**Remark 19.** For the case of \(\alpha = 1\), the channel gains in the direct links and cross links are the same. This point has been shown to be discontinuous in the generalized degrees of freedom for the one-way \(K\)-user IC [22] (the value is \(1/K\)) because at this point all receivers receive exactly same signals for the linear deterministic model. However, if one considers time-varying channels which grow at the same rate but are not necessarily identical, the value of this point, which is known as the degrees of freedom, is shown to be \(K/2\) [49], which is achieved by interference alignment. The same result has been shown for almost all (excluding a set of
measure zero) constant channels in [60]. For the K-pair-user two-way IC, we show the degrees of freedom is K in section 4.2 for both time-varying or constant channels.

3.3.2.2 Constant gap result for Gaussian model

We next derive an outer bound for the Gaussian interactive two-way K-user IC and show that, for certain “medium interference” regimes to be specified, this lies to within a constant gap of the outer bounds for two one-way, non-interactive K-user Gaussian ICs (for symmetric channels) of [23, Eq. (42)] which are identical to those of the 2-user IC [21], which in turn have been shown to lie within a constant gap of again, non-adaptive inner bounds for all channel gains outside a small outage set (we leave details of this “small outage set” to [23]). This means that, in the two regimes considered, “adaptation” or “interaction” may only provide a bounded gain.

**Theorem 19.** The symmetric sum-capacity of K-pair-user Gaussian two-way interference channel with interaction in the moderately weak interference ($\frac{2}{3} \log \text{SNR} \leq \log \text{INR} \leq \log \text{SNR}$ or $\frac{2}{3} \leq \alpha \leq 1$) and the strong interference ($\log \text{SNR} \leq \log \text{INR} \leq 2 \log \text{SNR}$ or $1 \leq \alpha \leq 2$) regimes is within $\log(K) + \frac{K}{2} - 1$ bits to the outer bound of two simultaneously operating one-way (non-interactive) Gaussian K-user interference channels, which in turn may be shown to be within a constant gap to non-adaptive inner bounds for all channel gains outside a small outage set, as done in [23].
Proof. We first derive a new outer bound for our channel model and then show the gap result.

Let \( Z_3, \ldots, Z_{2K-1} \) be the vector of noises \( Z_3, Z_5, \ldots, Z_{2K-1} \). Define \( \tilde{Z}_1 = Z_1 - Z_4, l = 6, 8, \ldots, 2K \). Let \( \tilde{Z}_6, \ldots, 2K \) denote \( \tilde{Z}_6, \tilde{Z}_8, \ldots, \tilde{Z}_{2K} \).

\[
\begin{align*}
n(R_{12} + R_{34} - \epsilon) \\
\leq I(M_{34}; Y^n_{12} | M_A, Z^n_3, \ldots, Z^n_{2K-1}) + I(M_{12}; Y^n_{14} | M_{34}, M_A, Z^n_3, \ldots, Z^n_{2K-1}) \\
\leq I(M_{34}; Y^n_{14}, Z^n_6, \ldots, Z^n_{2K} | M_A, Z^n_3, \ldots, Z^n_{2K-1}) \\
+ I(M_{12}; Y^n_{24}, Y^n_{14}, Z^n_6, \ldots, Z^n_{2K} | M_{34}, M_A, Z^n_3, \ldots, Z^n_{2K-1}) \\
= H(Y^n_{14}, Z^n_6, \ldots, Z^n_{2K} | M_A, Z^n_3, \ldots, Z^n_{2K-1}) \\
+ H(Y^n_{24} | Y^n_{14}, Z^n_6, \ldots, Z^n_{2K}, M_{34}, M_A, Z^n_3, \ldots, Z^n_{2K-1}) \\
- H(Y^n_{14}, Y^n_{24}, Z^n_6, \ldots, Z^n_{2K} | M_{12}, M_{34}, M_A, Z^n_3, \ldots, Z^n_{2K-1})
\end{align*}
\]

where (a) follows as all messages and noises are independent; (b) by adding the side information \( \tilde{Z}_6, \ldots, 2K \). We bound the three terms above respectively. For the first term:

\[
\begin{align*}
H(Y^n_{14}, Z^n_6, \ldots, Z^n_{2K} | M_A, Z^n_3, \ldots, Z^n_{2K-1}) &\leq H(Y^n_{14}) + H(Z^n_6, \ldots, Z^n_{2K}) \\
&\leq H(g_{14}X^n_1 + g_{34}X^n_3 + \ldots + g_{2K-1,4}X^n_{2K-1} + Z^n_4) + H(Z^n_6, \ldots, Z^n_{2K}) \\
\stackrel{(a)}{\leq} n \log 2\pi e (1 + SNR + (K - 1)INR + 2(K - 1)\sqrt{SNR \times INR}) \\
&\quad + (K - 1)(K - 2)INR + n(K - 2) \log(2\pi e)(2)
\end{align*}
\]

where in (a) we have used the fact that Gaussians maximize entropy subject to power constraints and the symmetric channel model. Due to interaction, the inputs \( X_l, l \in \{1, 3, \ldots, 2K-1\} \)
may be correlated, and so we have upper bounded this term by assuming all the transmitters have the same power and they are maximally (fully) correlated.

The second term can be bounded as follows:

$$H(Y^n_2 | Y^n_4, Z^n_6, ..., Z^n_{2K-1}, M_{34}, M_A, Z^n_3, ..., Z^n_{2K-1})$$

$$\leq H(Y^n_2, Y^n_3, Y^n_5, ..., Y^n_{2K-1} | Y^n_4, Z^n_6, ..., Z^n_{2K-1}, M_{34}, M_A, Z^n_3, ..., Z^n_{2K-1})$$

$$= \sum_{i=1}^{n} [H(Y_{2,i}, Y_{3,i}, Y_{5,i}, ..., Y_{2K-1,i} | Y^n_2, Y^n_3, ..., Y^n_{2K-1})]$$

$$= n \sum_{i=1}^{n} [H(g_{12}X_{1,i} + Z_{2,i} | g_{14}X_{1,i} + Z_{4,i})]$$

$$= n \log 2\pi e \left(1 + \frac{\text{SNR} + \text{INR}}{1 + \text{INR}}\right)$$

where in step (b) we construct $X_{6,i}$ in the conditioning because (1), $g_{14}X_{1,i} + Z_{2,i}$ can be decoded from $Y^n_4$ since $X_1^n, X_3^n, ..., X_{2K-1}^n$ are known; (2), $Z_{6,i}^{i-1}$ is known so that $g_{14}X_{1,i} + Z_{6,i}^{i-1}$ can be constructed; and (3), we consider symmetric model, i.e. $g_{14} = g_{16}$. Therefore $g_{16}X_{1,i}^{i-1} + Z_{6,i}^{i-1}$ is known and then we can construct $X_{6,i}$. Similarly $X_{8,i}, ..., X_{2K,i}$ can be constructed.

Finally, the negative third term can be lower bounded as:

$$H(Y^n_2, Y^n_4, Z^n_6, ..., Z^n_{2K} | M_{12}, M_{34}, M_A, Z^n_3, ..., Z^n_{2K-1})$$
Combining everything and considering the symmetric rate yields the following outer bound for the K-pair-user two-way Gaussian interference channel:

\[
R_{\text{sym}} = \frac{R_{12} + R_{34}}{2} \leq \frac{1}{2} \log(1 + \text{SNR} + (K - 1)^2 \text{INR}) + 2(K - 1)\sqrt{\text{SNR} \times \text{INR}} + \frac{1}{2} \log \left(1 + \frac{\text{SNR}}{1 + \text{INR}}\right) + \frac{K - 2}{2}
\]

We show that our outer bound is to within a constant gap to existing non-adaptive outer bounds for K-user one-way Gaussian interference channel provided in [23, Eq. (42)].

1) The following non-adaptive bound is for moderately weak interference regime given by \(\frac{1}{2} \log \text{SNR} \leq \log \text{INR} \leq \log \text{SNR}\) or \(\frac{1}{2} \leq \alpha \leq 1\):

\[
R_{\text{sym}1} \leq \frac{1}{2} \log(1 + \text{SNR}) + \frac{1}{2} \log(1 + \frac{\text{SNR}}{1 + \text{INR}}) \quad (3.38)
\]

In this case, the gap may be bounded as

\[
R_{\text{sym}} - R_{\text{sym}1} = \frac{1}{2} \log \frac{1 + (K - 1)^2 \text{INR} + \text{SNR} + 2(K - 1)\sqrt{\text{SNR} \times \text{INR}}}{1 + \text{SNR}} + \frac{K - 2}{2} \frac{\text{INR}}{\text{SNR}} \leq \frac{1}{2} \log \left(1 + \frac{(K^2 - 1)\text{SNR}}{1 + \text{SNR}}\right) + \frac{K - 2}{2}
\]
\[ \leq \frac{1}{2} \log(K^2) + \frac{K - 2}{2} = \log K + \frac{K}{2} - 1. \]

2) The following non-adaptive bound is for strong but not very strong interference regime given by \( \log \text{SNR} \leq \log \text{INR} \leq 2 \log \text{SNR} \) or \( 1 \leq \alpha \leq 2 \):

\[ R_{\text{sym}2} \leq \frac{1}{2} \log(1 + \text{SNR} + \text{INR}) \]  

(3.39)

In this case, the gap may be bounded as

\[
R_{\text{sym}} - R_{\text{sym}2} = \\
\frac{1}{2} \log \left( 1 + (K-1)^2 \frac{\text{SNR} + 2(K-1) \sqrt{\text{SNR} \times \text{INR}}}{1 + \text{INR}} + \frac{K - 2}{2} \right) \\
\leq \frac{1}{2} \log \left( 1 + \frac{(K^2 - 1) \text{INR}}{1 + \text{INR}} \right) + \frac{K - 2}{2} \leq \log K + \frac{K}{2} - 1. \]

3.4 Summary

In this chapter, we investigated a few examples of wireless Gaussian two-way networks. Specifically, we first showed that adaptation can only increase the sum-rate of the two-way Gaussian MAC/BC by up to \( \frac{1}{2} \) bit per direction. We then considered the Gaussian two-way IC with 4 terminals and 4 messages. There, it was shown that partial adaptation is useless in very strong interference, and for all other regimes non-adaptive schemes achieved to within constant gaps (at most 2 bits) of fully, or partially, adaptive schemes. Finally, we introduced the
K-pair-user two-way IC with interaction, for which we derived new outer bounds and demonstrated a capacity result for the linear deterministic model and a constant gap result for the Gaussian model, both in two “medium” interference regimes. Again, these results indicate that adaptation, or interaction between users is useless or may only provide limited capacity gains for the considered two-way networks.
CHAPTER 4

DEGREES OF FREEDOM OF WIRELESS TWO-WAY FULL-DUPEX NETWORKS

We have explored the Gaussian two-way networks in the previous chapter, and their capacities are shown to within a constant gap in certain parameter regimes and adaptation constraints. However, determining the capacity region in general is known to be an extremely difficult problem. In this chapter, we study the degrees of freedom, which is an alternative approximate capacity characterization, of several wireless two-way full-duplex networks. The study of full-duplex communication networks is also motivated by the fact that full-duplex (in-band) operation is becoming practically realizable. Specifically, we consider the $K$-pair-user full-duplex two-way interference channel with and without a MIMO relay.

4.1 System Model

We describe the $K$-pair-user full-duplex TWIC without and with a relay in this section.

4.1.1 $K$-pair-user Full-duplex Two-Way IC

The model of the $K$-pair-user TWIC has been introduced in section 3.3.1, as shown in Figure 7. Note with some abuse of notations (from previous chapters) we let $k$ denote the index of time slots in this chapter.
At each time slot $k$, the system input/output relationships are described as:

$$Y_p[k] = \sum_{m=1}^{K} h_{2m,p}[k]X_{2m}[k] + Z_p[k], \quad p = 1, 3, ..., 2K - 1$$ (4.1)

$$Y_q[k] = \sum_{m=1}^{K} h_{2m-1,q}[k]X_{2m-1}[k] + Z_q[k], \quad q = 2, 4, ..., 2K$$ (4.2)

where $X_l[k], Y_l[k], l \in \{1, 2, ..., 2K\}$ are the inputs and outputs of user $l$ at time slot $k$, and $h_{ij}[k], i, j \in \{1, 2, ..., 2K\}$ is the channel coefficient from node $i$ to node $j$ at time slot $k$.\(^1\) The network is subject to complex Gaussian noise $Z_l[k] \sim CN(0,1), l \in \{1, 2, ..., 2K\}$ which are independent across users and time slots. We consider time-varying channel coefficients, which for each channel use are all drawn from a continuous distribution (which need not be the same for all channel gains and time instances, as long as they are continuous) and whose absolute values are bounded between a nonzero minimum value and a finite maximum value. Note one can also alternatively consider a frequency selective rather than time-varying system model.

\(^1\)Note that if user $l$ is transmitting, we have already assumed that its own “self-interference” signal has been ideally subtracted off its received signal, and is hence not present in the above description of channel inputs and outputs. This idealization will of course form an upper bound on what is possible if full self-interference cancellation is not possible, which is outside the scope of this work.
We further assume per user, per symbol power constraints $E[|X_i[k]|^2] \leq P, i \in \{1, 2, \ldots, 2K\}, k \in \{1, 2, \ldots, n\}$, for block length $n$.\(^1\) User $2i - 1$ and $2i$ wish to exchange messages for $i = 1, 2, \ldots, K$ (user 1 sends to 2, 2 to 1, ..., 2K-1 to 2K, 2K to 2K-1) with interactive encoding functions

$$X_i[k] = f(M_{ij}, Y_{i}^{k-1}), \quad k = 1, 2, \ldots, n$$

at rate $R_{i,j} = \frac{\log_2 |M_{ij}|}{n}$, where $Y_{i}^{k-1}$ denotes the vector $(Y_i[1], \ldots, Y_i[k - 1])$ from time slot, or channel use 1 to $k - 1$ received at user $i$, and $n$ denotes the total number of channel uses (the blocklength). In other words, all users in this network can adapt current channel inputs to previously received channel outputs. The nodes $2i - 1$ and $2i$ have decoding functions which map $(Y_{2i-1}^n, M_{2i-1, 2i})$ to an estimate of $M_{2i, 2i-1}$ and $(Y_{2i}^n, M_{2i, 2i-1})$ to $M_{2i-1, 2i}$, respectively. A rate tuple $(R_{i+1,i}(P), R_{i+1,i}(P))_{i \in \{1, 3, \ldots, 2K-1\}}$, where we use the argument $P$ simply to remind the reader that this rate is indeed a function of the power constraint $P$, is said to be achievable if there exist a set of interactive encoders and decoders such that the desired messages may be estimated with arbitrarily small probability of error when the number of channel uses $n$ tends to infinity. The sum DoF characterizes the sum capacity of this Gaussian channel at high SNR and is defined as the maximum over all achievable $(R_{i+1,i}(P), R_{i+1,i}(P))_{i \in \{1, 3, \ldots, 2K-1\}}$ of

$$d_{\text{sum}} = \sum_{i=1,3,\ldots,2K-1} (d_{i+1,i} + d_{i+1,i})$$

\(^1\)In our outer bound in Theorem 20, several of the terms may be extended to per user average power constraints, but we leave the per symbol power constraints for simplicity in this initial study, as is often done in degree of freedom results.
\[
= \limsup_{P \to \infty} \sum_{i=1,3,\ldots,2K-1} \frac{R_{i,i+1}(P) + R_{i+1,i}(P)}{\log(P)}.
\]

Notice the implicit definitions of the DoF of the link from user \(i\) to user \(i+1\), \(d_{i,i+1}\) and the reverse \(d_{i+1,i}\).

### 4.1.2 K-pair-user Full-duplex Two-way IC with a MIMO Relay

We consider a K-pair-user two-way interference channel with a MIMO relay as shown in Figure 8. All the system settings are the same as in the previous subsection except there is a MIMO relay which helps in communicating messages and managing interference in the network. As before, all nodes including the relay are able to operate in full-duplex mode, or transmit and receive at the same time over the same channel, and perfectly cancel out their self-interference.

The relay is assumed to have \(M\) antennas and to operate either in a non-causal or “instantaneous” fashion, or in a causal fashion. By “instantaneous” (non causal, relay-without-delay [68]) we refer to its ability to decode and forward signals received at the previous and current (but not future) time slots. We note that this requirement is significantly less strict than a cognitive relay, which would know all users’ signals prior to transmission and does not obtain the messages over the air. We will comment more on the usage / impact of a cognitive relay in subsection 4.3.2. Here messages are obtained over the air; the only idealization is the non causality or access to received signals from the current time slot. An alternative motivation for this type
Figure 8. K-pair-user two-way interference channel with a MIMO relay. $M_{ij}$ denotes the message known at node $i$ and desired at node $j$; $\tilde{M}_{ij}$ denotes that $j$ would like to decode the message $M_{ij}$ from node $i$ [3].
of instantaneous relay may be found in [69]. Mathematically, we may describe non causal and
causal relaying functions, for each \( k = 1, 2, \ldots, n \), as

\[
\text{Non-causal / instantaneous relaying: } \quad X_R[k] = g_k(Y_R[1], Y_R[2], \ldots, Y_R[k])
\]

\[
\text{Causal relaying: } \quad X_R[k] = g_k(Y_R[1], Y_R[2], \ldots, Y_R[k-1]),
\]

where \( X_R[k] \) is a \( M \times 1 \) (M antennas) vector signal transmitted by the relay at time slot \( k \); \( g_k() \) is a deterministic function; and \( Y_R[l], l \in \{1, 2, \ldots, k\} \) is the \( M \times 1 \) vector of signals received at the relay at time slot \( l \). The relay is subject to per symbol transmit power constraints over all antennas \( E[||X_R[k]||^2] \leq P_R, \forall k \in \{1, 2, \ldots, n\} \), and global channel state information knowledge is assumed at all nodes. At each time slot \( k \), the system input/output relationships are:

\[
Y_p[k] = \sum_{m=1}^{k} h_{2m,p}[k]X_{2m}[k] + h_{Rp}[k]X_R[k] + Z_p[k], \quad p = 1, 3, \ldots, 2K - 1 \tag{4.3}
\]

\[
Y_q[k] = \sum_{m=1}^{k} h_{2m-1,q}[k]X_{2m-1}[k] + h_{Rq}[k]X_R[k] + Z_q[k], \quad q = 2, 4, \ldots, 2K \tag{4.4}
\]

\[
Y_R[k] = \sum_{m=1}^{2K} h_{m,R}[k]X_m[k] + Z_R[k] \quad \tag{4.5}
\]
where we use the same notation as in (Equation 4.1) and (Equation 4.2). In addition, $h_{ij}[k], i,j \in \{1, 2, \ldots, 2K, R\}$ is the $M \times 1$-dimensional channel coefficient vector from node $i$ to node $j$ at time slot $k$ ($i$ or $j$ must be the relay node $R$), and $Z_R[k] \sim CN(0, I)$ is the complex Gaussian noise vector at the relay. The terms in bold represent vectors (due to the MIMO relay). We use $^*$ to denote conjugate transpose and $^T$ to denote transpose.

4.1.3 Types of Signals

Let $s_{ij}$ denote the independent information symbols (signals) from transmitter $i$ to receiver $j$; these are real numbers which will be combined into the signals $X_i[k]$ transmitted by node $i$ at channel use $k$. The received signal at any given node may be broken down into four types of signals:

- the self-interference signal (SI, sent by itself, known to itself);
- the interference signal (sent by the undesired user(s) from the opposite side);
- the desired signal (sent by the desired user);
- the undesired signal (sent by the undesired user(s) from the same side), respectively.

For example, at receiver 1, $s_{12}$ is a self-interference signal (SI); $s_{43}, s_{56}, \cdots, s_{2K,2K-1}$ are the interference signals; $s_{21}$ is the desired signal, and $s_{34}, s_{56}, \cdots, s_{2K-1,2K}$ are the undesired signals. Note that we differentiate between interference and undesired signals (both of which in fact do not carry any messages desired by a particular node) as they will be treated in different ways in the following: interference signals may originate from other users (via direct links) or the relay and neutralized (combined with the direct links over which they are received to cancel
the interference) by choice of the relay beam forming vector, while undesired signals would be received from the relay node only, but will be nulled by proper choice of beam forming vectors at the relay.

Note we have already removed self-interference signals from the input/output equations (Equation 4.1)-(Equation 4.4), but SI terms may still be transmitted by the relay (or other users due to adaptation) and hence received.

4.2 **DoF of K-pair-user Full-duplex Two-way IC**

In this section we show that the degrees of freedom of the K-pair-user two-way IC is K, i.e. K/2 in each direction (the DoF of a one-way K-user IC is K/2 [49]), for both time-varying and (almost all) constant channel coefficients. This result indicates that while the full-duplex operation essentially doubles the DoF, interaction between users cannot further increase the DoF beyond what full-duplex allows. This may be intuitively explained as follows: 1) the DoF measures the number of clean information streams that may be transmitted at high SNR when the desired signals and interference signals are received roughly “at the same level” (SNR and INR scale to infinity at the same rate). In this case, rates cannot be improved by having users send messages of other users to re-route the message (i.e. message from user 1 to 2 could go via 1 to 4, 4 to 3 then 3 to 2 instead) as all links are equally strong. One would thus need to tradeoff one’s own rate to relay another user’s rate given the symmetry in the channels. 2) Adaptation allows for the correlation of messages at transmitters. In Gaussian channels, such correlation may be translated into coherent power gains inside the logarithm. The DoF metric is insensitive to coherent power gains as it measures pre-logarithm gains, not constant power factor gains
inside the logarithm, and hence even correlation between inputs which adaptation/interaction would permit does not improve the DoF.

The main result of this section is stated in the following theorem:

**Theorem 20.** The full-duplex $K$-pair-user two-way interference channel has $K$ degrees of freedom.

*Proof.* We use the achievability scheme used in demonstrating the DoF of the one-way $K$-user IC ( [49] for time-varying channel and [60] for (almost all) constant channels) in each $\rightarrow$ and $\leftarrow$ direction simultaneously with non-interactive nodes (i.e. each direction uses this scheme and ignores the past received signals, a non-interactive scheme). By making the appropriate correspondences, we see that $K/2$ DoF are achievable in each direction, leading to a sum DoF of $K$. This assumes that self-interference is able to be perfectly cancelled.

Now we prove the converse, which is valid for both time-varying and constant channel gains. This outer bound carefully merges techniques used in the one-sided or Z interference channel [21,70–72] (asymmetric genies to the two receivers in one direction), in the symmetric $K$-user interference channel with feedback [24] (providing differences between noises as genies), and in determining the DoF of the $K$-user interference channel [49] (re-scaling of channel coefficients). This combination is new and relies on novel constructions, and is more involved than the individual pieces given the larger number of messages and noises involved and the fact that we allow for adaptation. This bound can also be extended to the symmetric Gaussian channel model and used to show a constant gap to capacity result as in [2] (i.e. adaptation or interaction,
in some regimes of the symmetric Gaussian channel, can only improve capacity to within an additive constant gap).

First rewrite channel outputs in (Equation 4.2) for \( q = 6, 8, \ldots, 2K \) as

\[
Y_q[k] = h_{14}[k]X_1[k] + \sum_{m=2}^{K} \frac{h_{14}[k]}{h_{1,q}[k]} h_{2m-1,q}[k]X_{2m-1}[k] + Z'_q[k]
\]  

(4.6)

where \( Z'_q[k] \sim \mathcal{CN}(0, h_{1,q}[k] h_{21,q}[k]) \). Let \( M_A \) denote all the messages except \( M_{12}, M_{34} \), and let \( Z_{3,\ldots,2K-1}[k] \) denote noises \( Z_3[k], Z_5[k], \ldots, Z_{2K-1}[k] \). Define \( \bar{Z}_q[k] = Z'_q[k] - Z_q[k], q = 6, 8, \ldots, 2K, \) which is \( \mathcal{CN}(0, 1 + h_{1,q}[k] h_{21,q}[k]) \). Let \( \bar{Z}_{6,\ldots,2K}[k] \) denote \( \bar{Z}_6[k], \bar{Z}_8[k], \ldots, \bar{Z}_{2K}[k] \). We start by bounding the sum of a pair of rates:

\[
\begin{align*}
\mathbb{P}(R_{12} + R_{34} - \epsilon) & \leq I(M_{34}; Y_q^n|M_A, Z^n_{3,\ldots,2K-1}) \\
& \quad + I(M_{12}; Y_2^n, Y_4^n|M_{34}, M_A, Z^n_{3,\ldots,2K-1}) \\
& \leq I(M_{34}; Y_q^n, \bar{Z}^n_{6,\ldots,2K}|M_A, Z^n_{3,\ldots,2K-1}) \\
& \quad + I(M_{12}; Y_2^n, Y_4^n, \bar{Z}^n_{6,\ldots,2K}|M_{34}, M_A, Z^n_{3,\ldots,2K-1}) \\
& = H(Y_q^n, \bar{Z}^n_{6,\ldots,2K}|M_A, Z^n_{3,\ldots,2K-1}) \\
& \quad - H(Y_q^n, \bar{Z}^n_{6,\ldots,2K}|M_{34}, M_A, Z^n_{3,\ldots,2K-1}) \\
& \quad + H(Y_2^n, Y_4^n, \bar{Z}^n_{6,\ldots,2K}|M_{34}, M_A, Z^n_{3,\ldots,2K-1}) \\
& \quad - H(Y_2^n, Y_4^n, \bar{Z}^n_{6,\ldots,2K}|M_{12}, M_{34}, M_A, Z^n_{3,\ldots,2K-1})
\end{align*}
\]
\[ H(Y^n_4, \bar{Z}^n_{\ell_6, ... 2K} | M_A, Z^n_{3, ... 2K-1}) \]
\[ + H(Y^n_2 | Y^n_4, \bar{Z}^n_{\ell_6, ... 2K}, M_{34}, M_A, Z^n_{3, ... 2K-1}) \]
\[ - H(Y^n_2, Y^n_4, \bar{Z}^n_{\ell_6, ... 2K} | M_{12}, M_{34}, M_A, Z^n_{3, ... 2K-1}) \] (4.7)

where (a) follows as all messages and noises are independent of each other; (b) by adding the side information \( \bar{Z}^n_{\ell_6, ... 2K} \). Now we bound the three terms above respectively. We start with the first term as in (Equation 4.8) – (Equation 4.12)

\[ H(Y^n_4, \bar{Z}^n_{\ell_6, ... 2K} | M_A, Z^n_{3, ... 2K-1}) \]
\[ \leq H(Y^n_4) + H(\bar{Z}^n_{\ell_6, ... 2K}) \] (4.9)
\[ \leq n(\log(P) + o(\log(P))) + H(\bar{Z}_{\ell_6, ... 2K}[k]) \] (4.10)
\[ \leq n(\log(P) + o(\log(P))) \] (4.11)
\[ + \sum_{k=1}^{n} \log(2\pi e)^{K-2} \left( 1 + \frac{h_{14}^2[k]}{h_{16}^2[k]} \right) \cdots \left( 1 + \frac{h_{14}^2[k]}{h_{1,2K}^2[k]} \right) \]
\[ = n(\log(P) + o(\log(P))) \] (4.12)

where in (a) we have used the fact that Gaussians maximize entropy subject to power constraints (which we recall are \( P \) at each user and time slot). Due to adaptation, the inputs \( X_{2m-1}, m \in \{1, 2, ..., K\} \) may be correlated, but even if all users are fully correlated and all the transmitters meet the power constraint \( P \), \( H(Y^n_4) \leq n(\log P + o(\log P)) \) as correlation only induces a power
gain inside the logarithm for a single antenna receiver. Here \( f(x) = o(\phi(x)) \) denotes the Landau little-O notation, i.e. that \( \lim_{x \to \infty} \frac{f(x)}{\phi(x)} = 0 \).

The second term can be bounded as follows in (Equation 4.13)-(Equation 4.15).

\[
H(Y^n_2|Y^n_4, \tilde{Z}^n_{6,\ldots,2K}) \leq H(Y^n_2, Y^n_3, Y^n_5, \ldots, Y^n_{2K-1}|Y^n_4, \tilde{Z}^n_{6,\ldots,2K})
\]

\[
= \sum_{k=1}^{n} [H(Y_2[k], Y_3[k], Y_5[k], \ldots, Y_{2K-1}[k]|Y^n_{2K-1}, Y^n_4, \tilde{Z}^n_{6,\ldots,2K})]
\]

\[
\leq \sum_{k=1}^{n} [H(h_{12}[k]X_1[k] + Z_2[k]|h_{14}[k]X_1[k] + Z_4[k])]
\]

\[
\leq \sum_{k=1}^{n} \log 2\pi e \left( 1 + \frac{h_{12}^2[k]P}{1 + h_{14}^2[k]P} \right)
\]

\[
= n(o(\log(P)))
\]

where in step (c) we construct \( X_6[k] \) in the conditioning because (1), \( h_{14}^{k-1}X_1^{k-1} + Z_4^{k-1} \) (this notation is meant to compactly represent the \( k-1 \) dimensional vector of the \( k-1 \) equations \( h_{14}[l]X_1[l] + Z_4[l], \) for \( l = 1, 2, \ldots, k-1 \)) can be decoded from \( Y^n_4 \) since \( X^n_3, X^n_5, \ldots, X^n_{2K-1} \) are

\(^1\text{We note that a bound of } n \log(P) + o(\log(P)) \text{ may also be shown to hold for average rather than per symbol power constraints of } P \text{ at each transmitter using Jensen's inequality and using that } 2\sqrt{P_i P_j} \leq P_i + P_j.\)
known; (2), $Z_k^{k-1}$ is known so that $h_{14}^{k-1}X_1^{k-1} + Z_6^{[k-1]}$ can be constructed; and (3), together with the knowledge of $X_3^{k-1}, X_5^{k-1}, \ldots, X_{2K-1}^{k-1}$ and (1), $Y_6^{[k-1]}$ is constructed; finally (4), perfect CSI at receivers, i.e. knowing $Y_6^{[k-1]}$ is equivalent to knowing $Y_{6-1}^{k-1}$, and combining this with the knowledge of $M_{65}$, according to the interactive encoding function we can construct $X_6[k]$. Similarly $X_8[k], \ldots, X_{2K}[k]$ can be constructed. (d) follows since Gaussians maximize conditional entropies, as in for example [50, Equation (30)].

Finally, the negative third term can be lower bounded as follows:

\[
H(Y_2^n, Y_4^n, Z_6^n, 2K | M_{12}, M_{34}, M_A, Z_3^n, 2K-1) \\
\geq H(Y_2^n, Y_4^n, Z_6^n, 2K | M_{12}, M_{34}, M_A, Z_3^n, 2K-1, X_1^n, X_3^n, \ldots, X_{2K-1}^n) \\
= H(Z_2^n, Z_4^n, Z_6^n, 2K) \\
= H(Z_2^n, Z_4^n, Z_6^n, \ldots, Z_{2K}^n) \\
= n \log 2\pi e + n \log 2\pi e + \sum_{k=1}^{n} \log(2\pi e)^{K-2} \frac{h_{14}^2[k]}{h_{16}^2[k]} \cdots \frac{h_{14}^2[k]}{h_{1,2K}^2[k]} \\
= n(\Theta(1)),
\]

where $f(x) = O(\phi(x))$ denotes that $|f(x)| < A\phi(x)$ for some constant $A$ and all values of $x$.

Now combining everything, and taking the limit,

\[
d_{12} + d_{34} \leq \limsup_{P \to \infty} \frac{R_{12} + R_{34}}{\log(P)} = 1 + 0 + 0 - 0 = 1. \quad (4.16)
\]
From the above we see that the DoF per pair of users transmitting in the same direction is 1. Summing over all rate pairs leads to the theorem.

4.3 **DoF of K-pair-user Full-duplex Two-way IC with an Instantaneous MIMO Relay**

In this section, we investigate the DoF of the K-pair-user two-way IC with an instantaneous MIMO relay with $M = 2K$ antennas in the system model described in Section 4.1.2. We then make a number of comments on how to reduce the number of antennas at the relay, at the expense of for example diminished achievable degrees of freedom, or requiring partial cognition of the messages at the relay.

4.3.1 **Achieving the Maximal DoF**

We show the main result of this section: the maximum $2K$ DoF of the K-pair-user two-way IC with an instantaneous $2K$-antenna relay is achievable.

**Theorem 21.** The full-duplex K-pair-user two-way interference channel with an instantaneous 2K-antenna relay has $2K$ degrees of freedom.

**Proof.** *Converse.* The converse is trivial since for a $2K$-user, $2K$ message unicast network where all sources and destinations have a single antenna, the maximum degrees of freedom cannot exceed $2K$ by cut-set arguments, even with adaptation/interaction at all nodes.

*Achievability.* We propose a simple “one-shot” scheme that achieves $2K$ DoF for the K-pair-user two-way IC with the help of an instantaneous $2K$-antenna relay. We consider the Gaussian channel model at high SNR, and hence noise terms are ignored from now on.
The $2K$ users each transmit a symbol $s_{ij}$ (from user $i$ to intended user $j$) and the relay receives:

$$Y_R = \sum_{i=1}^{2K} h_{i,R}s_{ij},$$

for the appropriate $j$ values, see Figure 8.

The $2K$-antenna relay (with global CSI) decodes all $2K$ symbols using a zero-forcing decoder [73], and due to the instantaneous property, transmits the following signal in the same time slot:

$$X_R = \sum_{i=1}^{2K} u_{ij}s_{ij}$$

where $u_{ij}$ denote the $2K \times 1$ beamforming vectors carrying signals from user $i$ to intended user $j$. Now at receiver 1 (for example),

$$Y_1 = \sum_{m=1}^{K} h_{2m,1}s_{2m,2m-1} + h_{R_1}^X X_R. \quad \text{(4.17)}$$

To prevent undesired signals from reaching receiver 1, the relay picks beamforming vectors such that

$$u_{ij} \in \text{null}(h_{R_1}^*) , \quad i = 3, 5, ..., 2K - 1, \ j \text{ as appropriate,} \quad \text{(4.18)}$$

where $\text{null}(A)$ denotes the null space of $A$. Since there are $2K$ antennas at the relay, $\text{null}(h_{R_1}^*)$ has dimension $2K - 1$. 
At receiver 1, the interference signals received from the relay are used to neutralize the interference signals received from the transmitters. To do this, we design the beamforming vectors to satisfy:

$$h_{2m,1} + h_{R1}^* u_{2m,2m-1} = 0, \quad m = 2, 3, ..., K.$$  \hspace{1cm} (4.19)

The $2K \times 1$ beamforming vectors satisfying the needed constraints always exist, by a dimensionality argument, along with the random channel coefficients. To see this, take $u_{34}$ as an example. We wish to construct $u_{34}$ such that the following conditions are satisfied:

$$u_{34} \in \text{null}(h_{R p}^*), \quad p = 1, 5, 7, ..., 2K - 1$$  \hspace{1cm} (4.20)

$$h_{3q} + h_{R q}^* u_{34} = 0, \quad q = 2, 6, 8, ..., 2K.$$  \hspace{1cm} (4.21)

From $u_{34} \in \text{null}(h_{R1}^*)$ (one condition in (Equation 4.20) for $p = 1$), we see that there are $2K - 1$ free parameters, which are reduced to 2 in order to satisfy the other $K - 2$ conditions in (Equation 4.20) for $p = 5, 7, \cdots, 2K - 1$, and the $K - 1$ conditions in (Equation 4.21). That is, $(2K - 1) - (K - 2) - (K - 1) = 2$. Thus, let $a, b$ be two scalars, let $A, B$ be $1 \times 2K$ vectors
such that the matrix below is invertible, then the following choice of beam forming vector (for example) will satisfy all conditions:

\[
\begin{bmatrix}
h^*_{R1} \\
h^*_{R2} \\
h^*_{R5} \\
h^*_{R6} \\
\vdots \\
h^*_{R,2K} \\
A \\
B
\end{bmatrix}^{-1}
\begin{bmatrix}
0 \\
-h_{32} \\
0 \\
-h_{36} \\
0 \\
\vdots \\
-h_{3,2K} \\
a \\
b
\end{bmatrix}.
\]

Note that all the beam forming vectors must also be chosen to satisfy the relay power constraint \( P_R \), but that we have sufficient degrees of freedom (choices of \( a, b \)) to ensure this, and that this will not affect the DoF in either case, as we will let \( P_R \to \infty \), essentially removing the power constraint.

Still at receiver 1, once the interference signals have been neutralized and the undesired signals have been nulled (by the above choice of beam forming vectors), and the self-interference (SI) signal \( s_{12} \) has been subtracted off, the received signal in (Equation 4.17) becomes

\[
Y_1 - SI = h_{21}s_{21} + h_{R1}^* u_{21}s_{21},
\]

(4.23)
from which the desired signal $s_{21}$ can be easily decoded as long as $h_{21} \neq -h_{11}^* u_{21}$, which we may guarantee by proper scaling of $u_{21}$. Similar decoding procedures are performed at all other receivers. Note that we have again assumed that self-interference in the full-duplex system is able to be perfectly removed. This provides an ideal upper bound to what is currently realizable, and including the effect of self-interference on rates is beyond the scope of this work, but an interesting topic for future investigation.

Remark 20. To achieve $2K$ DoF we have assumed full duplex operation. If instead all nodes operate in half-duplex mode, intuitively the DoF will be halved, i.e. $K$. Indeed, it is trivial to achieve $K$ DoF in a half-duplex setup: In the first time slot, all $2K$ users transmit a message, and the $2K$-antenna relay listens and decodes all $2K$ messages using a zero-forcing decoder. At time slot 2, the relay broadcasts a signal and all users listen. By careful choice of beamforming vectors as in (Equation 4.22), for example, each receiver receives only their desired message in this time slot. Therefore $2K$ desired messages are obtained in 2 time slots, i.e. $K$ DoF is achievable. Note however that in the half-duplex setting, the relay is causal rather than non-causal or instantaneous.

Remark 21. We have shown in the previous section that the DoF of the $K$-pair-user two-way interference channel is $K$; Theorem 21 implies that the addition of an instantaneous $2K$-antenna relay can increase the DoF of the $K$-pair-user two-way IC to $2K$ – it essentially cancels out all interference in both directions simultaneously. This may have interesting design implications for full duplex two-way interference networks – i.e. the ability of nodes to operate in full duplex would double the DoF of the two-way $K$-pair user IC from $K/2$ (each direction time-shares) to
the addition of a full-duplex, instantaneous MIMO relay with $2K$ antennas (for example, a pico-cell) would again double this to $2K$ DoF.

### 4.3.2 Comments on Reducing the Number of Antennas at the Instantaneous Relay

We now investigate how many DoF can be achieved by using a reduced number of antennas at the instantaneous relay. For simplicity, we first consider the (2-pair-user) two-way IC with an instantaneous 3-antenna relay, for which we propose another one-shot strategy to achieve 3 DoF. Whether this is the optimal achievable DoF is still open, i.e. unlike in all other sections so far, we have not obtained a converse.

**Theorem 22.** For the full-duplex two-way interference channel with an instantaneous 3-antenna relay, 3 degrees of freedom are achievable.

**Proof.** We now demonstrate how to transmit 3 symbols (one for each of three users) in one time slot using a 3 antenna relay. Then, this clearly achieves 3 DoF, which is larger than the 2 achievable without a relay, but smaller than the maximal value of 4 (whether anything larger than 3 is achievable is left open).

Let transmitter 1, 2 and 3 transmit symbols $s_{12}, s_{21}$ and $s_{34}$. Transmitter 4 stays silent. The relay, with three antennas is then able to use a zero-forcing receiver to obtain the three transmitted symbols, and then proceeds to transmit

$$X_R = u_{12}s_{12} + u_{21}s_{21} + u_{34}s_{34}.$$
The receivers 1, 2 and 4 (since transmitter 4 is not sending anything, receiver 3 has no desired message and we ignore it) then receive the signals:

\[
Y_1 = h_{21}s_{21} + h_{R1}^* (u_{12}s_{12} + u_{21}s_{21} + u_{34}s_{34})
\]
\[
Y_2 = h_{12}s_{12} + h_{32}s_{34} + h_{R2}^* (u_{12}s_{12} + u_{21}s_{21} + u_{34}s_{34})
\]
\[
Y_4 = h_{34}s_{34} + h_{14}s_{12} + h_{R4}^* (u_{12}s_{12} + u_{21}s_{21} + u_{34}s_{34})
\].

At receiver 1, to decode the desired \(s_{21}\), we first subtract off the self-interference term \(h_{R1}^* u_{12}s_{12}\) and then design \(u_{34}\) such that the undesired term in \(s_{34}\) disappears, i.e. make

\[
u_{34} \in \text{null}(h_{R1}^*).
\]

Then receiver 1 is able to decode \(s_{21}\) as long as \(h_{21} + h_{R1}^* u_{21} \neq 0\), which may be guaranteed by proper scaling of \(u_{21}\).

At receiver 2, to decode the desired \(s_{12}\), we first subtract off the self-interference term \(h_{R2}^* u_{21}s_{21}\), then neutralize the interference term from \(s_{34}\) by selecting \(u_{34}\) such that

\[
h_{32} + h_{R2}^* u_{34} = 0.
\]

Then receiver 2 is able to decode \(s_{12}\) as long as \(h_{12} + h_{R2}^* u_{12} \neq 0\), which may be guaranteed by proper scaling of \(u_{12}\).
Finally, at receiver 4, there is no self-interference term, only the desired term in $s_{34}$, an interference term in $s_{12}$ and an undesired signal term $s_{21}$. We may nullify the undesired signal term by selecting

$$u_{21} \in \text{null}(h_{R4}^\ast).$$  \hspace{1cm} (4.26)

Then, select $u_{12}$ to neutralize the interference by selecting

$$h_{14} + h_{R4}^\ast u_{12} = 0.$$  \hspace{1cm} (4.27)

Then receiver 2 is able to decode $s_{34}$ as long as $h_{34} + h_{R4}^\ast u_{34} \neq 0$, which may be guaranteed by proper scaling of $u_{34}$.

Each $u_{12}, u_{21}, u_{34}$ is a $3 \times 1$ vector. There is one linear constraint on $u_{12}$, one linear constraint on $u_{21}$, and two linear constraints on $u_{34}$. Hence, we have enough degrees of freedom to select all beamforming vectors to satisfy the constraints, and hence achieve 3 DoF in one time slot.

We note that this scheme sends one symbol for three of the four users, i.e. the rate $R_{43} = 0$ as no message is sent by transmitter 4. Though it does not matter from a DoF perspective (as this is defined as a sum of rates), one may symmetrize the rates by having different subsets of users transmit over 4 time slots. That is, in time slot 1, users 1,2,3 transmit. In time slot 2, users 1,2,4 transmit. In time slot 3, users 1,3,4 transmit, and in time slot 4 users 2,3,4 transmit.
In this case, 12 symbols will be transmitted over 4 time slots, and each of the 4 users is able to transmit (or receive) 3 signals in 4 time slots, again leading to 3 DoF.

The above result demonstrates that by reducing the number of antennas at the instantaneous relay from 4 to 3, we have also reduced the achievable DoF from 4 to 3. However, we want to point out that we do not currently have a converse, and more than 3 DoF may still be achievable (but clearly no more than the maximal 4). One may ask how else we might be able to reduce the number of antennas without impacting or decreasing the DoF. One way is to trade cognition for antennas, as we remark on next.

**Remark 22.** If we consider a cognitive relay (cognitive in the sense of having a-priori knowledge of messages, as first introduced in [74]), which would have access to all 4 users’ signals prior to transmission, the number of antennas at the relay can be reduced to 2, while still being able to achieve the maximum 4 degrees of freedom for the full-duplex two-way IC. The achievability scheme is trivial: the cognitive relay broadcasts all 4 signals (desired for each user) and all users listen. By careful choice of the four $2 \times 1$ beamforming vectors to nullify undesired and interference signals, and subtracting the self-interference signal, each receiver is able to obtain the desired signal. Therefore the maximal 4 DoF are achieved.

**Remark 23.** We can do even better: if the cognitive relay only knows 2 users’ signals, then we are still able to achieve the maximum 4 DoF with 2 antennas at the relay by a simple one-shot scheme. For example, assume user 1 and 2’s signals are known at the relay prior to
transmission (knowing any 2 of the 4 messages suffices). Now, each transmitter sends a message $s_{ij}$ and the relay receives 4 messages. Then the 2-antenna relay first subtracts transmitter 1 and 2’s messages and uses a zero-forcing decoder to decode the other two messages, and transmits

$$X_R = u_{12}s_{12} + u_{21}s_{21} + u_{34}s_{34} + u_{43}s_{43}.$$

At receiver 1 (for example):

$$Y_1 = h_{21}s_{21} + h_{41}s_{43}$$

$$+ h_{R1}^* u_{21}s_{21} + h_{R1}^* u_{43}s_{43} + h_{R1}^* u_{12}s_{12} + h_{R1}^* u_{34}s_{34}.$$

To decode the desired message $s_{21}$, we subtract off the self-interference signal $s_{12}$; nullify the undesired signal $s_{34}$ by designing the beamforming vector such that $h_{R1}^* u_{34} = 0$; and neutralize the interference signal $s_{43}$ by setting $h_{41} + h_{R1}^* u_{43} = 0$. A similar decoding procedure follows for the other receivers, where we note the $2 \times 1$ beamforming vectors can be always constructed by the 2-antenna relay. Therefore, each user is able to get 1 desired signal in 1 time slot and the maximal 4 DoF are achieved.

### 4.4 DoF of K-pair-user Full-duplex Two-way IC with a Causal MIMO Relay

It is known that for one-way channels where nodes are either sources of destinations of messages but not both as in a two-way setting, the usage of feedback, causal relays (possibly with multiple antennas), and cooperation does not increase the DoF of the network [50]. In the
previous section, we showed that a non-causal / instantaneous multi-antenna relay may increase the DoF of a K-pair user two-way interference channel to its maximal value of $2K$ (provided we have sufficient number of antennas). Here we show that, in sharp contrast, if the relay is actually causal, it does not increase the DoF of the K-pair-user two-way IC beyond that of a network without the relay present, which would have $K$ DoF ($K/2$ in each direction). This aligns with (and the proof uses similar techniques) the one-way results in [50] in the sense that causal relays again do not help. However, we note that full-duplex operation does increase the DoF for the two-way networks in this work, but does not for their one-way counterparts [50].

We thus consider a K-pair-user two-way IC with one causal MIMO relay which has $M$ antennas. The system model is the same as that in Section 4.1.2, where we recall that the relay is now causal, and hence

$$X_R[k] = g_k(Y_R[1], Y_R[2], \cdots Y_R[k-1]),$$

where $X_R[k]$ is an $M \times 1$ vector signal transmitted by the relay at time $k$, $g_k()$ are deterministic functions for each $k = 1, 2, \cdots n$, and $Y_R[k]$ is the $M \times 1$ vector of signals received by the relay at time slot $k$. Let $P = P_R$ for simplicity (we simply need $P$ and $P_R$ to scale to infinity at the same rate). The main result of this section is the following.

**Theorem 23.** The DoF of the K-pair-user full-duplex two-way interference channel with a causal MIMO relay is $K$. 
Proof. Achievability follows from the fact that the DoF of the $K$-pair-user two-way interference channel without a relay is $K$, as shown in Section 4.2.

Now we prove the converse. Inspired by [50], we first transform our $2K + 1$ node network to a $2K$-node network as shown in Figure 9. Since cooperation between nodes cannot reduce the DoF, we let the causal MIMO relay fully cooperate with one of the users, take user $2K - 1$ WLOG. In other words, we co-locate user $2K - 1$ and the relay or put infinite capacity links between these nodes. Then the capacity region of the original network is outer bounded by that of the following $2K$-node network which each have one message and desire 1 message as before: All users except user $2K - 1$ each have a single antenna, while user $2K - 1$ has $M + 1$ antennas (one from the original node $2K - 1$, and $M$ from the relay). Since the original relay is
connected to all $2K$ users, user $2K - 1$ in the transformed network is connected to all other users, in contrast to the original network where there is no direct link between users $1, 3, ..., 2K - 3$ and $2K - 1$. Then, letting the tilde $\tilde{\mathbf{A}}$ notation denote the inputs, outputs and channel gains of the new network, we have the correspondences (or equivalences $\equiv$ for inputs, since they may actually be different due to interaction based on different received signals)

\[
\tilde{X}_i \equiv X_i, \quad i = 1, 2, ..., 2K, \text{ except } 2K - 1, \quad \tilde{X}_{2K-1}^\top \equiv [X_{2K-1}, X_R^\top],
\]
\[
\tilde{Z}_i \equiv Z_i, \quad i = 1, 2, ..., 2K, \text{ except } 2K - 1, \quad \tilde{Z}_{2K-1}^\top \equiv [Z_{2K-1}, Z_R^\top],
\]
\[
\tilde{h}_{ij} = h_{ij}, \quad \text{for appropriate } i, j \text{ and } i, j \neq 2K - 1
\]
\[
\tilde{h}_{i,2K-1}^\top = [0, h_{iR}^\top], \quad i = 1, 3, ..., 2K - 3,
\]
\[
\tilde{h}_{i,2K-1}^\top = [h_{i,2K-1}, h_{iR}^\top], \quad i = 2, 4, ..., 2K,
\]
\[
\tilde{h}_{2K-1,j}^\top = [0, h_{Rj}^\top], \quad j = 1, 3, ..., 2K - 3,
\]
\[
\tilde{h}_{2K-1,j}^\top = [h_{2K-1,j}, h_{Rj}^\top], \quad j = 2, 4, ..., 2K,
\]

and the following input/output relationships at each channel use:

\[
\tilde{Y}_p[k] = \sum_{m=1}^{K} h_{2m,p}[k] \tilde{X}_{2m}[k] + \tilde{h}_{2K-1,p}^\top[k] \tilde{X}_{2K-1}[k] + \tilde{Z}_p[k],
\]
\[
p = 1, 3, ..., 2K - 3 \quad (4.28)
\]
\[
\tilde{Y}_q[k] = \sum_{m=1}^{K-1} h_{2m-1,q}[k] \tilde{X}_{2m-1}[k] + \tilde{h}_{2K-1,q}^\top[k] \tilde{X}_{2K-1}[k] + \tilde{Z}_q[k],
\]
\[
q = 2, 4, ..., 2K \quad (4.29)
\]
\( Y_{2K-1}[k] = \sum_{m=1,m \neq 2K-1}^{2K} h_m Y_{2K-1}[k] X_m[k] + Z_{2K-1}[k]. \) (4.30)

We have the interactive encoding functions at each node

\[ \tilde{X}_i[k] = \tilde{f}_i(M_{ij}, \tilde{Y}_i^{k-1}), \quad i = 1, 2, \ldots, 2K, \text{except } 2K-1 \] (4.31)

\[ \tilde{X}_{2K-1}[k] = \tilde{f}_{2K-1}(M_{2K-1,2K}, \tilde{Y}_{2K-1}^{k-1}) \] (4.32)

where (Equation 4.32) is where the causality of the relay is observed / incorporated.

Let \( M_A \) denote all the messages except \( M_{12}, M_{34} \), and let \( \tilde{Y}_{(2,\ldots,2K)/4} \) denote \( \tilde{Y}_2, \tilde{Y}_3, \tilde{Y}_5, \ldots, \tilde{Y}_{2K} \)

i.e. all outputs except \( \tilde{Y}_1 \) and \( \tilde{Y}_4 \). Note \( \tilde{Y}_{(2,\ldots,2K)/4} \) includes the outputs vector \( \tilde{Y}_{2K-1} \) at user \( 2K - 1 \). Similarly, \( \tilde{X}_{(2,\ldots,2K)/4} \) and \( \tilde{Z}_{(2,\ldots,2K)/4} \) denote all inputs and noises except those at nodes 1 and 4.

We now bound the sum-rate in each direction, considering the sum of a pair of rates, and starting with Fano’s inequality, we will have

\[
n(R_{12} + R_{34} - \epsilon) \leq I(M_{34}; \tilde{Y}_4^n | M_A) + I(M_{12}; \tilde{Y}_4^n, \tilde{Y}_{(2,\ldots,2K)/4} | M_{34}, M_A) \\
\leq H(\tilde{Y}_4^n | M_A) - H(\tilde{Y}_4^n | M_{34}, M_A) \\
+ H(\tilde{Y}_4^n, \tilde{Y}_{(2,\ldots,2K)/4} | M_{34}, M_A) \\
- H(\tilde{Y}_4^n, \tilde{Y}_{(2,\ldots,2K)/4} | M_{34}, M_A, M_{12}) \\
= H(\tilde{Y}_4^n | M_A) + H(\tilde{Y}_{(2,\ldots,2K)/4} | \tilde{Y}_4^n, M_{34}, M_A)
\]
\[ -H(\tilde{Z}_4^n, \tilde{Z}_{(2,\ldots,2K)/4}) \]
\[ = H(\tilde{Y}_4^n|M_A) - H(\tilde{Z}_4^n) + H(\tilde{Y}_{(2,\ldots,2K)/4}|M_{34}, M_A, \tilde{Y}_4^n) \]
\[ - H(\tilde{Z}_{(2,\ldots,2K)/4}) \]
\[ \leq \sum_{k=1}^{n} [H(\tilde{Y}_4[k]) - H(\tilde{Z}_4[k]) \]
\[ + H(\tilde{Y}_{(2,\ldots,2K)/4}|k, \tilde{Y}_{(2,\ldots,2K)/4}, M_{34}, M_A, \tilde{Y}_4^n, \tilde{X}_4^n, \tilde{X}_{(2,\ldots,2K)/4}) \]
\[ - H(\tilde{Z}_{(2,\ldots,2K)/4}[k]) \]
\[ \leq n(\log P + o(\log P)) \]
\[ + \sum_{k=1}^{n} [H(\tilde{h}_{12}[k]|\tilde{X}_1[k] + \tilde{Z}_2[k], \tilde{Z}_3[k], \ldots, \tilde{h}_{1,2K-1}[k]|\tilde{X}_1[k] \]
\[ + Z_{2K-1}[k], \tilde{h}_{1,2K}[k]|\tilde{X}_1[k] + \tilde{Z}_{2K}[k]|\tilde{h}_{14}\tilde{X}_1[k] + \tilde{Z}_4[k]) \]
\[ - H(\tilde{Z}_{(2,\ldots,2K)/4}[k]) \]
\[ \leq n(\log P + o(\log P)) + no(\log P), \]

where the last step follows as it may be shown that the Gaussian distribution maximizes conditional entropy, as done in [50, Equation (30), (31)], similar to [75, Lemma 1], and similar to (Equation 4.14), (Equation 4.15). Note also that the conditional entropy term involves a single-input, multiple output term, and hence is again bounded by no(\log P), due to the conditioning.
Similarly, in the opposite direction, let $M_B$ denote all the messages except $M_{21}, M_{43}$:

\[
\begin{align*}
&n(R_{21} + R_{43} - \epsilon) \\
&\leq I(M_{21}; \tilde{Y}_1^n | M_B) + I(M_{43}; \tilde{Y}_1^n, \tilde{Y}_{(2,...,2K)/4}| M_{21}, M_B) \\
&\leq H(\tilde{Y}_1^n | M_B) - H(\tilde{Y}_1^n | M_{21}, M_B) \\
&\quad + H(\tilde{Y}_1^n, \tilde{Y}_{(2,...,2K)/4}| M_{21}, M_B) \\
&\quad - H(\tilde{Y}_1^n, \tilde{Y}_{(2,...,2K)/4}| M_{21}, M_B, M_{43}) \\
&= H(\tilde{Y}_1^n | M_B) + H(\tilde{Y}_{(2,...,2K)/4}| \tilde{Y}_1^n, M_{21}, M_B) \\
&\quad - H(\tilde{Z}_1^n, \tilde{Z}_{(2,...,2K)/4}) \\
&= H(\tilde{Y}_1^n | M_B) - H(\tilde{Z}_1^n) + H(\tilde{Y}_{(2,...,2K)/4}| M_{21}, M_B, \tilde{Y}_1^n) \\
&\quad - H(\tilde{Z}_{(2,...,2K)/4}) \\
&\leq \sum_{k=1}^{n} [H(\tilde{Y}_1^n) - H(\tilde{Z}_1^n)] \\
&\quad + H(\tilde{Y}_{(2,...,2K)/4}| k, \tilde{Y}_{(2,...,2K)/4}, M_{21}, M_B, \tilde{X}_1^n, \tilde{Y}_{(2,...,2K)/4}) \\
&\quad - H(\tilde{Z}_{(2,...,2K)/4}| k)] \\
&\leq n(\log P + o(\log P)) \\
&\quad + \sum_{k=1}^{n} [H(\tilde{Z}_2[k], \tilde{h}_{43}[k]| \tilde{X}_4[k] + \tilde{Z}_3[k], \tilde{h}_{45}[k]| \tilde{X}_4[k] + \tilde{Z}_5[k], \cdots] \\
&\quad + h_{4,2K-1}[k]| \tilde{X}_4[k] + Z_{2K-1}[k], \tilde{Z}_{2K}[k]| \tilde{h}_{41}[\tilde{X}_4[k] + \tilde{Z}_1[k]] \\
&\quad - H(\tilde{Z}_{(2,...,2K)/4}| k)] \\
&\leq n(\log P + o(\log P)) + no(\log P),
\end{align*}
\]
Then,
\[
d_{12} + d_{34} + d_{21} + d_{43} \leq \limsup_{P \to \infty} \frac{R_{12} + R_{34} + R_{21} + R_{43}}{\log(P)} \\
\leq 1 + 0 + 1 + 0 + 0 = 2,
\]

Summing over all rate pairs (see Remark 24) leads to the theorem, which indicates that the causal MIMO relay cannot increase the DoF of the full-duplex two-pair user two-way IC.

\[\square\]

**Remark 24.** We are able to sum over all rate pairs because the asymmetry of the transformed network (multiple antennas at user $2K-1$ only, and user $2K-1$ is connected to all other nodes, unlike the even and odd numbered nodes) does not affect the DoF. Intuitively this is because for a SIMO or MISO point-to-point channel, the DoF is still 1. More rigorously, consider the sum rate pair $R_{12} + R_{2K-1,2K}$ and using similar notation (now $M_A$ denotes all messages except $M_{12}, M_{2K-1,2K}$), and following the same steps as in bounding $R_{12} + R_{34}$, we notice that the bounds do not depend on the asymmetry and again lead to 1 DoF per pair:¹

\[
n(R_{12} + R_{2K-1,2K} - \epsilon) \\
\leq I(M_{2K-1,2K}; Y_{2K}^{\infty}|M_A) \\
+ I(M_{12}; Y_{2K}^{\infty}, Y_{2K-1,2K}^{n}|M_{2K-1,2K}, M_A)
\]

¹ We leave out several steps and replace it with \(\cdots\) to avoid repetition, as these follow identically.
\[
\leq \ldots \\
\leq \sum_{k=1}^{n} [H(\tilde{Y}_{2K}[k]) - H(\tilde{Z}_{2K}[k]) \\
+ \sum_{k=1}^{n} [H(\tilde{h}_{12}[k]\tilde{X}_1[k] + \tilde{Z}_2[k], \tilde{h}_{14}[k]\tilde{X}_1[k] + \tilde{Z}_4[k], \tilde{Z}_5[k], ...) \\
- H(\tilde{Z}_{2,K-1}[k])]
\]
\[
\leq n(\log P + o(\log P)) + no(\log P),
\]

Thus we will have \(d_{12} + d_{2K-1,2K} \leq 1\). Similar arguments follow for the opposite direction.

4.5 Summary

We proposed and studied the K-pair-user two-way interference channel with and without a MIMO relay where all nodes operate in full duplex. We demonstrated that the degrees of freedom of the K-pair-user two-way IC without a relay is \(K\), which indicates that full-duplex operation doubles the DoF over the setting with half-duplex nodes for this two-way setting, but that interaction, or adapting transmission based on previously received signals at the users cannot further increase the DoF beyond what full-duplex allows, i.e. the DoF is just that of two one-way, non-interactive ICs. We next showed that if we introduce a \(2K\) antenna, full-duplex and non-causal relay, that the DoF may again be doubled over the full-duplex, relay-free counterpart (or quadrupled over the half-duplex counterpart). We demonstrated a one-shot scheme to achieve the maximal \(2K\) DoF. In sharp contrast, if the relay is causal rather than non-causal, we derived a new converse showing that the DoF cannot be increased
beyond K for a K-pair-user two-way full-duplex IC. We commented on how one may decrease
the number of antennas at the relay node, at the expense of either a reduced achievable DoF
or cognition at the relay. However, a converse for the K-pair user TWIC with an instantaneous
relay with fewer than 2K antennas is open. Overall, this chapter has shown that in K-pair-user
two-way interference channels, full-duplex operation at least doubles the achievable DoF (over
half-duplex systems), interaction does not help (unless some channel gains are zero), and a full-
duplex relay may further increase the DoF (quadrupling the DoF over a half-duplex system) if
it is instantaneous and has a sufficient number of antennas.
CHAPTER 5

CONCLUSION AND FUTURE WORK

In this thesis, we investigated wireless two-way full-duplex communication networks from an information theoretic perspective. Specifically, we explored deterministic and Gaussian noise versions of the two-way MAC/BC channel, the two-way Z channel, the two-way IC, and an extension of that, the K-pair-user two-way IC. We focussed on demonstrating channel models for which adaptation/interaction, or the ability of nodes to adapt their current channel inputs based on previously received channel outputs, is useless or may only provide limited gains from a capacity region perspective, i.e. non-adaptive schemes achieve outer bounds derived or to within constant gaps for the fully or partially adaptive models.

The question of when adaptation is useless in general networks remains a challenging open question. However, based on some of the examples seen here, we believe that the following properties may be needed to make the claim that “adaptation is useless” for a particular network: 1) the self-interference can be cancelled (excludes the binary multiplier channel), 2) no loop in the networks (excludes the relaying of data along stronger paths), and 3) no “coherent” gains (excluding possible gains by having users use adaptation to create joint input distributions in for example Gaussian networks).

In addition, we characterized the degrees of freedom of several wireless two-way full-duplex networks, in particular, the K-pair-user TWIC with and without a MIMO relay is studied. Our results show that 1) Full-duplex operation doubles the DoF over the setting with half-

123
duplex nodes. 2) Interaction, or adapting transmission based on previously received signals at users cannot further increase the DoF. 3) A causal MIMO relay cannot increase the DoF of the $K$-pair-user full-duplex TWIC. 4) A non-causal $2K$-antenna relay is able to cancel all the interference in both directions simultaneously can achieve the maximum $2K$ DoF.

Overall, in working towards interactive and cooperative communications, we have studied the fundamental limits of wireless two-way full-duplex communication networks using information theory. Our theoretical findings may provide upper bounds on the data rate limits of multi-user two-way full-duplex communications in real cellular networks, which may become a reality in the near future.

There are many interesting problems that may be considered for future research in the direction started by this thesis. To name a few, 1) Investigate whether adaptation is useless for a more general class of noisy two-way networks where the noise can be arbitrarily random, i.e. non-stationary, non-ergodic. This is a natural extension of the "deterministic, invertible and cardinality constrained" models studied in Chapter 2, and we can start with the point-to-point two-way channel with arbitrary noise. 2) Further study the relationship between interaction and feedback, i.e., whether, if perfect output feedback cannot increase capacity for a certain channel, this implies interaction between users is useless. We believe the answer is affirmative and may possibly be verified by using private and common randomness arguments at all nodes. 3) From an information theoretic perspective, determining a single channel for which we may show capacity exactly and in which adaptation is useful. We believe this has to date not been shown for a two-way network and is likely very challenging. 4) Characterize the degrees of
freedom of general interactive two-way full-duplex networks, e.g. consider a fully-connected, MIMO-user network with and without different types of relays. This work may bring insight into the impact of techniques such as massive MIMO and advanced interference management, which have been considered as candidate technologies for improving the spectral efficiency for next generation wireless communications.
APPENDICES
Appendix A

PROOF OF THEOREM 4

Proof. Time-sharing may again be used to achieve this region. For the converse,

Proof of bound (Equation 2.8):

\[ n(R_{12} + R_{32} + R_{34} - \epsilon) \]

\[ \leq I(M_{12}; Y_{2}^{n}|M_{21}, M_{23}, M_{43}) + I(M_{32}, M_{34}; Y_{2}^{n}, Y_{4}^{n}|M_{43}, M_{12}, M_{21}, M_{23}) \]

\[ \leq I(M_{12}; Y_{2}^{n}|M_{21}, M_{23}, M_{43}) + I(M_{32}, M_{34}; Y_{2}^{n}|M_{43}, M_{12}, M_{21}, M_{23}) \]

\[ \leq H(Y_{2}^{n}|M_{21}, M_{23}, M_{43}) - H(Y_{2}^{n}|M_{12}, M_{21}, M_{23}, M_{43}) \]

\[ + H(Y_{4}^{n}|M_{43}, M_{12}, M_{21}, M_{23}, Y_{2}^{n}) \]

\[ \leq \sum_{i=1}^{n} [H(Y_{2,i}) + H(Y_{4,i}|M_{12}, M_{21}, M_{23}, M_{43}, Y_{4}^{i-1}, Y_{2}^{n})] \]

\[ \leq \sum_{i=1}^{n} [H(Y_{2,i}) + H(X_{3,i} \oplus X_{4,i}|M_{12}, M_{21}, M_{23}, M_{43}, Y_{4}^{i-1}, X_{4,i}^{i-1}, X_{3,i}^{i-1}, Y_{2}^{n}, X_{2}^{n})] \]

\[ \leq \sum_{i=1}^{n} [H(Y_{2,i}) + H(X_{3,i}|M_{12}, M_{21}, M_{23}, M_{43}, Y_{4}^{i-1}, X_{4,i}^{i-1}, X_{3,i}^{i-1}, X_{1,i}^{n} \oplus X_{2}^{n} \oplus X_{3}^{n}, X_{2}^{n}, X_{1}^{n})] \]

\[ = \sum_{i=1}^{n} [H(Y_{2,i})] \leq n \]

where (a) follows from the chain rule. We drop two negative entropy terms in inequality (b) and notice that the second and the third entropy terms cancel each other. In (c), we apply
the chain rule first, then we drop the conditioning part of the first entropy term. In (d), we construct \(X_4^i = f_4(M_{43}, Y_{4}^{i-1})\) and note that \(X_3^{i-1}\) may be obtained from \(Y_{4}^{i-1} = X_3^{i-1} \oplus X_4^{i-1}\), given \(X_4^{i-1}\). Adding \(X_2^n\) follows from the fact \(X_2^n = f_2(M_{21}, M_{23}, Y_{2}^{n-1})\). In (e), we cancel \(X_{4,i}\) in the second entropy term since we know \(X_4^i\). In addition, given \(M_{12}\) and \(X_2^n\), we may construct \(X_1^n\) as illustrated in Figure 10. Now, we may obtain \(X_3^n\) from \(Y_2^n = X_1^n \oplus X_2^n \oplus X_3^n\), so that the second entropy term in zero. Bound (Equation 2.9) follows by symmetry. \(\square\)
Appendix B

EVALUATION OF THE SUM-RATE OUTER BOUND WITH FULL ADAPTATION IN GAUSSIAN TWO-WAY IC OF THEOREM 13

Letting $E[X_jX_k^*] = \lambda_{jk}$, suppressing the subscript $i$, and assuming a symmetric channel, the first two terms in (Equation 3.19) may be bounded as

\[
H(g_{12}X_1 + g_{32}X_3 + Z_2|X_2) - H(Z_2) \\
\leq H(g_{12}X_1 + g_{32}X_3 + Z_2) - H(Z_2) \\
\leq \log 2\pi e (\text{Var}(g_{12}X_1 + g_{32}X_3) + 2\text{Cov}(g_{12}X_1, g_{32}X_3) + 1) - \log 2\pi e(1) \\
= \log (\text{SNR} + \text{INR} + 2|\lambda_{13}|\cos \theta \sqrt{\text{SNR} \times \text{INR}} + 1) \quad (B.1)
\]

where $\theta$ is the angle of $g_{12}g_{32}^\dagger \lambda_{13}$. Similarly, the last two terms may be bounded as

\[
H(g_{34}X_3 + Z_4|g_{32}X_3 + Z_2, X_4) - H(Z_4) \leq \log \left( \frac{\text{Var}(g_{34}X_3 + Z_4|g_{32}X_3 + Z_2, X_4)}{\sigma_4^2} \right) \\
\leq \log \left( 1 + \frac{\text{SNR}(1 - |\lambda_{34}|^2)}{\text{INR}(1 - |\lambda_{34}|^2) + 1} \right). \quad (B.2)
\]

To obtain (Equation 3.18) one may verify that the sum of (Equation B.1) and (Equation B.2) is maximized at $\lambda_{34} = 0$ and $\lambda_{13} = 1, \theta = \theta = 0$. 
Appendix C

COPYRIGHT RELEASE STATEMENTS

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line 2011 IEEE.

2) In the case of illustrations or tabular material, we require that the copyright line [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.

3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior authors approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/credit notice should be placed prominently in the references: [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]

2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
Appendix C (Continued)

3) In placing the thesis on the author’s university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity’s name goes here]’s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.
CITED LITERATURE


NAME: Zhiyu Cheng

EDUCATION: B.S., Information Engineering, East China University of Science and Technology (ECUST), Shanghai, China, 2009

M.S., Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, Illinois, 2012

Ph.D., Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, Illinois, 2014

HONORS: SCID Innovation Scholarship, Shanghai Chemical Industry District (5% acceptance), 2007

First Prize (top 2%) in National Mathematical Contest in Modeling, 2007

Meritorious Winner (top 13%) in international Mathematical Contest in Modeling, 2008

Outstanding Graduate Award, Shanghai Education Committee (5% acceptance), 2009

Academic Excellence Scholarship, ECUST (15% acceptance), 2005-2009

Graduate Fellowship, China Scholarship Council (5% acceptance), 2009-2013

PROFESSIONAL EXPERIENCE: R&D Intern, Motorola China Research Center, Shanghai, China, November 2008-May 2009

Instructor, College of Engineering, University of Illinois at Chicago, Chicago, Illinois, August 2011-December 2011
Teaching Assistant, Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, Illinois, August 2011-May 2012

Summer Research Intern, Bell Laboratories, Alcatel-Lucent, Murray Hill, NJ, June 2013-August 2013

Research Assistant, Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, Illinois, January 2010-May 2014


"Two-way Networks: when Adaptation is Useless", at University of Illinois at Urbana-Champaign, February 20, 2014. Hosted by 9th Coordinated Science Lab Student Conference.


Z. Cheng and N. Devroye, Degrees of Freedom of the Two-way Interference Channel with a Non causal Multi-antenna Relay, presented at the IEEE Global Communications Conference (Globecom), Atlanta, GA, December 2013.

Z. Cheng and N. Devroye, Two-way Networks: when Adaptation is Useless, the IEEE Transactions on Information Theory, March 2014.


POSTERS: "Outer Bounds to the Capacity Region of the Two-Way Z Channel," presented at the IEEE 2011 School of Information Theory, University of Texas at Austin, TX, May 27-30, 2011.


MEMBERSHIP & SERVICE: Student Member of the IEEE, www.ieee.org
Member of the IEEE Information Theory Society, www.itsoc.org
Member of the IEEE Communications Society, www.comsoc.org

Reviewer for the IEEE Transactions on Information Theory, the IEEE Journal on Selected Areas in Communications, the IEEE Transactions on Wireless Communications, the IEEE Transactions on Communications, the IEEE Transactions on Vehicular Technology, the IEEE Communications Surveys and Tutorials, EURASIP Journal on Wireless Communications and Networking, and a few IEEE conferences.