# Multiplexing Gains under Mixed-Delay Constraints on Wyner's Soft-Handoff Model

**Homa Nikbakht [1],\*, Michèle Angela Wigger [1],\* and Shlomo Shamai (Shitz) [2],\***

[1]  Laboratoire Traitement et Communication de l'Information (LTCI), Telecom Paris, Institut polytechnique de Paris, 91120 Palaiseau, France
[2]  Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa 32000, Israel
\*  Correspondence: homa.nikbakht@telecom-paris.fr (H.N.); michele.wigger@telecom-paristech.fr (M.A.W.); sshlomo@ee.technion.ac.il (S.S.)

**Abstract:** This paper analyzes the multiplexing gains (MG) achievable over Wyner's soft-handoff model under mixed-delay constraints, that is, when delay-sensitive and delay-tolerant data are simultaneously transmitted over the network. In the considered model, delay-sensitive data cannot participate or profit in any ways from transmitter or receiver cooperation, but delay-tolerant data can. Cooperation for delay-tolerant data takes place over rate-limited links and is limited to a fixed number of cooperation rounds. For the described setup, inner and outer bounds are derived on the set of MG pairs that are simultaneously achievable for delay-sensitive and delay-tolerant data. The bounds are tight in special cases and allow us to obtain the following conclusions. For large cooperation rates, and when both transmitters and receivers can cooperate, it is possible to simultaneously attain maximum MG for delay-sensitive messages and maximum sum MG for all messages. For comparison, in scheduling schemes (also called time-sharing schemes), the largest achievable sum MG decreases linearly with the MG of delay-sensitive messages. A similar linear decrease is proved for any coding scheme, not only for scheduling schemes, if only transmitters or only receivers can cooperate (but not both) and if delay-sensitive messages have moderate MG. In contrast, if the MG of delay-sensitive messages is small, the maximum sum MG can be achieved even with only transmitter or only receiver cooperation. To summarise, when cooperation rates are high and both transmitters and receivers can cooperate or when delay-sensitive messages have small MG, then transmitting delay-sensitive messages causes no penalty on the sum-MG. In other regimes, this penalty increases proportionally to the delay-tolerant MG in the sense that increasing the delay-sensitive MG by $\Delta$ penalises the largest achievable delay-tolerant MG by $2\Delta$ and thus the sum MG by $\Delta$.

**Keywords:** mixed-delay constraints; multiplexing gain region; Wyner's soft-handoff model

## 1. Introduction

One of the major challenges of today's wireless communication networks is to design coding schemes for transmission of heterogeneous traffic types. For example, different data streams (pertaining to different applications) can be subject to different delay constraints. Such mixed delay constraints in wireless networks have recently been studied in References [1–5]. In particular, Reference [1] proposes a broadcasting approach over a single-antenna fading channel to communicate a stream of "fast" messages, which have to be sent over a single coherence block, and a stream of "slow" messages, which can be sent over multiple blocks. A similar approach was taken in Reference [2] but for a broadcast scenario with $K$ users. Instead of superposing "slow" on "fast" messages, this latter work proposes a scheduling approach to give preference to the communication of "fast" messages. A scheduling

algorithm that prioritizes "fast" messages over "slow" messages was also proposed in Reference [3]. In particular, "fast" messages can be stored in the buffer for only one scheduling period.

A related scenario was introduced in Reference [4] for cloud radio access networks (C-RAN). In Reference [4], the messages sent by the mobile users close to the base stations (BS) are directly decoded at these BSs, whereas messages from users located further away are decoded at the cloud processor. In our terminology, the messages sent by close by users are "fast" messages because they will incur smaller decoding delay, and the messages sent by further away users are "slow" messages because their decoding is performed at the cloud processor and thus takes more time. In the spirit of this interpretation, we considered a similar approach in Reference [5], but where in our setup each user can send both "fast" and "slow" messages and "fast" messages have to be decoded immediately at the BS, whereas "slow" messages are decoded at the central processor. Moreover, in Reference [5] the channel from the mobile users to the BSs is modelled as a fading channel. The results in Reference [5] show that for small front haul capacities from the BSs to the cloud processor it is beneficial, in terms of sum-rate, to send both "fast" and "slow" messages. However, when the rate of "fast" messages is already large, then increasing it further, deteriorates the sum-rate of the system. In this regime, the stringent delay constraints on the "fast" messages penalise the overall performance.

In this paper, we consider a cellular network without a cloud processor, but where neighbouring BS or/and neighbouring mobile users can cooperate over dedicated *cooperation links* that do not interfere with the main communications channel. These cooperation links can model the backhauls between BSs or bluetooth or microwave links between neighbouring mobiles. We consider a scenario in which the transmitters have both delay-sensitive and delay-tolerant messages to transmit to their corresponding receivers. In our setup, delay-sensitive messages, the "fast" messages, cannot profit from cooperation due to their stringent delay constraints. That means they cannot participate in the transmitter (Tx) cooperation phase and they also have to be decoded prior to the receiver (Rx) cooperation phase. Delay-tolerant messages, the "slow" messages, can profit from both Tx- and Rx-cooperation. This specific problem formulation has strong similarity with (and was also inspired by) the model studied by Huleihel and Steinberg [6,7] where cooperation links may be absent and some of the messages are only sent if these links are present.

The focus of this paper is on the pairs of *Multiplexing Gains (MG)*, also called degree of freedom or capacity prelog, that are simultaneously achievable for the "fast" and "slow" messages in our setup. We consider Wyner's soft-handoff model also known as one-dimensional Wyner model [8,9]. Notice that cooperation in this network has been studied in various works including [10–17]. The focus of [10] is to identify associations between Rxs and Txs, that maximize the average MG across both uplink and downlink sessions, using cooperative transmission and reception schemes between Txs. More closely related are the works in [11,12]. In particular [12] is related to our setup when only "slow" messages are transmitted. In [12], neighbouring Txs can cooperate with each other over a fixed number of rounds and neighbouring Rxs can cooperate with each other over a fixed number of rounds. It is proved that for small cooperation prelogs, a single cooperation round at the Txs or at the Rxs achieves the same MG as when the number of cooperation rounds is unlimited. On the other hand, for large cooperation prelogs, the maximum per user MG increases with every additional cooperation round that is permitted either at the Txs or at the Rxs.

In the present work, we thus extend the work in Reference [12] to accommodate not only "slow" messages that can tolerate the delays from cooperation, but also "fast" messages that have to be encoded and decoded without further delay and thus cannot profit from cooperation. Notice that the standard approach to combine the transmissions of delay-tolerant and delay-sensitive data is to apply a smart scheduling algorithm and thus to time-share a scheme for only delay-tolerant data with a scheme for only delay-sensitive data. Since the maximum MG attained for only delay-tolerant data is larger (there are less constraints imposed on this transmission) than the maximum MG attained for only delay-sensitive data, this approach can achieve the maximum sum-MG only when exclusively sending delay-tolerant data. More specifically, for scheduling schemes, the sum-MG decreases linearly with the

MG of delay-sensitive data. In this paper, we determine the set of all achievable delay-sensitive and delay-tolerant MG pairs, that is, the *optimal MG region*, in the function of the prelogs of the cooperation links and the total number of cooperation rounds allowed for "slow" messages. The obtained results show that (for Wyner's soft-handoff model) when only Txs or only Rxs can cooperate, transmitting "fast" messages at low MG does not penalise the sum-MG of "slow" and "fast" messages. In contrast, when the MG of "fast" messages is large, this is not the case and increasing the MG of "fast" messages by Δ comes at the expense of decreasing the MG of "slow" messages by 2Δ and the sum MG by Δ. When the cooperation rates are sufficiently large and both Txs and Rxs can cooperate, then it is possible to accommodate the largest possible MG for delay-sensitive messages without decreasing the maximum sum-MG. The stringent delay constraints thus do not harm the overall performance in this scenario.

To achieve the described performance, we propose a new coding scheme where every second Tx sends a "fast" message and the other Txs send a "slow" message or no message at all. Due to the structure of Wyner's soft-handoff network, communication of "fast" messages is only interfered by transmissions of "slow" messages. This interference can thus be described during the Tx-cooperation phase and precanceled at the Txs sending the "fast" messages. On the other hand, Rxs that have to decode "fast" messages do this without further delay and describe their decoded messages during the Rx-cooperation phase to their adjacent Rxs. Since we alternated the transmission of "fast" and "slow" messages across Tx/Rx-pairs, these adjacent Rxs decode "slow" messages. With the obtained cooperation messages they can thus first subtract the interference from the "fast" messages and then decode their own "slow" messages. The described mechanism allows interference-free transmission of "fast" messages to be accommodated on every second Tx/Rx pair without disturbing the transmission of "slow" messages. Employing an optimal coding scheme for the transmission of "slow" messages on all other Tx/Rx pairs will then give the same over-all performance as when using an optimal coding scheme to send a "slow" message on each and every Tx/Rx pair. This explains why with Tx- and Rx-cooperation the maximum MG can be attained even with a "fast" MG of L/2, where L denotes the number of antennas at each Tx and Rx. Notice that this is the largest MG when only "fast" messages but no "slow" messages are transmitted.

### 1.1. Organization

The rest of this paper is organised as follows. We end this section with some remarks on notation. The following Section 2 describes the problem setup. Section 3 presents our results when only transmitters or only receivers can cooperate and Section 4 the results when transmitters *and* receivers can cooperate. Section 5 concludes the main body of the paper. Technical proofs of the converse results are referred to in appendices.

### 1.2. Notation

We use the shorthand notations "Rx" for "Receiver" and "Tx" for "Transmitter". The set of all integers is denoted by $\mathbb{Z}$, the set of positive integers by $\mathbb{Z}^+$ and the set of real numbers by $\mathbb{R}$. For other sets we use calligraphic letters, for example, $\mathcal{X}$. Random variables are denoted by uppercase letters, for example, $X$, and their realizations by lowercase letters, for example, $x$. For vectors we use boldface notation, that is, upper case boldface letters such as $\mathbf{X}$ for random vectors and lower case boldface letters such as $\mathbf{x}$ for deterministic vectors.) Matrices are depicted with sans serif font, for example, H. We also write $X^n$ for the tuple of random variables $(X_1, \dots, X_n)$ and $\mathbf{X}^n$ for the tuple of random vectors $(\mathbf{X}_1, \dots, \mathbf{X}_n)$.

## 2. Problem Description

Consider Wyner's soft-handoff network with $K$ Txs and $K$ Rxs that are aligned on two parallel lines so that each Tx $k$ has two neighbours, Tx $k-1$ and Tx $k+1$, and each Rx $k$ has two neighbours, Rx $k-1$ and Rx $k+1$. Interference is short-range in the sense that the signal sent by Tx $k$ is observed

only by Rx $k$ and by the neighbouring Rx $k + 1$ (see Figure 1). Let Txs and Rxs be equipped with $L > 0$ antennas each. The time-$t$ channel output at Rx $k$ is then described as

$$Y_{k,t} = H_{k,k}X_{k,t} + H_{k-1,k}X_{k-1,t} + Z_{k,t},$$ (1)

where $X_{k,t}$ and $X_{k-1,t}$ are the real $L$-dimensional vectors sent by Tx $k$ and Tx $k - 1$ at time $t$; $\{Z_{k,t}\}$ is a noise sequence consisting of i.i.d. standard Gaussian vectors; $H_{k,k}$ and $H_{k-1,k}$ are fixed full rank channel matrices; and $X_{0,t} = 0$ for all $t$.
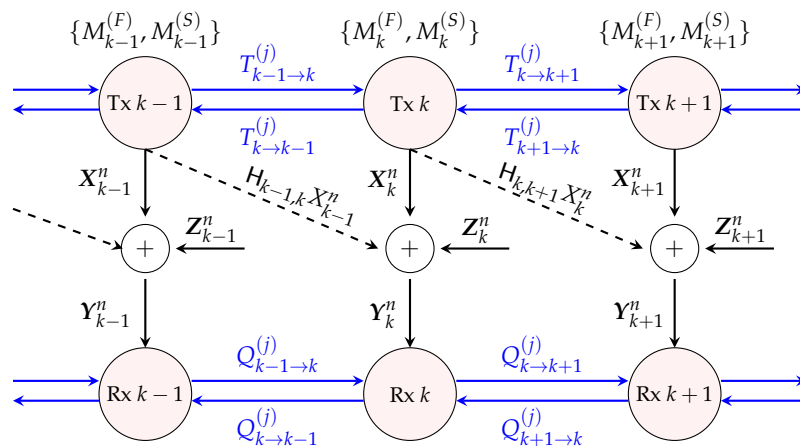


**Figure 1.** System model with transmitter (Tx-) and receiver (Rx-)cooperation.

Each Tx $k \in \{1, \dots, K\}$ wishes to send a pair of independent messages $M_k^{(F)}$ and $M_k^{(S)}$ to Rx $k$. The "fast" message $M_k^{(F)}$ is uniformly distributed over the set $\mathcal{M}_k^{(F)} \triangleq \{1, \dots, \lfloor 2^{nR_k^{(F)}} \rfloor\}$ and needs to be decoded subject to a stringent delay constraint, as we explain shortly. The "slow" message $M_k^{(S)}$ is uniformly distributed over $\mathcal{M}_k^{(S)} \triangleq \{1, \dots, \lfloor 2^{nR_k^{(S)}} \rfloor\}$ and is subject to a less stringent decoding delay constraint. Here, $n$ denotes the blocklength of transmission and $R_k^{(F)}$ and $R_k^{(S)}$ the rates of transmissions of the "fast" and "slow" messages.

We consider three different cooperation scenarios:

1.  Neighbouring Txs cooperate by communicating during $D_{Tx} > 0$ rounds over dedicated cooperation links. Rxs cannot cooperate, and so the number of Rx-cooperation rounds is $D_{Rx} = 0$. (This scenario is termed "Tx-cooperation Only")
2.  Neighbouring Rxs cooperate by communicating during $D_{Rx} > 0$ rounds over dedicated cooperation links. Txs cannot cooperate, and so the number of Tx-cooperation rounds is $D_{Tx} = 0$. (Termed "Rx-cooperation Only")
3.  Neighbouring Txs cooperate during $D_{Tx} > 0$ rounds over dedicated cooperation links and neighbouring Rxs cooperate during $D_{Rx} > 0$ rounds. (Termed "Tx- and Rx-cooperation").

The cooperative communication is subject to a total delay constraint

$$D_{Tx} + D_{Rx} \leq D,$$ (2)

where $D > 0$ is a given parameter of the system. In the "Tx-cooperation Only" scenario $D_{Rx}$ has to be 0 and thus $D_{Tx} \leq D$. Similarly in the "Rx-cooperation Only" scenario, $D_{Tx} = 0$ and $D_{Rx} \leq D$. For "Tx- and Rx-cooperation" the values of $D_{Tx}$ and $D_{Rx}$ are design parameters and can be chosen arbitrary such that (2) is satisfied. As we will see, in our setup the cooperative communication only concerns "slow" messages, because "fast" messages are subject to stringent delay constraint and thus have to be transmitted and decoded without further delay.

We describe the encoding at the Txs. In the case of Tx-cooperation, neighbouring Txs can communicate to each other over dedicated noise-free, but rate-limited, links. Communication takes place over $D_{Tx} > 0$ rounds and can depend only on the "slow" messages but not on the "fast" messages. In each cooperation round $j \in \{1, \ldots, D_{Tx}\}$, Tx $k$ produces a cooperation message $T_{k \to \ell}^{(j)}$ for each of its neighbours $\ell \in \{k-1, k+1\}$ by computing

$$T_{k \to \ell}^{(j)} = \xi_{k \to \ell}^{(n)} \left( M_k^{(S)}, \{ T_{\ell' \to k}^{(1)}, \ldots, T_{\ell' \to k}^{(j-1)} \}_{\ell' \in \{k-1, k+1\}} \right), \quad j \in \{1, \ldots, D_{Tx}\}, \ \ell \in \{k-1, k+1\}, \quad (3)$$

for some function $\xi_{k \to \ell}^{(n)}$ on appropriate domains. Tx $k$ sends the messages $T_{k \to \ell}^{(1)}, \ldots, T_{k \to \ell}^{(D_{Tx})}$ over the cooperation link to Tx $\ell \in \{k-1, k+1\}$. The rate-limitation on the cooperation link imposes

$$\sum_{j=1}^{D_{Tx}} H(T_{k \to \ell}^{(j)}) \leq \mu_{Tx} \cdot \frac{n}{2} \log(P), \qquad k \in \{1, \ldots, K\}, \ \ell \in \{k-1, k+1\}, \quad (4)$$

for a given $\mu_{Tx} > 0$.

Tx $k$ finally computes its channel inputs $\mathbf{X}_k^n = (\mathbf{X}_{k,1}, \ldots, \mathbf{X}_{k,n}) \in \mathbb{R}^{L \times n}$ as a function of its "fast" and "slow" messages and of all the $2D_{Tx}$ cooperation messages that it obtained from its neighbouring transmitters:

$$\mathbf{X}_k^n = \tilde{f}_k^{(n)} \left( M_k^{(F)}, M_k^{(S)}, \{ T_{\ell' \to k}^{(1)}, \ldots, T_{\ell' \to k}^{(D_{Tx})} \}_{\ell' \in \{k-1, k+1\}} \right). \quad (5)$$

In the setup without Tx-cooperation, Tx $k$ computes its channel inputs $\mathbf{X}_k^n$ simply as a function of its "fast" and "slow" messages:

$$\mathbf{X}_k^n = f_k^{(n)} \left( M_k^{(F)}, M_k^{(S)} \right). \quad (6)$$

In any case (i.e., with and without Tx-cooperation), the channel inputs have to satisfy the average block-power constraint

$$\frac{1}{n} \sum_{t=1}^{n} ||\mathbf{X}_{k,t}||^2 \leq P, \quad \forall k \in \{1, \ldots, K\}, \quad (7)$$

almost surely.

We now describe the decoding. In the case of Rx-cooperation, decoding takes place in two phases. During the first *fast-decoding phase*, each Rx $k$ decodes its intended "fast" message $M_k^{(F)}$ based on its own channel outputs $\mathbf{Y}_k^n = (\mathbf{Y}_{k,1}, \ldots, \mathbf{Y}_{k,n}) \in \mathbb{R}^{L \times n}$. So, it produces:

$$\hat{M}_k^{(F)} = g_k^{(n)}(\mathbf{Y}_k^n), \quad (8)$$

where $g_k^{(n)}$ denotes a decoding function on appropriate domains.

In the subsequent *slow-decoding phase*, Rxs first communicate with their neighbours during $D_{Rx} \geq 0$ rounds over dedicated noise-free and rate-limited links, and then they decode their intended "slow" messages based on their outputs and based on this exchanged information. Specifically, in each cooperation round $j \in \{1, \ldots, D_{Rx}\}$, each Rx $k$, for $k \in \{1, \ldots, K\}$, produces a cooperation message $Q_{k \to \ell}^{(j)}$ for each of its neighbours $\ell \in \{k-1, k+1\}$:

$$Q_{k \to \ell}^{(j)} = \psi_{k,\ell}^{(n)} \left( \mathbf{Y}_k^n, \left\{ Q_{\ell' \to k}^{(1)}, \ldots, Q_{\ell' \to k}^{(j-1)} \right\}_{\ell' \in \{k-1, k+1\}} \right), \quad (9)$$

for an encoding function $\psi_{k,\ell}^{(n)}$ on appropriate domains. Rx $k$ then sends the messages $Q_{k \to \ell}^{(1)}, \ldots, Q_{k \to \ell}^{(D_{Rx})}$ over the cooperation link to Rx $\ell \in \{k-1, k+1\}$. The rate-limitation on the cooperation link imposes

$$\sum_{j=1}^{D_{Rx}} H(Q_{k\to\ell}^{(j)}) \le \mu_{Rx} \cdot \frac{n}{2}\log(P), \qquad k \in \{1,\dots,K\}, \ \ell \in \{k-1,k+1\}, \tag{10}$$

for some given $\mu_{Rx} > 0$.

After the last cooperation round, each Rx $k$ decodes its desired "slow" messages as

$$\hat{M}_k^{(S)} = b_k^{(n)}\left(\mathbf{Y}_k^n, \left\{Q_{\ell'\to k}^{(1)},\dots,Q_{\ell'\to k}^{(D_{Rx})}\right\}_{\ell'\in\{k-1,k+1\}}\right), \tag{11}$$

where $b_k^{(n)}$ denotes a decoding function on appropriate domains.

For each of the three cooperation scenarios, given cooperation prelogs $\mu_{Rx}, \mu_{Tx} \ge 0$ and maximum delay D, a MG pair $(\mathsf{S}^{(F)}, \mathsf{S}^{(S)})$ is called *achievable*, if for every positive integer $K$ there exists a sequence of average rates $\{R_K^{(F)}(P), R_K^{(S)}(P)\}_{P>0}$ so that

$$\mathsf{S}^{(F)} \triangleq \varlimsup_{K\to\infty} \varlimsup_{P\to\infty} \frac{R_K^{(F)}(P)}{\frac{1}{2}\log(P)}, \tag{12}$$

$$\mathsf{S}^{(S)} \triangleq \varlimsup_{K\to\infty} \varlimsup_{P\to\infty} \frac{R_K^{(S)}(P)}{\frac{1}{2}\log(P)}, \tag{13}$$

and so that for each average rate pair $(R_K^{(F)}(P), R_K^{(S)}(P))$ it is possible to find a set (in the blocklength $n$) of encoding, cooperation, and decoding functions satisfying constraints (2), (4), (7), and (10) and with vanishing probability of error:

$$p(\text{error}) \triangleq \mathbb{P}\left[\bigcup_{k\in\{1,\dots,K\}}\left((\hat{M}_k^{(F)} \ne M_k^{(F)}) \cup (\hat{M}_k^{(S)} \ne M_k^{(S)})\right)\right] \to 0 \quad \text{as} \quad n \to \infty. \tag{14}$$

The closure of the set of all achievable MG pairs $(\mathsf{S}^{(F)}, \mathsf{S}^{(S)})$ is called *optimal MG region*. In the case of Tx-cooperation only, it is denoted $\mathcal{S}_{Tx}^\star(\mu_{Tx}, D)$, in the case of Rx-cooperation only $\mathcal{S}_{Rx}^\star(\mu_{Rx}, D)$, and in the case of Tx- and Rx-cooperation $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$.

## 3. Rx- or Tx-Cooperation Only

In the following two subsections, we consider the Rx-cooperation only scenario and the Tx-cooperation only scenario. For each scenario we present coding schemes and the optimal MG region. The scenario with both Tx-and Rx-cooperation is treated in the next Section 4.

*3.1. Optimal MG Region and Coding Schemes for Rx-Cooperation Only*

**Theorem 1** (Optimal Multiplexing Gain Region: Rx-cooperation Only). *For any given $\mu_{Rx} > 0$, the MG region $\mathcal{S}_{Rx}^\star(\mu_{Rx}, D)$ is the set of all nonnegative pairs $(\mathsf{S}^{(F)}, \mathsf{S}^{(S)})$ satisfying*

$$2\mathsf{S}^{(F)} + \mathsf{S}^{(S)} \le \mathsf{L} \tag{15}$$

$$\mathsf{S}^{(F)} + \mathsf{S}^{(S)} \le \min\left\{\frac{\mathsf{L}}{2} + \mu_{Rx}, \ \mathsf{L}\cdot\frac{2D+1}{2D+2}\right\}. \tag{16}$$

**Proof.** The converse to (16) follows by extending the proof in Reference [12] to the multi-antenna case and by noting that the sum MG of "slow" and "fast" messages cannot be larger than the MG of a scenario with only "slow" messages. The converse to (15) is proved in Appendix A. For the achievability, define the following five MG pairs:

$$\left( \mathsf{S}^{(F)} \;=\; \frac{\mathsf{L}}{2},\; \mathsf{S}^{(S)} = 0 \right), \tag{17a}$$

$$\left( \mathsf{S}^{(F)} \;=\; 0,\; \mathsf{S}^{(S)} = \mathsf{L} \cdot \frac{2D+1}{2D+2} \right), \tag{17b}$$

$$\left( \mathsf{S}^{(F)} \;=\; 0,\; \mathsf{S}^{(S)} = \frac{\mathsf{L}}{2} + \mu_{\mathrm{Rx}} \right), \tag{17c}$$

$$\left( \mathsf{S}^{(F)} \;=\; \frac{\mathsf{L}}{2D+2},\; \mathsf{S}^{(S)} = \mathsf{L} \cdot \frac{2D}{2D+2} \right), \tag{17d}$$

$$\left( \mathsf{S}^{(F)} \;=\; \frac{\mathsf{L}}{2} - \mu_{\mathrm{Rx}},\; \mathsf{S}^{(S)} = 2\mu_{\mathrm{Rx}} \right). \tag{17e}$$

In the following Section 3.1.1 we show that when $\mu_{\mathrm{Rx}} \geq \mu_{\max}$, where

$$\mu_{\max} \triangleq \mathsf{L} \cdot \frac{D}{2D+2}, \tag{18}$$

the MG pairs (17a,b,d) are achievable. When $\mu_{\mathrm{Rx}} < \mu_{\max}$ the MG pairs (17a,c,e) are achievable. The proof of achievability of Theorem 1 then follows from simple time-sharing arguments. □

Figure 2 depicts the MG region in Theorem 1 for different values of $\mu_{\mathrm{Rx}}$. When there are only "slow" messages, the maximum MG is $\min\{\frac{\mathsf{L}}{2} + \mu_{\mathrm{Rx}}, \mathsf{L} \cdot \frac{2D+1}{2D+2}\}$. Notice that in any scheme, we can replace a "fast" message by a "slow" message. By a rate-transfer argument, the maximum sum-MG thus coincides with the maximum "slow" MG. Interestingly, this sum MG remains unchanged whenever the "fast" MG $\mathsf{S}^{(F)}$ is below a certain threshold. Mathematically, this is described by the slope of the boundary of the region being equal to $-1$ when

$$\mathsf{S}^{(F)} \leq \max\left\{ \frac{\mathsf{L}}{2} - \mu_{\mathrm{Rx}}, \frac{\mathsf{L}}{2D+2} \right\}. \tag{19}$$

For

$$\mathsf{S}^{(F)} > \max\left\{ \frac{\mathsf{L}}{2} - \mu_{\mathrm{Rx}}, \frac{\mathsf{L}}{2D+2} \right\}, \tag{20}$$

the slope is $-2$. In this latter regime, increasing the MG of "fast" messages by $\Delta$ requires decreasing the MG of "slow" messages by $2\Delta$. There is thus a penalty in sum MG caused by the more stringent delay constraints on "fast" messages.
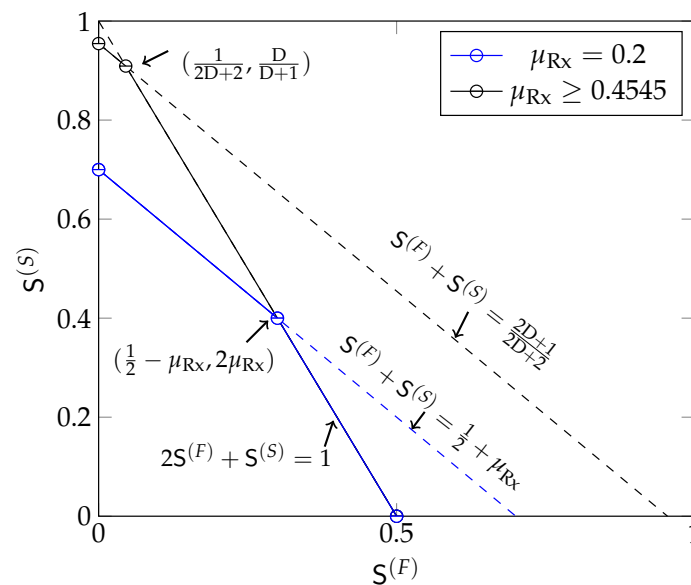


**Figure 2.** The optimal MG region $\mathcal{S}^{\star}_{\mathrm{Rx}}(\mu_{\mathrm{Rx}}, D)$ for Rx-cooperation only, for different values of $\mu_{\mathrm{Rx}}$ and $D = 10$ and $\mathsf{L} = 1$.

### 3.1.1. Schemes Proving Achievability of Theorem 1

We prove achievability of the MG pairs in (17).

1. MG pair in (17a): Periodically silence every second Tx. This splits the network into $\lceil K/2 \rceil$ non-interfering point-to-point links. Send a "fast" message over each of these links (see Figure 3), but no "slow" message at all. The described scheme achieves the MG pair in (17a) and requires no cooperation rate.

2. MG pairs in (17b,c): Let the Txs only send "slow" messages but no "fast" messages. Under this coding assumption, the setup at hand is a multi-antenna version of the setup in Reference [12], but specialized to 0 Tx-cooperation rounds and D Rx-cooperation rounds. The multi-antenna extension of the scheme proposed in Reference [12] (Section V) can thus be used to achieve the MG pair in (17b) if $\mu_{\text{Rx}} \geq \mu_{\max}$ and the MG pair in (17c) if $\mu_{\text{Rx}} < \mu_{\max}$.

For reference in the following subsection, we briefly review the scheme in Reference [12] (Section V) when specialized to Rx-cooperation only. For details, see Reference [12]. Consider first the case $\mu_{\text{Rx}} \geq \mu_{\max}$. In this case, the scheme periodically silences every $2D + 2$nd Tx. This splits the network into smaller subnets, each consisting of $2D + 1$ active Txs and $2D + 2$ active Rxs. We describe the communication in the first subnet, see also Figure 4; the others are treated in an analogous way.
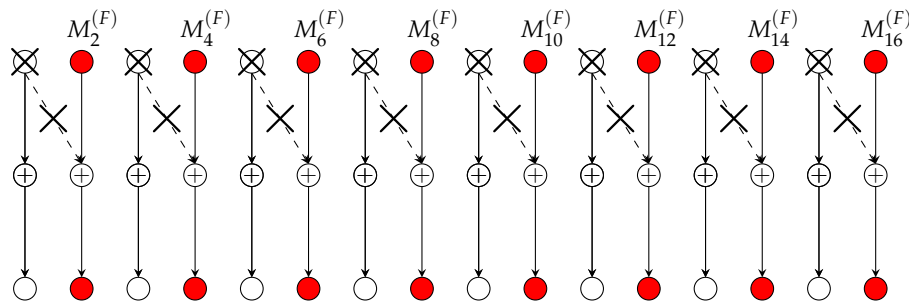


**Figure 3.** Scheme achieving Multiplexing Gain (MG) pair (17a) where only "fast" messages are transmitted.
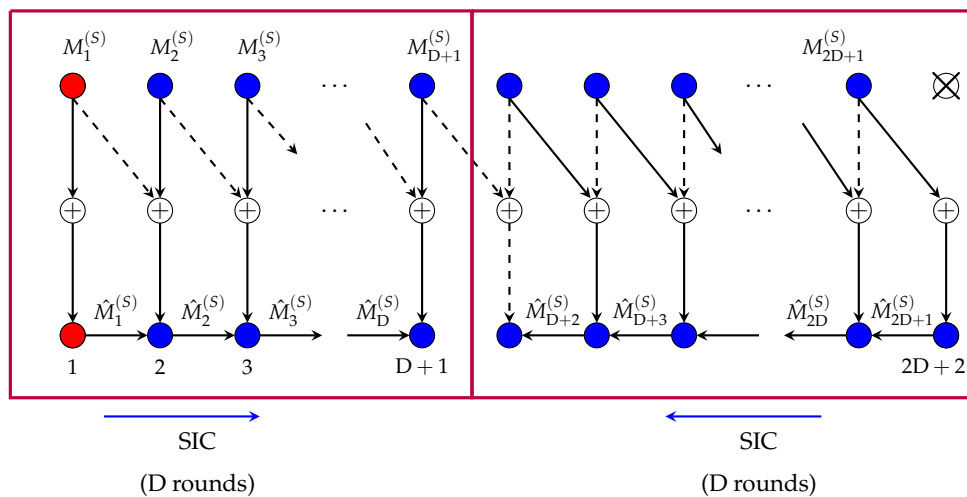


**Figure 4.** Scheme for Rx-cooperation only.

Each Tx $k \in \{1, \dots, 2D + 1\}$ in this first subnet encodes its "slow" message $M_k^{(S)}$ using an L-dimensional Gaussian codebook and then sends the resulting codeword using its L Tx-antennas over the channel. Decoding is performed as follows. Rx 1 decodes its desired message using an optimal point-to-point decoding method based on the interference-free channel outputs $Y_1^n = \mathsf{H}_{1,1} X_1^n + Z_1^n$. Then it sends its decoded message $\hat{M}_1^{(S)}$ over the cooperation link to Rx 2 during the first cooperation round. Rxs 2 to $D + 1$ apply successive interference cancellation (SIC) where they cancel the interference

from the preceding Tx with the cooperation message obtained from their left neighbour. After decoding its intended "slow" message, each Rx $k \in \{2, \ldots, D\}$ sends its decoded message $\hat{M}_k^{(S)}$ over the cooperation link to Rx $k + 1$ during cooperation round $k$.

We now describe decoding at Rxs $D + 2, \ldots, 2D + 2$. Recall that Tx $2D + 2$ is silenced. Therefore Rx $2D + 2$ observes the interference-free channel outputs $Y_{2D+2}^n = H_{2D+1,2D+2} X_{2D+1}^n + Z_{2D+2}^n$. Based on these outputs, Rx $2D + 2$ decodes the "slow" message $M_{2D+1}^{(S)}$ intended for Rx $2D + 1$ and transmits the decoded message $\hat{M}_{2D+1}^{(S)}$ to this Rx over the cooperation link in round 1. Rxs $D + 2$ to $2D + 1$ declare the cooperation message that they receive from their right neighbour as their desired message. They also employ SIC to decode the "slow" message intended for the neighbour to their left. Finally, after this decoding step, each Rx $k \in \{D + 3, \ldots, 2D + 2\}$ sends the decoded message $\hat{M}_{k-1}^{(S)}$ over the cooperation link to its left neighbour during cooperation round $2D + 3 - k$. Figure 4 illustrates the decodings and conferenced messages.

In the described scheme, $2D + 1$ Txs send a "slow" message using an L-dimensional Gaussian codebook of power P and all these messages can be decoded based on interference-free outputs. An average "slow" MG of $L \cdot \frac{2D+1}{2D+2}$ is thus achieved in each subnet. Moreover, 2D cooperation messages are sent in each subnet, each of prelog equal to the rate of a "slow" message, i.e., L. The *average* cooperation prelog *per link* is thus $L \cdot \frac{2D}{2(2D+2)} = \mu_{\max}$. If one time-shares $2D + 2$ different instances of the described scheme with a different subset of silenced users in each of them, the overall scheme achieves the MG pair $(S^{(F)} = 0, S^{(S)} = L \cdot \frac{2D+1}{2D+2})$ with each cooperation link being loaded at average cooperation prelog $\mu_{\max}$.

When $\mu_{Rx} < \mu_{\max}$, we can time-share the scheme achieving (17b) with a scheme that deactivates every second Tx and sends "slow" messages over the interference-free links. This latter scheme does not require any cooperation. Time-sharing is done according to the available cooperation prelog $\mu_{Rx}$: the first scheme that uses cooperation prelog $\mu_{\max}$ is used over a fraction $\frac{\mu_{Rx}}{\mu_{\max}}$ of time and the no-cooperation scheme over the remaining fraction $1 - \frac{\mu_{Rx}}{\mu_{\max}}$ of time. The combined scheme then requires cooperation prelog $\mu_{Rx}$ and achieves the MG pair in (17c).

3. MG pairs in (17d,e): Reconsider the coding scheme that achieves MG pair (17b) and that is described in the previous subsection and illustrated in Figure 4. A close inspection of the scheme reveals that in each subnet, decoding of the message sent by the left-most Tx does not rely on the conferenced information. This first message of each subnet thus satisfies our decoding requirement for "fast" messages.

We propose to apply the above scheme, but to let the first Tx of every subnet (the red Tx in Figure 4) send a "fast" message and the subsequent 2D Txs of the subnet send "slow" messages. This modified scheme requires the same cooperation prelog $\mu_{\max}$ as before and it achieves the MG pair in (17d).

For setups where $\mu_{Rx} < \mu_{\max}$, we propose to time-share the scheme achieving (17d) over a fraction $\frac{\mu_{Rx}}{\mu_{\max}}$ of time with the scheme achieving (17a) over the remaining fraction $1 - \frac{\mu_{Rx}}{\mu_{\max}}$ of time. This time-sharing scheme has cooperation prelog equal to $\mu_{Rx}$, and thus respects the constraint (10). Moreover, it achieves the MG pair in (17e).

### 3.2. Optimal MG Region and Coding Schemes for Tx-Cooperation Only

**Theorem 2** (Optimal MG region: Tx-cooperation Only). *For any given $\mu_{Tx} > 0$, the MG region $\mathcal{S}_{Tx}^\star(\mu_{Tx}, D)$ is the set of all nonnegative pairs $(S^{(F)}, S^{(S)})$ satisfying*

$$2S^{(F)} + S^{(S)} \leq L \tag{21}$$

$$S^{(F)} + S^{(S)} \leq \min \left\{ \frac{L}{2} + \mu_{Tx}, \ L \cdot \frac{2D+1}{2D+2} \right\}. \tag{22}$$

**Proof.** The converse to (22) follows by extending the proof in Reference [12] to the multi-antenna case and by noting that the sum MG cannot be larger than the MG of a scenario with only "slow" messages. The converse to (21) is proved in Appendix B. For the achievability, define the following MG pairs:

$$\left( \mathsf{S}^{(F)} = 0, \; \mathsf{S}^{(S)} = \mathsf{L} \cdot \frac{2D+1}{2D+2} \right), \tag{23a}$$

$$\left( \mathsf{S}^{(F)} = 0, \; \mathsf{S}^{(S)} = \frac{\mathsf{L}}{2} + \mu_{\text{Tx}} \right), \tag{23b}$$

$$\left( \mathsf{S}^{(F)} = \frac{\mathsf{L}}{2D+2}, \; \mathsf{S}^{(S)} = \mathsf{L} \cdot \frac{2D}{2D+2} \right), \tag{23c}$$

$$\left( \mathsf{S}^{(F)} = \frac{\mathsf{L}}{2} - \mu_{\text{Tx}}, \; \mathsf{S}^{(S)} = 2\mu_{\text{Tx}} \right). \tag{23d}$$

In the following Section 3.2.1 we show that when $\mu_{\text{Tx}} \geq \mu_{\max}$ the MG pairs (17a) and (23a,c) are achievable and when $\mu_{\text{Tx}} < \mu_{\max}$ the MG pairs (17a) and (23b,d) are achievable. The achievability proof of the theorem then follows by simple time-sharing arguments. □

**Remark 1.** *Notice the duality between Theorems 1 and 2, which show that cooperation is equally beneficial for only Tx- or only Rx-cooperation. As we will see in Section 4, it is however more beneficial, when Txs and Rxs can cooperate.*

### 3.2.1. Schemes Proving the Achievability of Theorem 2

We prove achievability of the MG pairs in (23). MG pair (17a) is achievable as described in the previous section (no cooperation is required at all).

1. MG pairs in (23a,b): Let the Txs only send "slow" messages but no "fast" messages. Under this coding assumption, the introduced setup corresponds to a multi-antenna version of the setup in Reference [12] but specialized to D Tx-cooperation rounds and 0 Rx-cooperation rounds. Achievability of MG pairs (23a,b) then follows immediately by specializing [12] (Theorem 1) to Tx-cooperation only. In the following we briefly describe the schemes achieving (23a,b). For details see Reference [12].

We silence every $2D + 2$nd Tx. This splits the network into non-interfering subnets, and in a given subnet we apply the scheme depicted in Figure 5. Specifically, Tx 1 encodes its message using an L-dimensional power-P Gaussian point-to-point codebook, and sends the resulting codeword $X_1^n$ using its L Tx-antennas over the channel. It also precodes the obtained sequence with the matrix $\mathsf{H}_{2,2}^{-1}\mathsf{H}_{1,2}$, quantises the precoded sequence $I_1^n \triangleq \mathsf{H}_{2,2}^{-1}\mathsf{H}_{1,2}X_1^n$ with a rate-$\frac{\mathsf{L}}{2}\log(1+P)$ quantiser to obtain a quantisation $\hat{I}_1^n$ at noise level, and sends the resulting quantisation message as a first-round cooperation message to Tx 2. For each $k = 2, \ldots, D+1$, Tx $k$ obtains a round-$(k-1)$ cooperation message from its left neighbour Tx $k-1$ that describes the quantised version $\hat{I}_{k-1}^n$ of $I_{k-1}^n \triangleq \mathsf{H}_{k,k}^{-1}\mathsf{H}_{k-1,k}X_{k-1}^n$. Based on this message, Tx $k$ reconstructs $\hat{I}_{k-1}^n$, encodes its "slow" message $M_k^{(S)}$ using a power P dirty-paper code (DPC) that mitigates the interference $\hat{I}_{k-1}^n$, and sends the resulting DPC sequence $X_k^n$ over the channel. Moreover, it precodes this input sequence with the matrix $\mathsf{H}_{k+1,k+1}^{-1}\mathsf{H}_{k,k+1}$, quantises the precoded sequence $I_k^n \triangleq \mathsf{H}_{k+1,k+1}^{-1}\mathsf{H}_{k,k+1}X_k^n$ with a rate-$\mathsf{L}/2\log(1+P)$ quantiser (for a quantisation at noise level) to obtain $\hat{I}_k^n$, and sends the quantisation message as a round-$k$ cooperation message over the link to its right neighbour. Tx $D+1$ produces its inputs in a similar way, that is, using DPC, but sends no cooperation message at all.
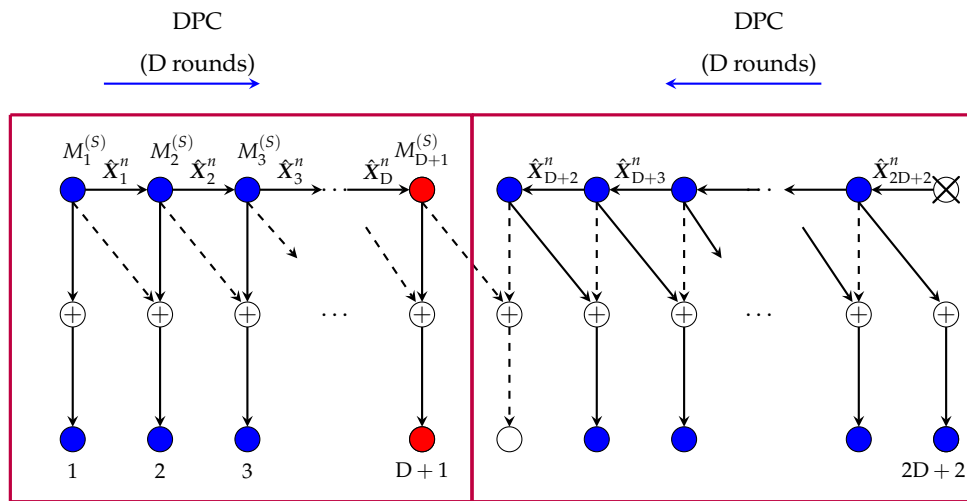
**Figure 5.** Scheme for Tx-cooperation only.

Rx 1 decodes $M_1^{(S)}$ based on the interference-free outputs

$$Y_1^n = H_{1,1} X_1^n + Z_1^n, \tag{24}$$

using a standard point-to-point decoding rule. Each Rx $k \in \{2, \dots, D+1\}$ decodes its desired message $M_k^{(S)}$ based on the premultiplied outputs

$$H_{k,k}^{-1} Y_k^n = H_{k,k}^{-1} H_{k-1,k} X_{k-1}^n + X_k^n + H_{k,k}^{-1} Z_k^n, \tag{25}$$

using an optimal DPC decoding rule. (Recall that $X_k^n$ was produced as a DPC sequence that mitigates $\hat{I}_{k-1}^n$, a quantised version of $I_{k-1}^n = H_{k,k}^{-1} H_{k-1,k} X_{k-1}^n$). Since quantisation was performed at noise level, each message $M_1^{(S)}, \dots, M_{D+1}^{(S)}$ can be sent reliably with MG L.

Each message $M_k$, with $k \in \{D+3 \dots 2D+2\}$, is sent over the path Tx $k \to$ Tx $k-1 \to$ Rx $k$. We describe the transmissions in more detail, starting with the last Tx in the subnet. Tx $2D+2$ does not send any channel inputs, that is, $X_{2D+2}^n = 0^n$. However, it first encodes its "slow" message $M_{2D+2}^{(S)}$ using an L-dimensional Gaussian point-to-point codebook, precodes the codeword $U_{2D+2}^n$ by the matrix $H_{2D+1,2D+2}^{-1}$, and then quantises this precoded codeword $S_{2D+1}^n \triangleq H_{2D+1,2D+2}^{-1} U_{2D+2}^n$ with a rate-$L/2 \log(1+P)$ to obtain a quantisation $\hat{S}_{2D+1}^n$ at noise level. It finally sends the quantisation message describing $\hat{S}_{2D+1}^n$ as a first-round cooperation message to Tx $2D+1$. Tx $2D+1$ reconstructs $\hat{S}_{2D+1}^n$ and sends it over the channel, that is, $X_{2D+1}^n = \hat{S}_{2D+1}^n$.

In a similar way, each Tx $k \in \{2D+1, \dots, D+2\}$ encodes its own "slow" message $M_k^{(S)}$ by means of DPC of power P that mitigates the interference $H_{k-1,k}^{-1} H_{k,k} X_k^n$ of the signal sent by Tx $k$ itself; precodes the obtained sequence $U_k^n$ with the matrix $H_{k-1,k}^{-1} H_{k,k}$; quantises the precoded sequence $S_{k-1}^n \triangleq H_{k-1,k}^{-1} H_{k,k} U_k^n$ to obtain a quantisation $\hat{S}_{k-1}$ at noise level; and sends the corresponding quantisation message as a $(2D+3-k)$-round cooperation message over the link to Tx $k-1$. Tx $k-1$ then reconstructs $\hat{S}_{k-1}^n$ and sends it over the channel: $X_{k-1}^n = \hat{S}_{k-1}^n$. Rxs $D+2, \dots, 2D+1$ decode their intended messages using an optimal DPC decoding rule based on the premultiplied outputs

$$H_{k-1,k}^{-1} Y_k^n = X_{k-1}^n + H_{k-1,k}^{-1} H_{k,k} X_k^n + H_{k-1,k}^{-1} Z_k^n. \tag{26}$$

Recall that $X_{k-1}^n$ is a quantised version (at noise level) of the precoded signal $S_{k-1}^n \triangleq H_{k-1,k}^{-1} H_{k,k} U_k^n$, where $U_k^n$ is a DPC sequence that mitigates the interference $H_{k-1,k}^{-1} H_{k,k} X_k^n$. Each of the messages $M_{D+3}^{(S)}, \dots, M_{2D+2}^{(S)}$ can thus be transmitted reliably at full MG L.

In the described scheme, an average "slow" MG of $\mathsf{L} \cdot \frac{2D+1}{2D+2}$ is thus achieved in each subnet. Moreover, 2D cooperation messages of prelog L are sent in each subnet, and the *average* cooperation prelog *per link* is $\mathsf{L} \cdot \frac{2D}{2(2D+2)} = \mu_{\max}$. If one time-shares $2D + 2$ different instances of the described scheme with a different subset of silenced users in each of them, the overall scheme achieves the MG pair $(S^{(F)} = 0, S^{(S)} = \mathsf{L}\frac{2D+1}{2D+2})$ with each cooperation link being loaded at average cooperation prelog $\mu_{\max}$.

When $\mu_{\mathrm{Tx}} < \mu_{\max}$, we propose to time-share above described scheme over a fraction $\frac{\mu_{\mathrm{Tx}}}{\mu_{\max}}$ of time with a scheme that deactivates every second Tx and sends "slow" messages over the interference-free links (which does not require any cooperation) over the remaining fraction $1 - \frac{\mu_{\mathrm{Tx}}}{\mu_{\max}}$ of time. The overall time-sharing scheme achieves the MG pair (23b) and loads each Tx-cooperation link at prelog $\mu_{\mathrm{Tx}}$.

2. MG pairs in (23c,d): A close inspection of the coding scheme described above and depicted in Figure 5 reveals that in each subnet, the message pertaining to the $D + 1$st Tx does not participate in the cooperation, see Figure 5. That means, all conferenced information is independent of this message. The message thus satisfies the constraints imposed on "fast" messages in our scenario. We thus propose to employ above scheme, but where the $D + 1$st Tx in each subnet (the red Tx in Figure 5) sends a "fast" message and the first and the last D Txs in the subnet send "slow" messages. This scheme requires again cooperation prelog $\mu_{\max}$ and achieves the MG pair in (23c).

When $\mu_{\mathrm{Tx}} < \mu_{\max}$, we can time-share this scheme over a fraction $\frac{\mu_{\mathrm{Tx}}}{\mu_{\max}}$ of time with the scheme achieving (17a) over the remaining fraction $1 - \frac{\mu_{\mathrm{Tx}}}{\mu_{\max}}$ of time. The time-shared scheme achieves the MG pair (23d) and loads each Tx-cooperation link at orelog $\mu_{\mathrm{Tx}}$.

## 4. Both Tx-and Rx-Cooperation

In this section we consider both Tx- and Rx-cooperation. Recall that the number of Tx- and Rx-cooperation rounds $D_{\mathrm{Tx}}$ and $D_{\mathrm{Rx}}$ is a design parameter over which we can optimize subject to the sum-constraint $D_{\mathrm{Tx}} + D_{\mathrm{Rx}} \leq D$. For simplicity, in this section we assume that the total number of cooperation rounds D is even.

In Section 4.1 we present our inner and outer bounds on the MG region. We also prove that they match in some cases. In the following subsections we then present the coding schemes that allow us to conclude our achievability result.

### 4.1. Results on MG Region

Let the maximum number of total cooperation rounds D be given. For any pair $D_{\mathrm{Rx}} \in \{1, \ldots, D-1\}$ and $D_{\mathrm{Tx}} \in \{1, \ldots, D-1\}$ summing to less than D, define

$$\mu_{\mathrm{Tx,L}}(D_{\mathrm{Tx}}) \triangleq \mathsf{L} \cdot \frac{D_{\mathrm{Tx}}}{2D+2}, \tag{27}$$

$$\mu_{\mathrm{Rx,L}}(D_{\mathrm{Rx}}) \triangleq \mathsf{L} \cdot \frac{D_{\mathrm{Rx}}}{2D+2}, \tag{28}$$

$$\mu_{\mathrm{Tx,H}}(D_{\mathrm{Tx}}) \triangleq \mathsf{L} \cdot \frac{\frac{D}{2} + \frac{3}{4}D_{\mathrm{Tx}} - \frac{1}{4}}{2D+2}, \tag{29}$$

$$\mu_{\mathrm{Rx,H}}(D_{\mathrm{Rx}}) \triangleq \mathsf{L} \cdot \frac{\frac{D}{2} + D_{\mathrm{Rx}} - 1}{2D+2}. \tag{30}$$

Notice that $\mu_{\mathrm{Tx,L}}(D_{\mathrm{Tx}}) \leq \mu_{\mathrm{Tx,H}}(D_{\mathrm{Tx}})$ and $\mu_{\mathrm{Rx,L}}(D_{\mathrm{Rx}}) \leq \mu_{\mathrm{Rx,H}}(D_{\mathrm{Rx}})$.

Also, define the five MG pairs:

$$S_{\text{NoCoop}}^{(F)} \triangleq \left(S^{(F)} = \frac{L}{2}, S^{(S)} = 0\right), \tag{31a}$$

$$S_{\text{NoCoop}}^{(S)} \triangleq \left(S^{(F)} = 0, S^{(S)} = \frac{L}{2}\right), \tag{31b}$$

$$S_{\text{Coop}} \triangleq \left(S^{(F)} = 0, S^{(S)} = L \cdot \frac{2D+1}{2D+2}\right), \tag{31c}$$

$$S_{\text{Partial}} \triangleq \left(S^{(F)} = L \cdot \frac{2}{2D+2}, S^{(S)} = L \cdot \frac{2D-1}{2D+2}\right), \tag{31d}$$

$$S_{\text{Interlaced}} \triangleq \left(S^{(F)} = \frac{L}{2}, S^{(S)} = L \cdot \frac{D}{2D+2}\right). \tag{31e}$$

Notice that all these MG pairs do not depend on the number of cooperation rounds $D_{\text{Tx}}$ and $D_{\text{Rx}}$. In what follows, we will be interested in convex combinations of these points and therefore define for each $\alpha \in [0,1]$:

$$S_{\text{Coop}}(\alpha) \triangleq \alpha \cdot S_{\text{Coop}} + (1-\alpha) \cdot S_{\text{NoCoop}}^{(S)}, \tag{32a}$$

$$S_{\text{Partial}}(\alpha) \triangleq \alpha \cdot S_{\text{Partial}} + (1-\alpha) \cdot S_{\text{NoCoop}}^{(F)}, \tag{32b}$$

$$S_{\text{Interlaced}}(\alpha) \triangleq \alpha \cdot S_{\text{Interlaced}} + (1-\alpha) \cdot S_{\text{NoCoop}}^{(F)}, \tag{32c}$$

$$S_{\text{Partial-Inter}}(\alpha) \triangleq \alpha \cdot S_{\text{Interlaced}} + (1-\alpha) \cdot S_{\text{Partial}}. \tag{32d}$$

Notice that $S_{\text{Coop}}(1) = S_{\text{Coop}}$ and $S_{\text{Partial}}(1) = S_{\text{Partial}}$ and $S_{\text{Interlaced}}(1) = S_{\text{Interlaced}}$. Moreover, $S_{\text{Partial-Inter}}(1) = S_{\text{Interlaced}}$ and $S_{\text{Partial-Inter}}(0) = S_{\text{Partial}}$.

**Theorem 3** (Achievable MG Region: Tx- and Rx-cooperation)**.** *For any choice of odd-valued integers $D_{Tx}, D_{Rx} \in \{1, 3, 5, \dots, D-1\}$ summing to D, the optimal MG region $\mathcal{S}^\star(\mu_{\text{Tx}}, \mu_{\text{Rx}}, D)$ contains some of the following regions, depending on the available cooperation prelogs $\mu_{\text{Tx}}$ and $\mu_{\text{Rx}}$.*

- *If $\mu_{\text{Tx}} \geq \mu_{Tx,H}(D_{Tx})$ and $\mu_{\text{Rx}} \geq \mu_{Rx,H}(D_{Rx})$, the optimal MG region $\mathcal{S}^\star(\mu_{\text{Tx}}, \mu_{\text{Rx}}, D)$ contains the trapezoidal region*

$$\text{conv hull}\left((0,0), S_{\text{NoCoop}}^{(F)}, S_{\text{Coop}}, S_{\text{Interlaced}}\right). \tag{33}$$

- *If $\mu_{\text{Tx}} \geq \mu_{Tx,L}(D_{Tx})$ and $\mu_{\text{Rx}} \geq \mu_{Rx,L}(D_{Rx})$, the optimal MG region $\mathcal{S}^\star(\mu_{\text{Tx}}, \mu_{\text{Rx}}, D)$ contains the pentagon*

$$\text{conv hull}\left((0,0), S_{\text{NoCoop}}^{(F)}, S_{\text{Coop}}, S_{\text{Partial-Inter}}(\alpha_1^\star), S_{\text{Interlaced}}(\beta_1^\star)\right), \tag{34}$$

  *where*

$$\alpha_1^\star \triangleq \min\left\{\frac{\mu_{\text{Tx}} - \mu_{Tx,L}(D_{Tx})}{\mu_{Tx,H}(D_{Tx}) - \mu_{Tx,L}(D_{Tx})}, \frac{\mu_{\text{Rx}} - \mu_{Rx,L}(D_{Rx})}{\mu_{Rx,H}(D_{Rx}) - \mu_{Rx,L}(D_{Rx})}, 1\right\} \tag{35}$$

  *and*

$$\beta_1^\star \triangleq \min\left\{\frac{\mu_{\text{Tx}}}{\mu_{Tx,H}(D_{Tx})}, \frac{\mu_{\text{Rx}}}{\mu_{Rx,H}(D_{Rx})}, 1\right\}. \tag{36}$$

- *For $\mu_{\text{Tx}} \leq \mu_{Tx,L}(D_{Tx})$ or $\mu_{\text{Rx}} \leq \mu_{Rx,L}(D_{Rx})$, the optimal MG region $\mathcal{S}^\star(\mu_{\text{Tx}}, \mu_{\text{Rx}}, D)$ contains the region*

$$\text{conv hull}\left((0,0), S_{\text{Coop}}(\alpha_2^\star), S_{\text{Partial}}(\alpha_2^\star), S_{\text{Interlaced}}(\beta_2^\star)\right). \tag{37}$$

  *where*

$$\alpha_2^\star \triangleq \min\left\{\frac{\mu_{\text{Tx}}}{\mu_{Tx,L}(D_{Tx})}, \frac{\mu_{\text{Rx}}}{\mu_{Rx,L}(D_{Rx})}\right\} \tag{38}$$

*and*

$$\beta_2^\star = \min \left\{ \frac{\mu_{Tx}}{\mu_{Tx,H}(D_{Tx})}, \frac{\mu_{Rx}}{\mu_{Rx,H}(D_{Rx})} \right\}. \tag{39}$$

In Figure 6 we schematically illustrate above MG regions (33), (34) and (37). We see that for large cooperation prelogs our MG region is the trapezoid in Figure 6a. For smaller cooperation prelogs the MG region turns into a pentagon, see Figure 6b, because MG pair $S_{Interlaced}$ is not included anymore. Finally, for even smaller cooperation prelogs even the MG pair $S_{Coop}$ is not included anymore, but needs to be replaced by $S_{Coop}(0.93)$. Similarly, $S_{Partial-Inter}(0.6)$ needs to be replaced by $S_{Partial}(0.93)$.



(**a**) MG region in (33)  (**b**) MG region in (34)  (**c**) MG region in (37)
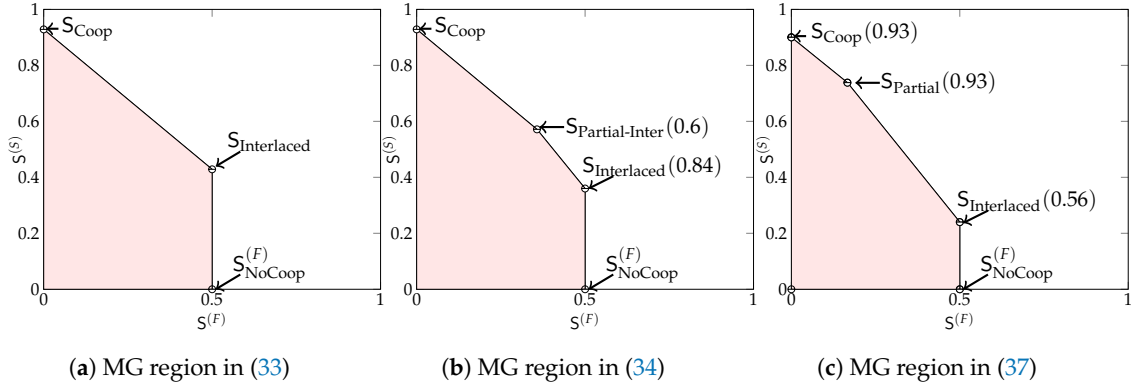
**Figure 6.** Examples of the three MG regions in (33), (34) and (37). Specifically, we used $D_{Tx} = 3$, $D_{Rx} = 3$, and $\mu_{Tx} \in \{0.4, 0.3, 0.2\}$ and $\mu_{Rx} \in \{0.4, 0.3, 0.2\}$.

The achievable MG region described in the theorem can also be written as a union over the choice of the Tx- and Rx-cooperation rounds $D_{Tx}$ and $D_{Rx}$ summing to no more than D. Notice however, that one cannot take the convex hull of this union because the way we defined the problem setup the choice of $D_{Tx}$ and $D_{Rx}$ needs to be fixed in advance and time-sharing between different choices is not possible.

**Proof of Theorem 3.** In the following Sections 4.2–4.4 we show how to achieve the MG pairs in (31c–e) with sufficiently large cooperation prelogs $\mu_{Tx}$ and $\mu_{Rx}$. In particular, to achieve (31c,d), cooperation prelogs $\mu_{Tx} \geq \mu_{Tx,L}(D_{Tx})$ and $\mu_{Rx} \geq \mu_{Rx,L}(D_{Rx})$ are required. To achieve (31e) cooperation prelogs $\mu_{Tx} \geq \mu_{Tx,H}(D_{Tx})$ and $\mu_{Rx} \geq \mu_{Rx,H}(D_{Rx})$ are required. MG pairs (31a,b) can be achieved without any Tx- or Rx-cooperation by simply silencing every second transmitter and sending either only "fast" or only "slow" messages over the remaining $K/2$ isolated point-to-point links.

The proof of the theorem follows then by simple time-sharing arguments. In particular, for any $\alpha \in [0,1]$ the MG pair $S_{Coop}(\alpha)$ can be achieved by time-sharing the scheme achieving $S_{Coop}$ over a fraction $\alpha$ of the time with the scheme achieving $S_{NoCoop}^{(S)}$ over the remaining fraction of time. Such a time-sharing scheme requires cooperation prelogs of $\mu_{Tx} \geq \alpha\mu_{Tx,L}$ and $\mu_{Rx} \geq \alpha\mu_{Rx,L}$. The MG pairs $S_{Partial}(\alpha)$ and $S_{Interlaced}(\alpha)$ are achieved by time-sharing the scheme achieving $S_{Partial}$ or the scheme achieving $S_{Interlaced}$ over a fraction $\alpha$ of the time with the scheme achieving $S_{NoCoop}^{(F)}$ over the remaining fraction of time. The time-sharing scheme leading to $S_{Partial}(\alpha)$ requires cooperation prelogs $\mu_{Tx} \geq \alpha\mu_{Tx,L}$ and $\mu_{Rx} \geq \alpha\mu_{Rx,L}$ and the time-sharing scheme leading to $S_{Interlaced}(\alpha)$ requires $\mu_{Tx} \geq \alpha\mu_{Tx,H}$ and $\mu_{Rx} \geq \alpha\mu_{Rx,H}$. The MG pair $S_{Partial-Inter}(\alpha)$ is achieved by time-sharing the scheme achieving $S_{Interlaced}$ over a fraction $\alpha$ of the time with the scheme achieving $S_{Partial}$ over the remaining fraction of time. This time-sharing schme requires cooperation prelogs $\mu_{Tx} \geq \alpha\mu_{Tx,H} + (1-\alpha)\mu_{Tx,L}$ and $\mu_{Rx} \geq \alpha\mu_{Rx,H} + (1-\alpha)\mu_{Rx,L}$.

Notice that for all of above time-sharing arguments, it is important that the MG pairs $S_{NoCoop}^{(F)}$, $S_{NoCoop}^{(S)}$, $S_{Coop}$, $S_{Partial}$, and $S_{Interlaced}$ can be achieved using the same values of $D_{Tx}$ and $D_{Rx}$. As will become clear in the following sections, these MG pairs can be achieved using any values $D_{Tx}, D_{Rx} \in \{1, 3, 5, \ldots D-1\}$ summing to D. The required cooperation prelogs however depend on the specific

choices of $D_{Tx}$ and $D_{Rx}$. This explains why the allowed time-sharing coefficients $\alpha$ depend on the number of cooperation rounds $D_{Tx}$ and $D_{Rx}$. $\quad\square$

**Remark 2.** *If in Theorem 3 we allow the parameters $D_{Tx}, D_{Rx}$ to take on any values in $\{1, 2, \ldots, D-1\}$ summing to $D$ and we remove the MG points $S_{Interlaced}$, $S_{Interlaced}(\beta_1^\star)$, $S_{Interlaced}(\beta_2^\star)$, and $S_{Partial-Inter}(\alpha_1^\star)$, we obtain a different achievable region, which can be larger for certain system parameters.*

*To see that this modified region is also achievable, notice that our schemes achieving $S_{Coop}$ and $S_{Partial}$ described in Sections 4.2 and 4.3 can be run with any number of Tx- and Rx- cooperation rounds $D_{Tx}$ and $D_{Rx}$, irrespective of whether they are odd or even. Their performance remains unchanged. In contrast, the scheme achieving $S_{Interlaced}$ that we present in Section 4.4 requires that both $D_{Tx}$ and $D_{Rx}$ are both odd.*

*In Figure 7 we schematically illustrate the MG regions that are achieved for $D_{Tx}$ or $D_{Rx}$ even. Specifically, Figure 7a shows the MG region for large cooperation prelogs and Figure 7b for small cooperation prelogs.*

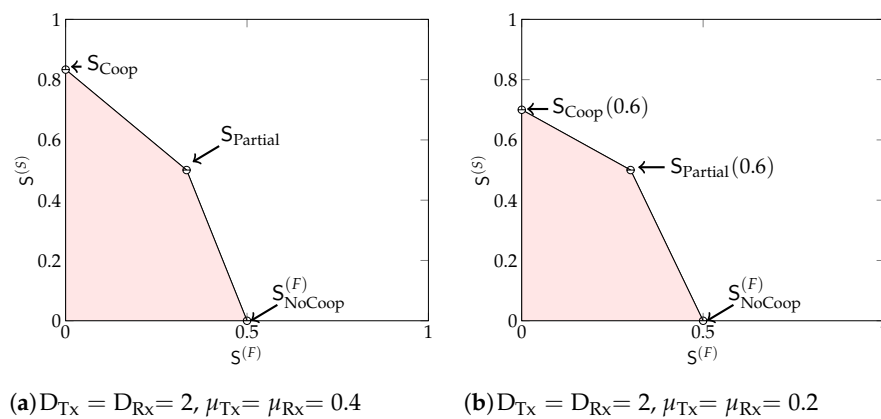We also have the following converse result.



(a) $D_{Tx} = D_{Rx} = 2$, $\mu_{Tx} = \mu_{Rx} = 0.4$

(b) $D_{Tx} = D_{Rx} = 2$, $\mu_{Tx} = \mu_{Rx} = 0.2$

**Figure 7.** Examples of the MG regions discussed in Remark 2 for even values of $D_{Tx}$ and $D_{Rx}$.

**Proposition 1** (Outer Bound on Optimal MG Region: Both Tx- and Rx-cooperation). *Any MG pair $(S^{(F)}, S^{(S)})$ in $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ satisfies*

$$S^{(F)} \leq \frac{L}{2}, \tag{40a}$$

$$S^{(F)} + S^{(S)} \leq \min\left\{\frac{L}{2} + \mu_{Tx} + \mu_{Rx}, \; L \cdot \frac{2D+1}{2D+2}\right\}. \tag{40b}$$

**Proof.** Follows from the converse result in Reference [12] and by a rate-transfer argument from "fast" to "slow" messages. $\quad\square$

Figure 8 depicts our inner and outer bounds (Theorem 3, Remark 2, and Proposition 1) on the optimal MG region with $\mu_{Tx} = \mu_{Rx} = 0.45$ and $D = 10$ for different values of $D_{Tx}$ and $D_{Rx}$. For $D_{Rx} = D/2 = 5$ and $D_{Tx} = D/2 = 5$,

$$\mu_{Tx} \geq \mu_{Tx,H} \quad \text{and} \quad \mu_{Rx} \geq \mu_{Rx,H}, \tag{41}$$

and the inner bound is given by the trapezoidal region defined in (33). It coincides with the outer bound, and thus establishes the exact MG region. Notice that in this case, the MG region is solely constrained by the fact that the MG of "fast" messages cannot exceed $\frac{L}{2}$ and that the sum MG of all messages cannot exceed $L \cdot \frac{2D_{Rx}+2D_{Tx}+1}{2D_{Rx}+2D_{Tx}+2}$. Imposing a stringent constraint on the decoding delay of the "fast" messages in this case never penalises the sum-MG of the system. Our inner bounds obtained for odd-valued cooperation-rounds $(D_{Tx}, D_{Rx}) \in \{(1, 9), (3, 7), (7, 3), (9, 1)\}$ coincide with the outer bound only if $S^{(F)} \leq L \cdot \frac{2(1-\alpha_1^\star)+\alpha_1^\star \cdot (D+1)}{2D+2}$, where $\alpha_1^\star$ depends on the choice of $(D_{Tx}, D_{Rx})$ and is defined

in (35). The inner bounds for even-valued cooperation-rounds $(D_{Tx}, D_{Rx}) \in \{(2,8), (4,6), (6,4), (8,2)\}$ all coincide and attain the outer bound only if $S^{(F)} \leq L \cdot \frac{2}{2D+2}$.



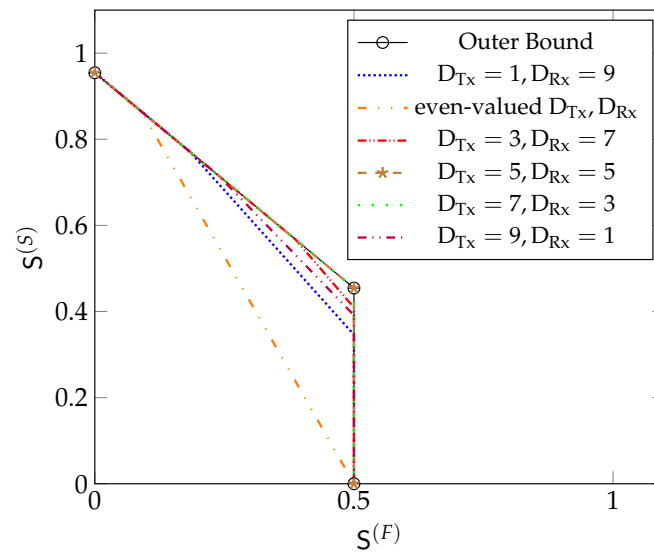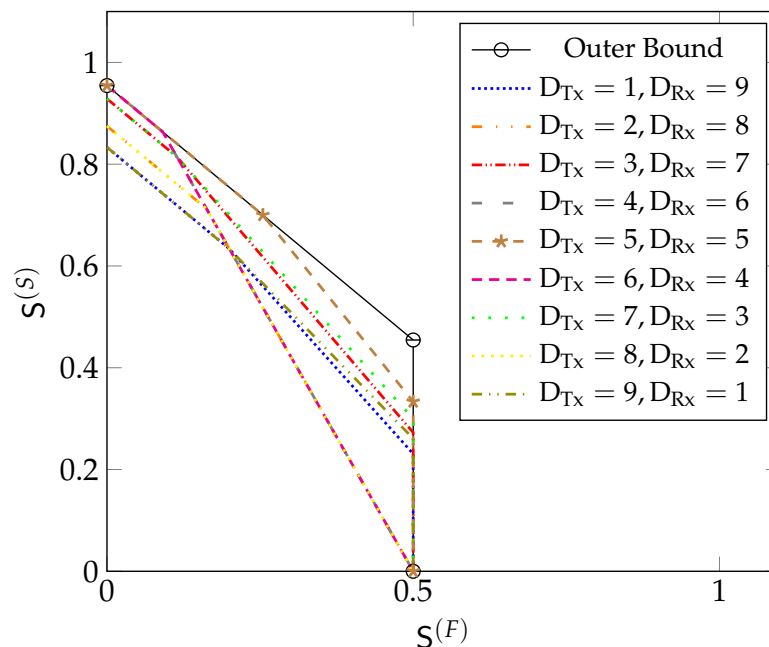**Figure 8.** Bounds on $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ for $\mu_{Tx} = 0.45$, $\mu_{Rx} = 0.45$, $D = 10$ and $L = 1$ with both Tx- and Rx-cooperation.

Figure 9 depicts our inner and outer bounds on the optimal MG region for the same $D = 10$ but smaller values of $\mu_{Tx} = \mu_{Rx} = 0.3$. In Figure 9, we see that by decreasing $\mu_{Tx}$ and $\mu_{Rx}$ from 0.45 to 0.3, our inner and outer bounds do not coincide for all values of $S^{(F)}$. Our inner bound for $D_{Rx} = 5$ and $D_{Tx} = 5$ contains all other inner bounds and it matches the outer bound in the regime $S^{(F)} \leq L \cdot \frac{2(1 - \alpha_1^\star) + \alpha_1^\star \cdot (D+1)}{2D+2}$, where for the definition of $\alpha_1^\star$ in (35) one should set $D_{Tx} = D_{Rx} = 5$.



**Figure 9.** Bounds on $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ for $\mu_{Tx} = 0.3$, $\mu_{Rx} = 0.3$, $D = 10$ and $L = 1$ with both Tx- and Rx-cooperation.

Figure 10 depicts our inner and outer bounds on the optimal MG region for the same $D = 10$ but when $\mu_{Tx} = 0.3$ is smaller than $\mu_{Rx} = 0.45$. Here, the inner bound obtained for $(D_{Tx} = 3, D_{Rx} = 7)$

includes all other inner bounds and it matches the outer bound in the regime $\mathsf{S}^{(F)} \leq \mathsf{L} \cdot \frac{2(1-\alpha_1^\star)+\alpha_1^\star \cdot (D+1)}{2D+2}$, where $\alpha_1^\star$ is defined in (35) with $D_{Tx} = 3$ and $D_{Rx} = 7$.
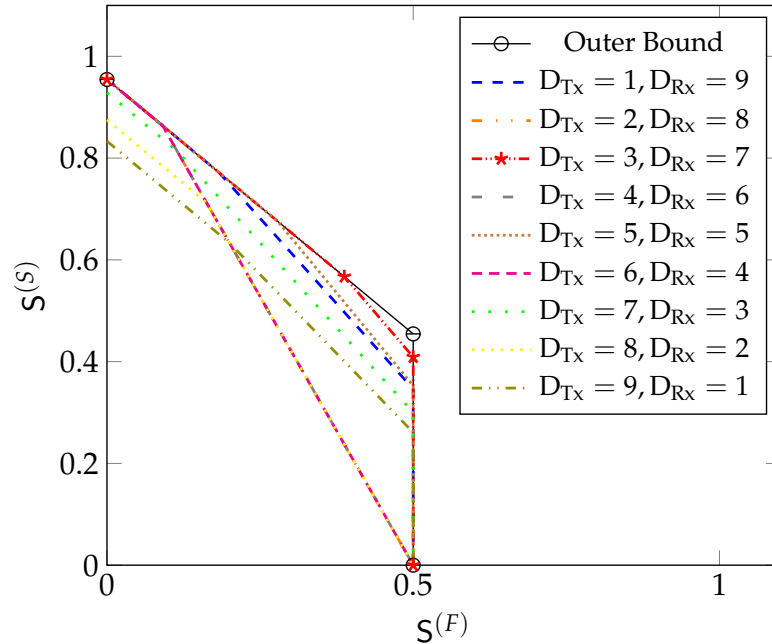
The following corollaries generalise these observations.



**Figure 10.** Bounds on $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ for $\mu_{Tx} = 0.3$, $\mu_{Rx} = 0.45$, $D = 10$ and $\mathsf{L} = 1$ with both Tx- and Rx-cooperation.

**Corollary 1.** *If there exist integers* $D_{Tx}, D_{Rx} \in \{1, 3, 5, \ldots, D-1\}$ *summing to D such that the two constraints*

$$\mu_{Tx} \geq \mu_{Tx,H}(D_{Tx}) \tag{42a}$$

$$\mu_{Rx} \geq \mu_{Rx,H}(D_{Rx}) \tag{42b}$$

*are simultaneously satisfied, then the optimal MG region* $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ *coincides with the trapezoidal region in* (33). *That means,* $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ *is the set of all nonnegative pairs* $(\mathsf{S}^{(F)}, \mathsf{S}^{(S)})$ *satisfying*

$$\mathsf{S}^{(F)} \leq \frac{\mathsf{L}}{2}, \tag{43}$$

$$\mathsf{S}^{(F)} + \mathsf{S}^{(S)} \leq \mathsf{L} \cdot \frac{2D+1}{2D+2}. \tag{44}$$

**Proof.** Follows directly from the achievability result in Theorem 3, see (33), and the converse result in Proposition 1. For the converse result notice in particular that under constraints (42) the sum $\mu_{Tx} + \mu_{Rx}$ exceeds $\mathsf{L} \cdot \frac{D}{2D+2}$. $\square$

**Remark 3.** *Under conditions* (42) *there is no penalty in sum-MG due to the stringent decoding constraint on "fast" messages. These "fast" messages can be submitted at maximum MG without decreasing the overall performance of the system.*

The following corollaries present partial characterizations of the optimal MG region $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ for $\mathsf{S}^{(F)}$ below a certain threshold.

**Corollary 2.** *If a pair of integers $D_{Tx}, D_{Rx} \in \{1, 3, \ldots, D-1\}$ summing to D satisfies*

$$\mu_{Tx} \geq \mu_{Tx,L}(D_{Tx}), \tag{45a}$$

$$\mu_{Rx} \geq \mu_{Rx,L}(D_{Rx}), \tag{45b}$$

*then the optimal MG region $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ contains the MG pair $(\mathsf{S}^{(F)}, \mathsf{S}^{(S)})$ with*

$$\mathsf{S}^{(F)} \leq \frac{\mathsf{L}}{2}\left(1 - \frac{D-1}{D+1}(1 - \alpha_1^\star)\right) \tag{46}$$

*(where $\alpha_1^\star$ is defined in (35) and depends on the choice of $D_{Tx}, D_{Rx}$ and on $\mu_{Tx}, \mu_{Rx}$) if, and only if,*

$$\mathsf{S}^{(F)} + \mathsf{S}^{(S)} \leq \mathsf{L} \cdot \frac{2D+1}{2D+2}. \tag{47}$$

*Similarly, if a pair of integers $D_{Tx}, D_{Rx} \in \{2, 4, \ldots, D-2\}$ summing to D satisfies (45), then the optimal MG region $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ contains the MG pair $(\mathsf{S}^{(F)}, \mathsf{S}^{(S)})$ with*

$$\mathsf{S}^{(F)} \leq \frac{\mathsf{L}}{2} \cdot \frac{2}{D+1} \tag{48}$$

*if, and only if,*

$$\mathsf{S}^{(F)} + \mathsf{S}^{(S)} \leq \mathsf{L} \cdot \frac{2D+1}{2D+2}. \tag{49}$$

Noting the fundamental bound $\mathsf{S}^{(F)} \leq \frac{\mathsf{L}}{2}$, one observes that when there is a odd-valued pair $D_{Tx}, D_{Rx} \in \{1, 3, \ldots, D-1\}$ such that (45) holds and $\alpha_1^\star = 1$, then the first part of Corollary 2 recovers Corollary 1 and determines the entire optimal MG region $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$.

**Proof.** Achievability of (47) follows from Theorem 3, see (34), because the two components of $\mathsf{S}_{\text{Inter-Partial}}(\alpha_1^\star) = (\mathsf{S}^{(S)}, \mathsf{S}^{(F)})$ satisfy:

$$\mathsf{S}^{(F)} = \alpha_1^\star \frac{\mathsf{L}}{2} + (1 - \alpha_1^\star)\frac{\mathsf{L}}{2}\frac{2}{D+1} = \frac{\mathsf{L}}{2}\left(1 - \frac{D-1}{D+1}(1 - \alpha_1^\star)\right) \tag{50}$$

and

$$\mathsf{S}^{(S)} + \mathsf{S}^{(F)} = \mathsf{L} \cdot \frac{2D+1}{2D+2}. \tag{51}$$

Achievability of (49) can be proved in a similar way from Remark 2. The converse to both results follows from Proposition 1 because constraint (47) implies that the sum $\mu_{Tx} + \mu_{Rx}$ exceeds $\mathsf{L} \cdot \frac{D}{2D+2}$. □

**Corollary 3.** *If*

$$\mu_{Rx} + \mu_{Tx} < \mathsf{L} \cdot \frac{D}{2D+2} \tag{52}$$

*and if a pair $D_{Tx}, D_{Rx} \in \{1, 2, 3, \ldots, D-1\}$ (both odd and even values are allowed) summing to D satisfies*

$$\frac{\mu_{Tx}}{D_{Tx}} = \frac{\mu_{Rx}}{D_{Rx}}, \tag{53}$$

*then the optimal MG region $\mathcal{S}^\star(\mu_{Tx}, \mu_{Rx}, D)$ contains the MG pair $(\mathsf{S}^{(F)}, \mathsf{S}^{(S)})$ with*

$$\mathsf{S}^{(F)} \leq \frac{\mathsf{L}}{2}\left(1 - \mu_{Tx}\frac{2(D-1)}{D_{Tx}}\right) \tag{54}$$

*if, and only if,*

$$S^{(F)} + S^{(S)} \leq \frac{L}{2} + \mu_{Tx} + \mu_{Rx}. \tag{55}$$

**Proof.** The result (55) follows from the converse result in Proposition 1 and the achievability results in Theorem 3, see (37), and Remark 2. More specifically, to prove achievability let $D_{Tx}$ and $D_{Rx}$ be such that Condition (53) is satisfied. Then,

$$\frac{\mu_{Tx,L}(D_{Tx})}{\mu_{Tx}} = \frac{\mu_{Rx,L}(D_{Rx})}{\mu_{Rx}} \tag{56}$$

and Condition (52) implies that both inequalities

$$\mu_{Tx} < \mu_{Tx,L} \qquad \text{and} \qquad \mu_{Rx} < \mu_{Rx,L} \tag{57}$$

are satisfied. Moreover, $\alpha_2^\star$ as defined in (38) satisfies

$$\alpha_2^\star = \frac{\mu_{Tx}}{\mu_{Tx,L}} = \mu_{Tx} \frac{2D+2}{L \cdot D_{Tx}} = \frac{\mu_{Rx}}{\mu_{Rx,L}} = \mu_{Rx} \frac{2D+2}{L \cdot D_{Rx}}. \tag{58}$$

Notice next that for the two MG pairs $S_{Coop}(\alpha_2^\star)$ and $S_{Partial}(\alpha_2^\star)$, which are achievable by either Theorem 3 or Remark 2, the sum of the two components satisfies

$$S^{(F)} + S^{(S)} = \alpha_2^\star \cdot \frac{L}{2} \cdot \frac{2D+1}{D+1} + (1 - \alpha_2^\star) \frac{L}{2} = \frac{L}{2} + \frac{\mu_{Tx}}{D_{Tx}} D = \frac{L}{2} + \mu_{Tx} + \mu_{Rx}, \tag{59}$$

where in the last equation we used (53). Moreover, the "fast" MG $S^{(F)}$ in $S_{Coop}(\alpha_2^\star)$ equals 0, whereas in $S_{Partial}(\alpha_2^\star)$ it equals

$$S^{(F)} = \alpha_2^\star \cdot \frac{L}{2} \cdot \frac{2}{D+1} + (1 - \alpha_2^\star) \frac{L}{2} = \left(1 - \mu_{Tx} \frac{2(D-1)}{D_{Tx}}\right) = \left(1 - \mu_{Rx} \frac{2(D-1)}{D_{Rx}}\right). \tag{60}$$

Since one can always choose to transmit at smaller MGs and because the convex hull of all achievable MG pairs is also achievable, this concludes the proof of achievability. □

**Remark 4.** *For both corollaries, in the regimes where we could characterize the optimal MG region, i.e., for "fast" MGs below a certain threshold, the sum-MG is at its maximum. We can thus conclude that for sufficiently small $S^{(F)}$ the sum-MG is not decreased due to the stringent constraint on the "fast" messages.*

In the following subsections we present the coding schemes achieving the MG regions in Theorem 3 and Remark 2.

*4.2. Scheme Achieving (31c)*

Let each Tx only send "slow" messages but no "fast" messages. Under this coding assumption, our setup is a multi-antenna version of the setup in [12]. Achievability of (31c) then follows immediately by the multi-antenna version of [12] (Theorem 1). We redescribe the coding schemes achieving (31c) for completeness and reference in the next subsection.

We silence every $2D + 2$nd Tx, which splits the network into smaller subnets. In each subnet, we combine the SIC idea explained for the setup with only Rx-cooperation (see Section 3.1.1) with the DPC coding idea that was explained for the setup with only Tx-cooperation (see Section 3.2.1). The scheme for the first subnet is illustrated in Figure 11 and will be explained in the following. Communication in the other subnets is similar.

The Tx/Rx pairs of the first subnet are assigned to four groups, depending on their mode of operation. Notice that the Tx/Rx pair $D_{Rx} + 2D_{Tx} + 2$ is assigned to both groups $\mathcal{G}_3$ and $\mathcal{G}_4$, whereas

all other Tx/Rx pairs are assigned to only one group. The reason is that message $M^{(S)}_{\text{D}_{\text{Rx}}+2\text{D}_{\text{Tx}}+2}$ is split into two parts $(M^{(S,3)}_{\text{D}_{\text{Rx}}+2\text{D}_{\text{Tx}}+2}, M^{(S,4)}_{\text{D}_{\text{Rx}}+2\text{D}_{\text{Tx}}+2})$ of equal rates, and part $M^{(S,3)}_{\text{D}_{\text{Rx}}+2\text{D}_{\text{Tx}}+2}$ is communicated in the same way as the messages for Tx/Rx pairs in group $\mathcal{G}_3$, whereas $M^{(S,4)}_{\text{D}_{\text{Rx}}+2\text{D}_{\text{Tx}}+2}$ is communicated in the same way as the messages for Tx/Rx pairs in group $\mathcal{G}_4$.
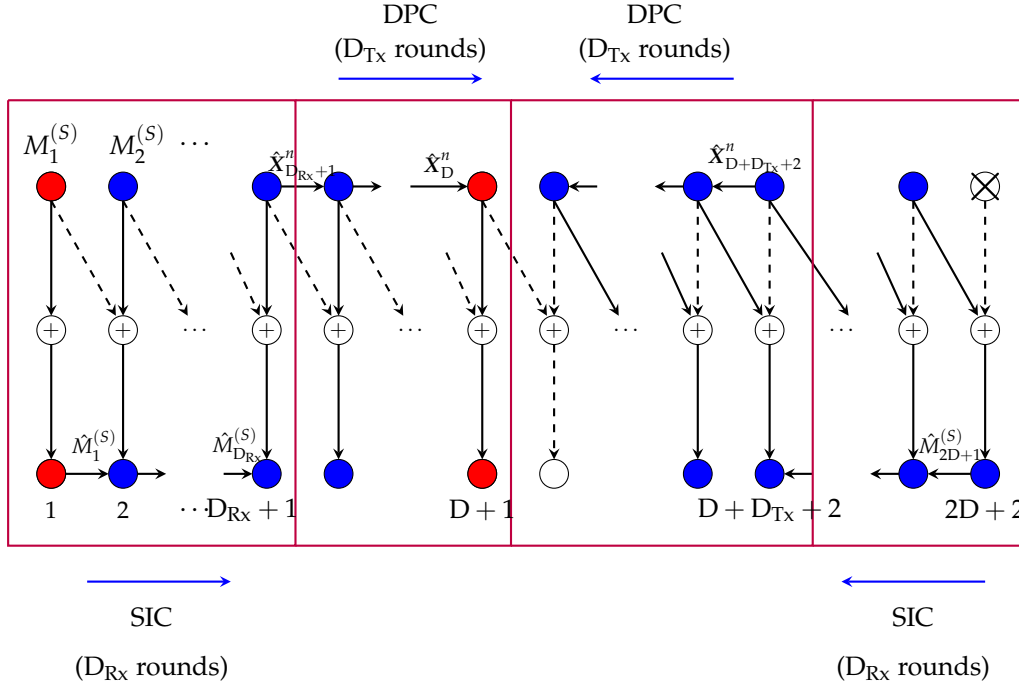


**Figure 11.** Scheme with Rx- and Tx-cooperation.

*Group $\mathcal{G}_1 \triangleq \{1, \ldots, D_{Rx} + 1\}$:* Each Tx $k \in \mathcal{G}_1$ encodes its "slow" message $M^{(S)}_k$ using a codeword $X^n_k(M^{(S)}_k)$ from a Gaussian point-to-point code of power P, and transmits this codeword over the channel: $X^n_k = X^n_k(M^{(S)}_k)$. Each Rx $k \in \mathcal{G}_1$ uses the cooperation message received from its left neighbour Rx $k - 1$ for SIC, i.e., to delete the interference term $\mathsf{H}_{k-1,k} X^n_{k-1}(\hat{M}^{(S)}_{k-1})$ from its output sequence $Y^n_k$:

$$\hat{Y}^n_k = Y^n_k - \mathsf{H}_{k-1,k} X^n_{k-1}(\hat{M}^{(S)}_{k-1}), \tag{61}$$

and to decode its desired message $M^{(S)}_k$ based on $\hat{Y}^n_k$. Rx $k$ also describes its decoded message $\hat{M}_k$ over the cooperation link to Rx $k + 1$, so as to facilitate SIC at this next Rx.

To facilitate the transmissions in the next group, the last Tx of group $\mathcal{G}_1$, Tx $D_{Rx} + 1$, precodes its channel inputs $X^n_{\text{D}_{Rx}+1}$ with the matrix $\mathsf{H}^{-1}_{\text{D}_{Rx}+2,\text{D}_{Rx}+2}\mathsf{H}_{\text{D}_{Rx}+1,\text{D}_{Rx}+2}$, quantises the produced sequence $I^n_{\text{D}_{Rx}+1} \triangleq \mathsf{H}^{-1}_{\text{D}_{Rx}+2,\text{D}_{Rx}+2}\mathsf{H}_{\text{D}_{Rx}+1,\text{D}_{Rx}+2}X^n_{\text{D}_{Rx}+1}$ with a rate-$\text{L}/2\log(1 + \text{P})$ quantiser to obtain the quantisation $\hat{I}^n_{\text{D}_{Rx}+1}$ at noise level and sends the resulting quantisation index as a first-round cooperation message to the first Tx in group $\mathcal{G}_2$, i.e. to Tx $D_{Rx} + 2$.

*Group $\mathcal{G}_2 \triangleq \{D_{Rx} + 2, \ldots, D_{Rx} + D_{Tx} + 1\}$:* Each Tx $k \in \mathcal{G}_2$ obtains a cooperation message from its left neighbour Tx $k - 1$ that describes the quantised version $\hat{I}^n_{k-1}$ of $I^n_{k-1} \triangleq \mathsf{H}^{-1}_{k,k}\mathsf{H}_{k-1,k}X^n_{k-1}$. Based on this message, Tx $k$ reconstructs $\hat{I}^n_{k-1}$, encodes its "slow" message $M^{(S)}_k$ using a power P DPC that mitigates the interference $\hat{I}^n_{k-1}$, and sends the resulting DPC sequence $X^n_k$ over the channel. Moreover, it precodes this input sequence with the matrix $\mathsf{H}^{-1}_{k+1,k+1}\mathsf{H}_{k,k+1}$, quantises the precoded sequence $I^n_k \triangleq \mathsf{H}^{-1}_{k+1,k+1}\mathsf{H}_{k,k+1}X^n_k$ with a rate-$\text{L}/2\log(1 + \text{P})$ quantiser (for a quantisation at noise level) to obtain $\hat{I}^n_k$, and sends the quantisation message as a round-$k$ cooperation message over the link to its right neighbour. Tx $D_{Tx} + D_{Rx} + 1$ produces its inputs in a similar way, i.e., using DPC,

but sends no cooperation message at all. Rxs in $\mathcal{G}_2$ use a standard DPC decoding rule based on the premultiplied outputs

$$\mathsf{H}_{k,k}^{-1}Y_k^n = \mathsf{H}_{k,k}^{-1}\mathsf{H}_{k-1,k}X_{k-1}^n + X_k^n + \mathsf{H}_{k,k}^{-1}Z_k^n, \tag{62}$$

to decode their intended "slow" messages. (Recall that $X_k^n$ was produced as a DPC sequence that mitigates $\hat{I}_{k-1}^n$, a quantised version of $\mathsf{H}_{k,k}^{-1}\mathsf{H}_{k-1,k}\hat{X}_{k-1}^n$). Since quantisation was performed at noise level, each message $M_{\mathrm{D_{Rx}}+2}^{(S)}, \ldots, M_{\mathrm{D}+1}^{(S)}$ can be sent reliably with MG L.

*Group* $\mathcal{G}_3 \triangleq \{D_{Rx} + D_{Tx} + 2, \ldots, D_{Rx} + 2D_{Tx} + 2\}$: This group of Tx/Rx pairs participates in the transmission of the "slow" messages

$$M_{\mathrm{D_{Rx}}+D_{Tx}+3}^{(S)}, \ldots, M_{\mathrm{D_{Rx}}+2D_{Tx}+1}^{(S)}, M_{\mathrm{D_{Rx}}+2D_{Tx}+2}^{(S,3)}. \tag{63}$$

In particular, Tx $D_{Rx} + 2D_{Tx} + 2$ does not send an own message to its corresponding Rx.

Each of the messages in (63) is transmitted over the communication path Tx $k \to$ Tx $k-1 \to$ Rx $k$ for some $k \in \{D_{Rx} + D_{Tx} + 3, \ldots, D_{Rx} + 2D_{Tx} + 2\}$.

For each $k \in \{D_{Rx} + D_{Tx} + 3, \ldots, D_{Rx} + 2D_{Tx} + 2\}$, Tx $k$ encodes its own "slow" message $M_k^{(S)}$ by means of DPC of power P that mitigates the interference $\mathsf{H}_{k-1,k}^{-1}\mathsf{H}_{k,k}X_k^n$ of the signal sent by Tx $k$ itself; precodes the obtained sequence $U_k^n$ with the matrix $\mathsf{H}_{k-1,k}^{-1}\mathsf{H}_{k,k}$; quantises the precoded sequence $S_{k-1}^n \triangleq \mathsf{H}_{k-1,k}^{-1}\mathsf{H}_{k,k}U_k^n$ to obtain a quantisation $\hat{S}_{k-1}$ at noise level; and sends the corresponding quantisation message as a $(2D + 3 - k)$-round cooperation message over the link to Tx $k - 1$. Tx $k - 1$ then reconstructs $\hat{S}_{k-1}^n$ and sends it over the channel: $X_{k-1}^n = \hat{S}_{k-1}^n$. The construction of the transmit signal $X_{\mathrm{D_{Rx}}+2D_{Tx}+2}^n$ mentioned above, is explained in the following paragraph. RXs $D_{Rx} + D_{Tx} + 3, \ldots, D_{Rx} + 2D_{Tx} + 2$ decode their intended "slow" messages using an optimal DPC decoding rule based on the premultiplied outputs

$$\mathsf{H}_{k-1,k}^{-1}Y_k^n = X_{k-1}^n + \mathsf{H}_{k-1,k}^{-1}\mathsf{H}_{k,k}X_k^n + \mathsf{H}_{k-1,k}^{-1}Z_k^n. \tag{64}$$

Recall that $X_{k-1}^n$ is a quantised version (at noise level) of the precoded signal $S_{k-1}^n \triangleq \mathsf{H}_{k-1,k}^{-1}\mathsf{H}_{k,k}U_k^n$ for $U_k^n$ a DPC sequence that mitigates the interference $\mathsf{H}_{k-1,k}^{-1}\mathsf{H}_{k,k}X_k^n$. Each of the messages $M_{\mathrm{D}+3}^{(S)}, \ldots, M_{2D+2}^{(S)}$ can thus be transmitted reliably at full MG L.

*Group* $\mathcal{G}_4 \triangleq \{D_{Rx} + 2D_{Tx} + 2, \ldots, 2D_{Rx} + 2D_{Tx} + 2\}$: This group of Tx/Rx pairs participates in the transmission of the "slow" messages

$$M_{\mathrm{D_{Rx}}+2D_{Tx}+2}^{(S,4)}, M_{\mathrm{D_{Rx}}+2D_{Tx}+3}^{(S)}, \ldots, M_{2D_{Rx}+2D_{Tx}+1}^{(S)}. \tag{65}$$

Tx $2D_{Rx} + 2D_{Tx} + 2$ thus is not sending an own message to its corresponding Rx. The messages in (65) are transmitted over the path Tx $k \to$ Rx $k+1 \to$ Rx $k$, for some $k \in \{D_{Rx} + 2D_{Tx} + 2, \ldots, 2D_{Rx} + 2D_{Tx} + 2\}$.

Each Tx $k \in \{D_{Rx} + 2D_{Tx} + 2, \ldots, 2D_{Rx} + 2D_{Tx} + 1\}$ encodes its "slow" message $M_k^{(S)}$ (or $M_k^{(S,4)}$ if $k = D_{Rx} + 2D_{Tx} + 2$) using a codeword from a Gaussian codebook of power P, and sends this codeword over the channel $X_k^n = X_k^n(M_k^{(S)})$ (or $X_k^n = X_k^n(M_k^{(S,4)})$ if $k = D_{Rx} + 2D_{Tx} + 1$).

Rx $2D_{Rx} + 2D_{Tx} + 2$ decodes $M_{2D_{Rx}+2D_{Tx}+1}^{(S)}$ based on an interference-free output $Y_{2D_{Rx}+2D_{Tx}+2}^n = \mathsf{H}_{2D_{Rx}+2D_{Tx}+1,2D_{Rx}+2D_{Tx}+2}X_{2D_{Rx}+2D_{Tx}+1}^n + Z_{2D_{Rx}+2D_{Tx}+2}^n$, and sends the decoded message $\hat{M}_{2D_{Rx}+2D_{Tx}+1}^{(S)}$ over the cooperation link to the intended Rx $2D_{Rx} + 2D_{Tx} + 1$. For $k = 2D_{Rx} + 2D_{Tx} + 1, \ldots, D_{Rx} + 2D_{Tx} + 3$, Rx $k$ uses the cooperation message received from its right neighbour Rx $k+1$ to decode $M_{k-1}^{(S)}$ (or $M_{k-1}^{(S,4)}$ if $k = D_{Rx} + 2D_{Tx} + 2$) using SIC, i.e., to first delete the interference $\mathsf{H}_{k,k}X_k^n$ from $Y_k^n$ and then decode message $M_{k-1}^{(S)}$ (or $M_{k-1}^{(S,4)}$ if $k = D_{Rx} + 2D_{Tx} + 2$) from an interference-free signal. Rx $k$ then sends the decoded message $\hat{M}_{k-1}^{(S)}$ (or $\hat{M}_{k-1}^{(S,4)}$ if $k = D_{Rx} + 2D_{Tx} + 2$) over the cooperation link to its left neighbour Rx $k - 1$, which is the intended Rx for this message.

In the described scheme, each transmitted message is either decoded based on interference-free outputs or using DPC. Since precoding matrices do not depend on the power and quantizations are performed at noise levels, all messages can be transmitted reliably at MG L. Tx $D_{Rx} + 2D_{Tx} + 2$ sends two "slow" messages and $2D_{Rx} + 2D_{Tx} - 1$ other Txs send one "slow" message. An average "slow" MG of $L \cdot \frac{2D_{Rx} + 2D_{Tx} + 1}{2D_{Rx} + 2D_{Tx} + 2}$ is thus achieved in each subnet. Moreover, $2D_{Rx} + 2D_{Tx}$ cooperation messages of prelog L are sent in each subnet:

1.  Rxs in $\mathcal{G}_1$ send $D_{Rx}$ Rx-cooperation messages with prelog L;
2.  Txs in $\mathcal{G}_2$ send $D_{Tx}$ Tx-cooperation messages with prelog L;
3.  Txs in $\mathcal{G}_3$ send $D_{Tx}$ Tx-cooperation messages with prelog L;
4.  Rxs in $\mathcal{G}_4$ send $D_{Rx}$ Rx-cooperation messages with prelog L.

The *average* cooperation prelog *per link* at the Tx-side is $\mu_{Tx,L}$ and at the Rx-side it is $\mu_{Rx,L}$. If one time-shares $2D + 2$ different instances of the described scheme with a different subset of silenced users in each of them, the overall scheme still achieves the MG pair $(S^{(F)} = 0, S^{(S)} = L\frac{2D+1}{2D+2})$ in (31d), each Tx-cooperation link is loaded at exactly this average cooperation prelog $\mu_{Tx,L}$, and each Rx-cooperation link is loaded at the average cooperation prelog $\mu_{Rx,L}$.

### 4.3. Scheme Achieving MG Pair (31d)

Consider the scheme described in the previous Section 4.2 and depicted in Figure 11. Notice that the first Tx in each subnet does not at all participate in the cooperation, and decoding of its message also does not rely on cooperation messages. The same observation applies also to the $D_{Rx} + D_{Tx} + 1$st Tx of each subnet and its message. The first and the $D_{Rx} + D_{Tx} + 1$st message of each subnet (the red Txs in Figure 11) thus satisfy the requirements on "fast" messages. We propose to use this scheme but let the first and the $(D_{Rx} + D_{Tx} + 1)$st messages in each subnet be "fast" messages and all other messages be "slow" messages. This achieves the MG pair (31d).

The required cooperation rates equal $\mu_{Tx,L}$ and $\mu_{Rx,L}$, as explained in the previous Section 4.2.

### 4.4. Schemes Achieving MG Pair (31e)

We periodically silence every $2D + 2$-nd Tx to split the network into smaller subnets. Then we send a "fast" message on all odd Txs and a "slow" message on all even Txs, except for the previously silenced Txs (which are all even). See Figure 12.

In what follows, we describe and analyze transmissions over the first subnet. Other subnets are treated analogously.

*Odd Txs $1, 3, 5, \ldots, 2D + 1$:* Each odd Tx encodes its "fast" message $M_k^{(F)}$ using a codeword $U_k^n(M_k^{(F)})$ from a Gaussian codebook of power P that depends on the Tx and the channel realizations and is explained later. Tx 1 simply sends this Gaussian codeword $X_1^n = U_1^n(M_1^{(F)})$. Any other odd Tx $k$ first considers the cooperation message it received from its left neighbour Tx $k - 1$ and reconstructs $\hat{X}_{k-1}^n$, a quantised version of Tx $k - 1$th input $X_{k-1}^n$. Tx $k$ then sends the input signal

$$X_k^n = U_k^n(M_k^{(F)}) - H_{k,k}^{-1}H_{k-1,k}\hat{X}_{k-1}^n. \tag{66}$$

Odd Txs relay some of the cooperation messages they obtain from their neighbours, as will become clear in the following, but they do not create new cooperation messages.
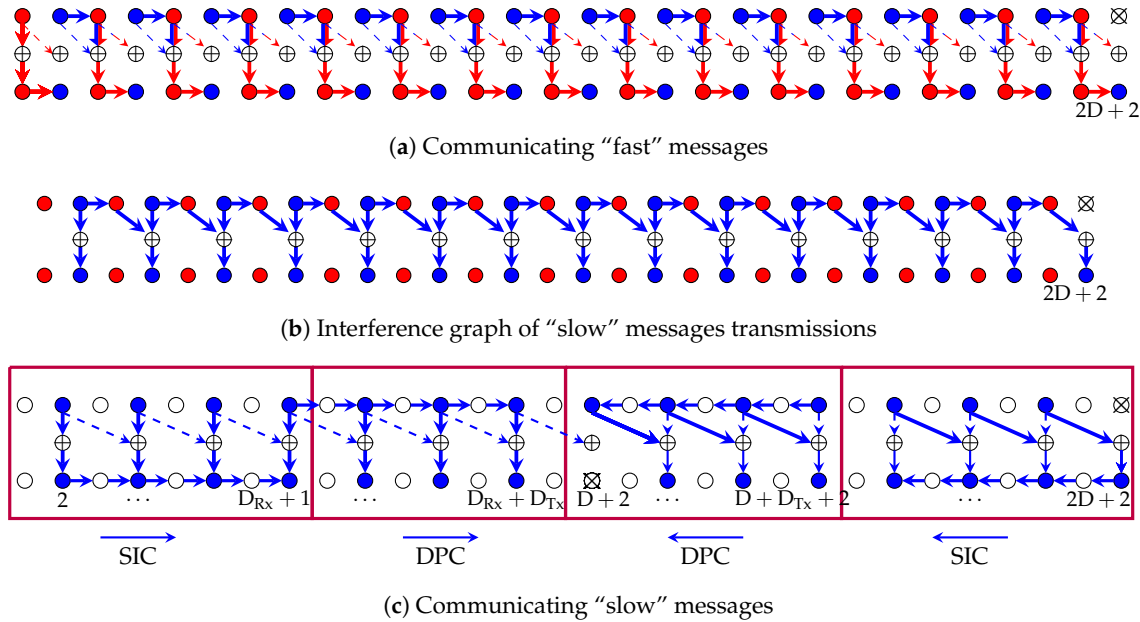
*Odd Rxs $1, 3, 5, \ldots, 2D + 1$:* Given the precanceling at odd Txs described above, each odd Rx $k$ observes an almost interference-free signal:

$$Y_k^n = H_{k,k}U_k^n + H_{k-1,k}(X_{k-1}^n - \hat{X}_{k-1}^n) + Z_{k,k}, \tag{67}$$

where notice that $\hat{X}_{k-1}^n$ is a quantised version of $X_{k-1}^n$ at noise level. Each odd Rx $k$ therefore decodes its desired fast message $M_k^{(F)}$ using standard point-to-point decoding. It also sends the decoded message $\hat{M}_k^{(F)}$ over the cooperation link to its right neighbour Rx $k+1$ as a first round cooperation message.

Odd Rxs also relay some of the cooperation messages they obtain from their neighbours, as will become clear in the following.



(**a**) Communicating "fast" messages



(**b**) Interference graph of "slow" messages transmissions



(**c**) Communicating "slow" messages

**Figure 12.** An illustration of the scheme achieving MG pair (31e). Notice that since D is even, the last Tx of $\mathcal{G}_2$ sends a "fast" message. And since $D_{Rx}$ is odd, also the first Tx in $\mathcal{G}_4$ sends a "fast" message.

Before describing the operations at the even Tx/Rx pairs, we make the following observations based on the operations at the odd Tx/Rx pairs. Irrespective of the operations performed at the even Txs, each even Rx $k$ observes the sum of a signal depending only on "slow" messages and a signal depending only on its left-neighbour's "fast" message (the signal $\mathsf{H}_{k-1,k}U_{k-1}^n$). Since odd Rxs convey their decoded "fast" messages to their right-neighbour, even Rxs can cancel the signals depending on "fast" messages whenever they have been decoded correctly. There is thus no loss in reliable communication rate caused by the transmission of "fast" messages. And transmission of "slow" messages at even Txs can be designed as if no "fast" messages were present. However, if "slow" Rxs wish to send cooperation messages that do not depend on the "fast" transmissions, they have to wait for the second round.

*Even Txs* $2, 4, 6, \ldots, 2D$: Each even Tx $k$, for $k = 2, \ldots, 2D$, performs the same steps as Tx $k$ in the scheme described in Section 4.2, but where the scheme needs to be adapted to include only even Txs. In particular, if an even Tx $k$ previously sent a quantisation message to its direct left- or right-neighbour Tx $k-1$ or $k+1$, now it will send it to the previous or following *even* Tx $k-2$ or Tx $k+2$. (This simply means that the odd Tx lying between them has to relay the cooperation message as we already mentioned previously.) Similarly, when using DPC, if Tx $k$ previously mitigated the quantised sequence $\hat{I}_{k-1}^n$ or $\hat{S}_{k+1}^n$, now it mitigates the quantised sequence $\hat{I}_{k-2}^n$ or $\hat{S}_{k+2}^n$. Notice that since D is even, Tx D is the last even Tx in $\mathcal{G}_2$ (so the last Tx in $\mathcal{G}_2$ sending a "slow" message). Tx-cooperation in group $\mathcal{G}_2$ thus takes place only during the first $D_{Tx} - 1$ rounds. The only Tx-cooperation message in round $D_{Tx}$ is the message sent from Tx $D+3$ to Tx $D+2$ in group $\mathcal{G}_3$.

In addition, if this is not already done as part of the scheme in Section 4.2, any even Tx $k$ also quantizes its channel inputs $X_k^n$ at rate $\mathsf{L} \cdot 1/2 \log(1 + \mathsf{P})$ to generate the quantised sequence $\hat{I}_k^n$. The quantisation message describing $\hat{I}_k^n$ is then sent as a $D_{Tx}$-round cooperation message over the link to Tx $k+1$ to allow this Tx to precancel this interference in the way that was described previously.

Even Tx $D + 2$ (the first Tx in group $\mathcal{G}_3$) does not need to send this round-$D_{Tx}$ cooperation message because its right neighbour Tx $D + 3$ already learns the Tx signal $X_{D+2}^n$ as part of the proposed scheme in Section 4.2. Since all even Txs (except for Tx $D + 2$) receive their last cooperation message in round $D_{Tx} - 1$, they can indeed compute their input perior to the last round $D_{Tx}$ and thus perform the proposed round-$D_{Tx}$ cooperation.

*Even Rxs* $2, 4, 6, \ldots, 2D + 2$: Using the round-1 Rx-cooperation messages from its left neighbour, each even Rx $k$, for $k = 2, \ldots, 2D + 2$, first subtracts the interference caused by the transmission of the "fast" message $M_{k-1}^{(F)}$ at its left neighbour. That means, it forms

$$\tilde{Y}_k^n = Y_k^n - H_{k-1,k} U_{k-1}^n. \tag{68}$$

It then proceeds with this modified output sequence $\tilde{Y}_k^n$ and performs all the steps as Rx $k$ did in the scheme in Section 4.2, but where the scheme again needs to be adapted to include only even Rxs and it also needs to be adapted to start only at cooperation round 2. This allows even Rxs to calculate (68) before performing the other steps. Notice that since the first Txs of $\mathcal{G}_1$ and $\mathcal{G}_4$ only send "fast" messages (the latter holds because $D_{Rx}$ is odd), there is no harm in waiting for this second round. To adapt the scheme in Section 4.2 only to even Rxs, any even Rx $k$ that previously sent its decoded message to its direct left- or right-neighbour Rx $k - 1$ or Rx $k + 1$, now sends it to the previous or following *even* Rx $k - 2$ or Rx $k + 2$. Similarly, any Rx $k$ that previously applied the SIC step to cancel the interference from Tx $k - 1$ or Tx $k + 1$, now cancels the interference from Tx $k - 2$ or Tx $k + 2$.

In the described scheme, all odd Txs of a subnet can send reliably a "fast" message of MG L and the even Txs $\{2, 4, \ldots, 2D\}$ each can send reliably a "slow" message of MG L. The scheme thus achieves the MG pair in (31e): $(S^{(F)} = \frac{L}{2}, S^{(S)} = L \cdot \frac{D}{2D+2})$.

We now analyze the cooperation prelog of the described scheme. Recall that in this scheme each even Tx sends a quantised version of its inputs to its right neighbour and each odd Rx sends its decoded message to its right neighbour. Since each of these cooperation messages is of prelog L the described messages consume a Tx-cooperation prelog of $L \cdot D$ and a Rx-cooperation prelog of $L \cdot D$.

*In addition*, for encoding and decoding of "slow" messages:

1. Rxs in $\mathcal{G}_1$ send $D_{Rx} - 1$ Rx-cooperation messages with prelog L. (The cooperation message from Rx 1 to Rx 2 has already been counted in the previous paragraph.)
2. Txs in $\mathcal{G}_2$ send $(D_{Tx} - 1)/2$ Tx-cooperation messages with prelog L. (The cooperation message from even to odd Txs in $\mathcal{G}_2$ have already been counted in the previous paragraph.)
3. Txs in $\mathcal{G}_3$ send $D_{Tx}$ Tx-cooperation messages with prelog L.
4. Rxs in $\mathcal{G}_4$ send $D_{Rx} - 1$ Rx-cooperation messages with prelog L. (The first Rx in $\mathcal{G}_4$ does not obtain a cooperation message because it is a "fast" Tx.)

To summarize, the described scheme requires an *average prelog per Tx-cooperation link* of $\mu_{Tx,H} = L \cdot \frac{\frac{D}{2} + \frac{3}{4}D_{Tx} - \frac{1}{4}}{2D+2}$ and an *average prelog per Rx-cooperation link* of $\mu_{Rx,H} = L \cdot \frac{\frac{D}{2} + D_{Rx} - 1}{2D+2}$. (Notice that this is larger than in the scheme in Section 4.2.) If one time-shares $2D + 2$ different instances of the described scheme with a different subset of silenced users in each of them, the required prelog on each Tx-cooperation link is exactly $\mu_{Tx,H}$ and the required prelog on each Rx-cooperation link is exactly $\mu_{Rx,H}$. This concludes the proof.

## 5. Summary and Concluding Remarks

We considered Wyner's soft-handoff network and characterized the MG region with transmitter and receiver cooperation when part of the messages are subject to stringent delay constraints. For the setup with *only* transmitter or *only* receiver cooperation we observed the following. Increasing the MG of delay-sensitive messages by $\Delta$ requires decreasing the MG of delay-tolerant messages approximately by $2\Delta$. This penalty does not arise when both transmitters *and* receivers can cooperate. More precisely, for small cooperation prelogs, when delay-sensitive messages have moderate or small MGs, then the

*sum*-MG is not decreased compared to when only delay-tolerant messages are transmitted. For large cooperation prelogs, this conclusion even holds when delay-sensitive messages have large MGs.

An interesting line of future work concerns extending the existing results to two-dimensional cellular models (i.e., to models where transmitters and receivers are not aligned on a grid). First results on the hexagonal Wyner model [18] indicate that similar conclusions hold as for Wyner's soft-handoff model investigated in this talk. Another interesting line of future work studies the impact of channel state information (CSI) at the transmitter as in Reference [19] but for the considered model with mixed-delay constraints. In particular a model where CSI is present for en/decoding delay-tolerant messages but not for en/decoding of delay-sensitive messages is a natural extension of the presented setup.

**Author Contributions:** Writing–original draft, H.N.; Writing–review and editing, M.A.W. and S.S. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Proof of the Converse to (15)

For convenience of notation, define for any $k \in \{1, \ldots, K\}$:

$$M_k \triangleq (M_k^{(F)}, M_k^{(S)}). \tag{A1}$$

For each power $P > 0$, fix a sequence (in the blocklength $n$) of encoding and decoding functions respecting the power constraints and the Rx-cooperation rate-limitations (recall that we consider a setup with only Rx-cooperation but no Tx-cooperation) such that the error probability $p(\text{error}) \to 0$ as the blocklength $n \to \infty$.

By Fano's Inequality there exists a sequence $\epsilon_n$ satisfying $\frac{\epsilon_n}{n} \to 0$ as $n \to \infty$ such that for any $k \in [1 : K - 1]$ and each blocklength $n$:

$$
\begin{aligned}
&R_k^{(F)} + R_k^{(S)} + R_{k+1}^{(F)} \\
&= \frac{1}{n}\Big[ H(M_k^{(F)}) + H(M_k^{(S)}) + H(M_{k+1}^{(F)}) \Big] \tag{A2} \\
&= \frac{1}{n}\Big[ H(M_k^{(F)}|M_{k-1}) \\
&\qquad + H\big(M_k^{(S)}|M_1, \ldots, M_{k-1}, M_k^{(F)}, M_{k+1}, \ldots, M_K\big) + H\big(M_{k+1}^{(F)}|M_{k-1}, M_{k+1}^{(S)}\big) \Big] \tag{A3} \\
&\leq \frac{1}{n}\Big[ I(M_k^{(F)}; Y_k^n|M_{k-1}) + I(M_k^{(S)}; Y_1^n, \ldots, Y_K^n|M_1, \ldots, M_{k-1}, M_k^{(F)}, M_{k+1}, \ldots, M_K) \\
&\qquad + I(M_{k+1}^{(F)}; Y_{k+1}^n|M_{k-1}, M_{k+1}^{(S)}) \Big] + \frac{\epsilon_n}{n} \tag{A4} \\
&\overset{(a)}{=} \frac{1}{n}\Big[ I(M_k^{(F)}; Y_k^n|M_{k-1}) + I(M_k^{(S)}; Y_k^n, Y_{k+1}^n|M_{k-1}, M_k^{(F)}, M_{k+1}) \\
&\qquad + I(M_{k+1}^{(F)}; Y_{k+1}^n|M_{k-1}, M_{k+1}^{(S)}) \Big] + \frac{\epsilon_n}{n} \tag{A5} \\
&\overset{(b)}{=} \frac{1}{n}\Big[ I(M_k^{(F)}, M_k^{(S)}; Y_k^n|M_{k-1}) + I(M_k^{(S)}; Y_{k+1}^n|Y_k^n, M_k^{(F)}, M_{k-1}, M_{k+1}) \\
&\qquad + I(M_{k+1}^{(F)}; Y_{k+1}^n|M_{k-1}, M_{k+1}^{(S)}) \Big] + \frac{\epsilon_n}{n} \tag{A6} \\
&\leq \frac{1}{n}\Big[ h(\mathsf{H}_{k,k} X_k^n + Z_k^n) - h(Z_k^n) + h(\mathsf{H}_{k,k+1} X_k^n + Z_{k+1}^n|\mathsf{H}_{k,k} X_k^n + Z_k^n) \\
&\qquad - h(Z_{k+1}^n) + h(Y_{k+1}^n|M_{k+1}^{(S)}) - h(\mathsf{H}_{k,k+1} X_k^n + Z_{k+1}^n) \Big] + \frac{\epsilon_n}{n} \tag{A7}
\end{aligned}
$$

$$\overset{(c)}{\leq} \quad \sum_{i=1}^{L} \frac{1}{2} \log \left( 1 + \Big( \sum_{j=1}^{L} |H_{k+1,k+1}(i,j)| \Big)^2 P + \Big( \sum_{j=1}^{L} |H_{k,k+1}(i,j)| \Big)^2 P \right)$$

$$+ \frac{1}{2} \log \det \left( I_L + H_{k,k+1} H_{k,k}{}^{-1} H_{k,k}{}^{-T} H_{k,k+1}^T \right) \tag{A8}$$

$$+ \frac{1}{2} \log \det \left( H_{k,k}^{-1} H_{k,k}^{-T} + H_{k,k+1}^{-1} H_{k,k+1}^{-T} \right) + \log \det(H_{k,k}) + \frac{\epsilon_n}{n},$$

where $I_L$ denotes the L-by-L identity matrix and $H_{k+1,k+1}(i,j)$ and $H_{k,k+1}(i,j)$ denote the elements of matrices $H_{k+1,k+1}$ and $H_{k,k+1}$ in row $i$ and column $j$. Here, $(a)$ follows because given source messages $M_{k-1}$ and $M_{k+1}$, the triple $(M_k, Y_k^n, Y_{k+1}^n)$ is independent of the rest of the outputs $Y_1^n, \ldots, Y_{k-1}^n, Y_{k+2}^n, \ldots, Y_K^n$ and source messages $M_1, \ldots, M_{k-2}, M_{k+2}, \ldots, M_K$; $(b)$ follows by the chain rule of mutual information and because $M_{k+1}$ is independent of the tuple $(M_{k-1}, M_k, Y_k^n)$; and $(c)$ is obtained by rearranging terms, and the following bounds (A12), (A15) and (A24).

We first bound the term $h(Y_{k+1}^n | M_{k+1}^{(S)})$, and start by noting that because conditioning can only reduce entropy and by the entropy-maximizing property of the Gaussian distribution:

$$h(Y_{k+1}^n | M_{k+1}^{(S)}) \quad \leq \quad \sum_{i=1}^{L} \sum_{t=1}^{n} h(\mathbf{Y}_{k+1,t}(i)) \leq \sum_{i=1}^{L} \sum_{t=1}^{n} \frac{1}{2} \log((2\pi e) \mathrm{Var}(\mathbf{Y}_{k+1,t}(i))), \tag{A9}$$

where $\mathbf{Y}_{k+1,t}(i)$ denotes the $i$-th entry of the vector $\mathbf{Y}_{k+1,t}$. Recall that in this setup without Tx-cooperation the input vectors $X_k^n$ and $X_{k+1}^n$ are independent. However the elements of each input vector can be arbitrarily correlated. The variance $\mathrm{Var}(\mathbf{Y}_{k+1,t}(i))$ is maximized if the elements of $X_{k+1,t}$ are fully correlated and thus :

$$\mathrm{Var}(\mathbf{Y}_{k+1,t}(i)) \leq 1 + \left( \sum_{j=1}^{L} |H_{k,k+1}(i,j)| \sqrt{P_{k,t}(j)} \right)^2 + \left( \sum_{j=1}^{L} |H_{k+1,k+1}(i,j)| \sqrt{P_{k+1,t}(j)} \right)^2, \tag{A10}$$

where $P_{k,t}(j)$ and $P_{k+1,t}(j)$ denote the variances of the $j$-th elements of input vectors $X_{k,t}^n$ and $X_{k+1,t}^n$. In the following we relax the power constraint (7) by requiring only that the power of the $n$ channel inputs produced by any given Tx-antenna cannot exceed $nP$:

$$\sum_{t=1}^{n} P_{k,t}(j) \leq nP, \quad k \in \{1, \ldots, K\} \text{ and } j \in \{1, \ldots, L\}. \tag{A11}$$

Since the right-hand side of (A10) is monotonically increasing and jointly concave in the powers $\{P_{k,t}(j)\}$ and $\{P_{k+1,t}(j)\}$, the upper bound on $\mathrm{Var}(\mathbf{Y}_{k+1,t}(i))$ is largest when $P_{k,t}(j) = P_{k+1,t}(j) = P$. Moreover since also the function $x \mapsto \log(1+x)$ is monotonically increasing, we conclude:

$$h(Y_{k+1}^n | M_{k+1}^{(S)}) \quad \leq \quad n \sum_{i=1}^{L} \frac{1}{2} \log(2\pi e) \left( 1 + \Big( \sum_{j=1}^{L} |H_{k,k+1}(i,j)| \Big)^2 P + \Big( \sum_{j=1}^{L} |H_{k+1,k+1}(i,j)| \Big)^2 P \right). \tag{A12}$$

We next bound the term

$$\frac{1}{n} h(H_{k,k+1} X_k^n + Z_{k+1}^n | H_{k,k} X_k^n + Z_k^n) \quad = \quad \frac{1}{n} h(Z_{k+1}^n - H_{k,k+1} H_{k,k}{}^{-1} Z_k^n | H_{k,k} X_k^n + Z_k^n) \tag{A13}$$

$$\leq \quad \frac{1}{n} h(Z_{k+1}^n - H_{k,k+1} H_{k,k}{}^{-1} Z_k^n) \tag{A14}$$

$$= \quad \frac{1}{2} \log \det(I_L + H_{k,k+1} H_{k,k}{}^{-1} H_{k,k}{}^{-T} H_{k,k+1}^T), \tag{A15}$$

where recall that $I_L$ denotes the L-by-L identity matrix.

For the last bound, define

$$T_k^n \triangleq \mathsf{H}_{k,k}^{-1} Z_k^n - \mathsf{H}_{k,k+1}^{-1} Z_{k+1}^n. \tag{A16}$$

Then:

$$\frac{1}{n}h(\mathsf{H}_{k,k}X_k^n + Z_k^n) - \frac{1}{n}h(\mathsf{H}_{k,k+1}X_k^n + Z_{k+1}^n)$$

$$\overset{(e)}{=} \frac{1}{n}h(X_k^n + \mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n + T_k^n) - \frac{1}{n}h(X_k^n + \mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n) + \log\left(\frac{\det(\mathsf{H}_{k,k})}{\det(\mathsf{H}_{k,k+1})}\right) \tag{A17}$$

$$\overset{(f)}{\le} \frac{1}{n}h(X_k^n + \mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n + T_k^n) - \frac{1}{n}h(X_k^n + \mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n|T_k^n) + \log\left(\frac{\det(\mathsf{H}_{k,k})}{\det(\mathsf{H}_{k,k+1})}\right) \tag{A18}$$

$$= \frac{1}{n}I(X_k^n + \mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n + T_k^n; T_k^n) + \log\left(\frac{\det(\mathsf{H}_{k,k})}{\det(\mathsf{H}_{k,k+1})}\right) \tag{A19}$$

$$\overset{(g)}{\le} \frac{1}{n}I(\mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n + T_k^n; T_k^n|X_k^n) + \log\left(\frac{\det(\mathsf{H}_{k,k})}{\det(\mathsf{H}_{k,k+1})}\right) \tag{A20}$$

$$\overset{(g)}{=} \frac{1}{n}I(\mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n + T_k^n; T_k^n) + \log\left(\frac{\det(\mathsf{H}_{k,k})}{\det(\mathsf{H}_{k,k+1})}\right) \tag{A21}$$

$$\overset{(h)}{=} \frac{1}{n}h(T_k^n) - h(\mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n) + \log\left(\frac{\det(\mathsf{H}_{k,k})}{\det(\mathsf{H}_{k,k+1})}\right) \tag{A22}$$

$$= \frac{1}{2}\log\frac{\det\left(\mathsf{H}_{k,k}^{-1}\mathsf{H}_{k,k}^{-\mathsf{T}} + \mathsf{H}_{k,k+1}^{-1}\mathsf{H}_{k,k+1}^{-\mathsf{T}}\right)}{\det(\mathsf{H}_{k,k+1}^{-1}\mathsf{H}_{k,k+1}^{-\mathsf{T}})} + \log\left(\frac{\det(\mathsf{H}_{k,k})}{\det(\mathsf{H}_{k,k+1})}\right) \tag{A23}$$

$$= \frac{1}{2}\log\det\left(\mathsf{H}_{k,k}^{-1}\mathsf{H}_{k,k}^{-\mathsf{T}} + \mathsf{H}_{k,k+1}^{-1}\mathsf{H}_{k,k+1}^{-\mathsf{T}}\right) + \log\det(\mathsf{H}_{k,k}), \tag{A24}$$

where $(e)$ holds by the definition of $T_k^n$ and because $h(\mathsf{A}X) = \log\det(\mathsf{A}) + h(X)$ for any matrix $\mathsf{A}$ and vector $X$; $(f)$ holds because conditioning can only reduce entropy; (inequalities) $(g)$ hold again because conditioning can only reduce entropy and by the independence of $T_k^n$ and $X_k^n$; and $(h)$ holds because by the independence of the noise vectors we have $h(T_k^n|\mathsf{H}_{k,k+1}^{-1}Z_{k+1}^n + T_k^n) = h(-\mathsf{H}_{k,k+1}^{-1}Z_k^n|\mathsf{H}_{k,k}^{-1}Z_k^n) = h(\mathsf{H}_{k,k+1}^{-1}Z_k^n)$.

Following similar steps as the ones leading to (A12), one can also prove that

$$R_1^{(F)} \le \frac{1}{n}I(M_1^{(F)}; Y_K^n|M_1^{(S)}) + \frac{\epsilon_n}{n} \tag{A25}$$

$$\le \sum_{i=1}^{\mathsf{L}}\frac{1}{2}\log\left(1 + \left(\sum_{j=1}^{\mathsf{L}}|\mathsf{H}_{1,1}(i,j)|\right)^2\mathsf{P}\right) + \frac{\epsilon_n}{n}, \tag{A26}$$

and

$$R_K^{(F)} + R_K^{(S)} \le \frac{1}{n}I(M_K^{(F)}, M_K^{(S)}; Y_K^n|M_{K-1}) + \frac{\epsilon_n}{n} \tag{A27}$$

$$\le \sum_{i=1}^{\mathsf{L}}\frac{1}{2}\log\left(1 + \left(\sum_{j=1}^{\mathsf{L}}|\mathsf{H}_{K,K}(i,j)|\right)^2\mathsf{P}\right) + \frac{\epsilon_n}{n}, \tag{A28}$$

where $\mathsf{H}_{1,1}(i,j)$ and $\mathsf{H}_{K,K}(i,j)$ denote row-$i$, column-$j$ elements of the matrices $\mathsf{H}_{1,1}$ and $\mathsf{H}_{K,K}$.

We sum up the bound in (A9) for all values of $k \in \{1, \ldots, K-1\}$, and combine it with (A26) and (A28). Taking $n \to \infty$, it follows that because the probability of error $p(\text{error})$ vanishes as $n \to \infty$ (and thus $\frac{\epsilon_n}{n} \to 0$ as $n \to \infty$):

$$\sum_{k=1}^{K} \left(2R_k^{(F)} + R_k^{(S)}\right)$$

$$= R_1^{(F)} + \sum_{k=1}^{K-1} \left(R_k^{(F)} + R_k^{(S)} + R_{k+1}^{(F)}\right) + R_K^{(F)} + R_K^{(S)} \tag{A29}$$

$$\leq \sum_{k=1}^{K-1} \left[ \sum_{i=1}^{L} \frac{1}{2} \log\left(1 + \left(\sum_{j=1}^{L} |H_{k+1,k+1}(i,j)|\right)^2 P + \left(\sum_{j=1}^{L} |H_{k,k+1}(i,j)|\right)^2 P\right) \right.$$

$$\left. + \frac{1}{2} \log\det\left(I_L + H_{k,k+1}H_{k,k}{}^{-1}H_{k,k}{}^{-\mathsf{T}}H_{k,k+1}^T\right) + \frac{1}{2}\log\det\left(H_{k,k}^{-1}H_{k,k}^{-\mathsf{T}} + H_{k,k+1}^{-1}H_{k,k+1}^{-\mathsf{T}}\right) + \log\det(H_{k,k})\right]$$

$$+ \sum_{i=1}^{L} \frac{1}{2}\log\left(1 + \left(\sum_{j=1}^{L}|H_{1,1}(i,j)|\right)^2 P\right) + \sum_{i=1}^{L}\frac{1}{2}\log\left(1 + \left(\sum_{j=1}^{L}|H_{K,K}(i,j)|\right)^2 P\right). \tag{A30}$$

Dividing by $K$ and $\frac{1}{2}\log(P)$ and taking $P, K \to \infty$, establishes the converse bound (15).

## Appendix B. Proof of Converse to (21)

Fix a sequence (in the blocklength $n$) of encoding and decoding functions respecting the power constraints and the Tx-cooperation rate-limitations (recall that we consider a setup with only Tx-cooperation but no Rx-cooperation) such that the error probability $p(\text{error}) \to 0$ as the blocklength $n \to \infty$.

Let $\mathbf{M}^{(S)} \triangleq (M_1^{(S)}, \ldots, M_K^{(S)})$. By Fano's Inequality there exists a sequence $\epsilon_n$ satisfying $\frac{\epsilon_n}{n} \to 0$ as $n \to \infty$ such that for any $k \in \{1, \ldots, K-1\}$ and any blocklength $n$:

$$R_k^{(F)} + R_{k+1}^{(S)} + R_{k+1}^{(F)}$$

$$= \frac{1}{n}\left[H(M_k^{(F)}|\mathbf{M}^{(S)}, M_{k-1}^{(F)}) + H(M_{k+1}|M_1^{(S)}, \ldots, M_k^{(S)}, M_{k+2}^{(S)}, \ldots, M_K^{(S)})\right] + \frac{\epsilon_n}{n}$$

$$\leq \frac{1}{n}\left[I(M_k^{(F)}; Y_k^n|\mathbf{M}^{(S)}, M_{k-1}^{(F)}) + I(M_{k+1}; Y_{k+1}^n|M_1^{(S)}, \ldots, M_k^{(S)}, M_{k+2}^{(S)}, \ldots, M_K^{(S)})\right] + \frac{\epsilon_n}{n}$$

$$= \frac{1}{n}\left[h(H_{k,k}X_k^n + Z_k^n|\mathbf{M}^{(S)}) - h(Z_k^n)\right.$$

$$\left. + h(Y_{k+1}^n|M_1^{(S)}, \ldots, M_k^{(S)}, M_{k+2}^{(S)}, \ldots, M_K^{(S)}) - h(H_{k,k+1}X_k^n + Z_{k+1}^n|\mathbf{M}^{(S)})\right] + \frac{\epsilon_n}{n}$$

$$\overset{(a)}{\leq} \sum_{i=1}^{L} \frac{1}{2}\log\left(1 + \left(\sum_{j=1}^{L}(|H_{k+1,k+1}(i,j)| + |H_{k,k+1}(i,j)|)\right)^2 P\right)$$

$$+ \frac{1}{2}\log\det\left(H_{k,k}^{-1}H_{k,k}^{-\mathsf{T}} + H_{k,k+1}^{-1}H_{k,k+1}^{-\mathsf{T}}\right) + \log\det(H_{k,k}) + \frac{\epsilon_n}{n}, \tag{A31}$$

where $(a)$ follows by similar steps as lead to (A15) and (A24), but where one has to account for the fact that due to the Tx-cooperation, the input vectors $X_k^n$ and $X_{k+1}^n$ can be correlated.

Similarly to (A24) one can further prove that

$$R_1^{(F)} \leq \sum_{i=1}^{L}\frac{1}{2}\log\left(1 + \left(\sum_{j=1}^{L}|H_{1,1}(i,j)|\right)^2 P\right) + \frac{\epsilon_n}{n}, \tag{A32}$$

and

$$R_K^{(F)} + R_K^{(S)} \leq \sum_{i=1}^{L}\frac{1}{2}\log\left(1 + \left(\sum_{j=1}^{L}(|H_{K,K}(i,j)| + |H_{K-1,K}(i,j)|)\right)^2 P\right) + \frac{\epsilon_n}{n}, \tag{A33}$$

where again one has to consider that because of the Tx-cooperation the various input vectors can be correlated.

We now sum up the bound in (A31) for all values of $k \in \{1, \ldots, K-1\}$ and combine it with (A32) and (A33). Taking $n \to \infty$, it follows that because the probability of error $p(\text{error})$ vanishes as $n \to \infty$ (and thus $\frac{\epsilon_n}{n} \to 0$ as $n \to \infty$):

$$
\sum_{k=1}^{K} \left( 2R_k^{(F)} + R_k^{(S)} \right)
$$

$$
= \ R_1^{(F)} + \sum_{k=1}^{K-1} \left( R_k^{(F)} + R_k^{(S)} + R_{k+1}^{(F)} \right) + R_K^{(F)} + R_K^{(S)} \right) \tag{A34}
$$

$$
\leq \ \sum_{k=1}^{K-1} \left[ \sum_{i=1}^{L} \frac{1}{2} \log \left( 1 + \Big( \sum_{j=1}^{L} (|\mathsf{H}_{k+1,k+1}(i,j)| + |\mathsf{H}_{k,k+1}(i,j)|) \Big)^2 \mathsf{P} \right) \right.
$$

$$
\left. + \frac{1}{2} \log \det \left( \mathsf{H}_{k,k}^{-1} \mathsf{H}_{k,k}^{-\mathsf{T}} + \mathsf{H}_{k,k+1}^{-1} \mathsf{H}_{k,k+1}^{-\mathsf{T}} \right) + \log \det(\mathsf{H}_{k,k}) \right]
$$

$$
+ \sum_{i=1}^{L} \frac{1}{2} \log \left( 1 + \Big( \sum_{j=1}^{L} |\mathsf{H}_{1,1}(i,j)| \Big)^2 \mathsf{P} \right) + \sum_{i=1}^{L} \frac{1}{2} \log \left( 1 + \Big( \sum_{j=1}^{L} (|\mathsf{H}_{K,K}(i,j)| + |\mathsf{H}_{K-1,K}(i,j)|) \Big)^2 \mathsf{P} \right).
$$

Dividing by $K$ and $\frac{1}{2} \log(\mathsf{P})$ and taking $\mathsf{P}, K \to \infty$, establishes the converse bound (21).

## References

1. Cohen, K.M.; Steiner, A.; Shamai (Shitz), S. The broadcast approach under mixed delay constraints. In Proceedings of the 2012 IEEE International Symposium on Information Theory Proceedings, Cambridge, MA, USA, 1–6 July 2012; pp. 209–213.
2. Zhang, R.; Cioffic, J.; Liang, Y.-C. MIMO broadcasting with delay-constrained and no-delay-constrained services. In Proceedings of the IEEE International Conference on Communications, Seoul, Korea, 16–20 May 2005; pp. 783–787.
3. Zhang, R. Optimal dynamic resource allocation for multi-antenna broadcasting with heterogeneous delay-constrained traffic. *IEEE J. Sel. Top. Signal Process.* **2008**, *2*, 243–255. [CrossRef]
4. Kassab, R.; Simeone, O.; Popovski, P. Coexistence of URLLC and eMBB services in the C-RAN uplink: An information-theoretic study. *arXiv* **2018**, arXiv:1804.06593.
5. Nikbakht, H.; Wigger, M.; Hachem, W.; Shamai (Shitz), S. Mixed delay constraints on a fading C-RAN uplink. In Proceedings of the ITW 2019: IEEE Information Theory Workshop, Visby, Sweden, 25–28 August 2019.
6. Huleihel, W.; Steinberg, Y. Channels with cooperation links that may be absent. *IEEE Trans. Inf. Theory* **2017**, *63*, 5886–5906. [CrossRef]
7. Itzhak, D.; Steinberg, Y. The broadcast channel with degraded message sets and unreliable conference. *arXiv* **2017**, arXiv:1701.05780.
8. Wyner, A.D. Shannon-theoretic approach to a Gaussian cellular multiple-access channel. *IEEE Trans. Inf. Theory* **1994**, *40*, 1713–1727. [CrossRef]
9. Hanly, S.V.; Whiting, P. Information-theoretic capacity of multi-receiver networks. *Telecommun. Syst.* **1993**, *1*, 1–42. [CrossRef]
10. Singhal, M.; Seyfi, T.; Gamal, A.E. Joint Uplink-Downlink Cooperative Interference Management with Flexible Cell Associations. *arXiv* **2018**, arXiv:1811.11986.
11. Lapidoth, A.; Levy, N.; Shamai (Shitz), S.; Wigger, M. Cognitive Wyner networks with clustered decoding. *IEEE Trans. Inf. Theory* **2014**, *60*, 6342–6367. [CrossRef]
12. Wigger, M.; Timo, R.; Shamai (Shitz), S. Conferencing in Wyner's asymmetric interference network: Effect of number of rounds. *IEEE Trans. Inf. Theory* **2016**, *63*, 1199–1226. [CrossRef]
13. Levy, N.; Shamai (Shitz), S. Clustered local decoding for Wyner-type cellular models. *IEEE Trans. Inf. Theory* **2009**, *55*, 4967–4985. [CrossRef]
14. Lapidoth, A.; Shamai (Shitz), S.; Wigger, M.A. On cognitive interference networks. In Proceedings of the 2007 IEEE Information Theory Workshop, Tahoe City, CA, USA, 2–6 September 2007; pp. 325–330.
15. He, B.; Yang, N.; Zhou, X.; Yuan, J. Base station cooperation for confidential broadcasting in multi-cell networks. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 5287–5299. [CrossRef]
16. Annapureddy, V.S.; El Gamal, A.; Veeravalli, V.V. Degrees of freedom of interference channels with CoMP transmission and reception. *IEEE Trans. Inf. Theory* **2012**, *58*, 5740–5760. [CrossRef]
17. Shamai (Shitz), S.; Wigger, M. Rate-limited transmitter-cooperation in Wyner's asymmetric interference network. In Proceedings of the 2011 IEEE International Symposium on Information Theory Proceedings, St. Petersburg, Russia, 31 July–5 August 2011; pp. 425–429.

18. Nikbakht, H.; Wigger, M.; Shamai (Shitz), S. Multiplexing gain region of sectorized cellular networks with mixed delay constraints. In Proceedings of the 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, 2–5 July 2019.

19. Wang, J.; Yuan, B.; Huang, L.; Jafar, S.A. GDoF of Interference Channel with Limited Cooperation under Finite Precision CSIT. *arXiv* **2019**, arXiv:1908.00703.