Radiodiffusion avec CSIT retardée : analyse de SNR fini et voie de retour hétérogène

Chao He

université PARIS-SACLAY

CentraleSupélec

NNT : 2016SACLC091

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS-SACLAY PRÉPARÉE À CENTRALESUPÉLEC

ÈCOLE DOCTORALE N°580

Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat: Réseaux, Information et Communications

par

M. Chao He

Radiodiffusion avec CSIT retardée: analyse de SNR fini et voie de retour hétérogène

Thèse présentée et soutenue à Gif-sur-Yvette, le 2 Decembre 2016:

Composition du Jury :

M.	**Michel Kieffer**,	Université Paris Sud	Président du Jury
Mme.	**Inbar Fijalkow**,	ENSEA	Rapporteur
M.	**Albert Guillén i Fàbregas**,	Universitat Pompeu Fabra	Rapporteur
M.	**David Gesbert**,	Eurecom	Examinateur
Mme.	**Michèle Wigger**,	Télécom ParisTech	Examinateur
M.	**Samir M. Perlaza**,	INSA Lyon	Examinateur
M.	**Sheng Yang**,	CentraleSupélec	Directeur de Thèse
M.	**Pablo Piantanida**,	CentraleSupélec	Co-encadrant de Thèse

I would like to dedicate this thesis...
...to my beloved parents

Acknowledgments

Over past three years and a few more months, countless hours of work, leading to this thesis, has been dedicated in order to take a tiny step further in one specific direction in the telecommunication domain. During the course of preparing this thesis, I was encouraged, helped, and guided by the warmhearted people around that accompanied my 'adventure'. Therefore, I would like to acknowledge all these people for their kindness.

First of all, I am truly grateful to my supervisors Sheng and Pablo, who always helped me along the road, for their generosity of time and enthusiasm. Working with them allowed me to deepen my understanding on the problems, to challenge myself by acquiring more techniques, and to look at things with curiosity. Their guidance is a treasure for my future.

Second, I would like to thank all the kind people I have met during my stay at CentraleSupélec (previously known as Supélec). My dearest officemates Chien-Chun, Meryem, Germán, Clément, Victor, Hernan, and Qianqian, and and all the other students and postdocs, Andrés, Stefano, Matha, Zheng, Dora, Gil, Kenza, Ejder, Bakarime, Salah Eddine, Asma, Jane, Hafiz, Maialen, Fei, Hoang, Abdulaziz and many others, with whom I shared delightful time. I have also a deep gratitude towards the personnel of the telecommunication department, Catherine, José and Huu-Hung for the assistance they gently offered.

I am also thankful to a special group of friends, Yijing, Wenjie, Wenmeng, Kai, Chao, Li, Shanshan, Xin, Yan, Huaxia, and the people surrounded them, like Xi, Xiaofei, and Lulu. We enjoyed chatting, dinnering, playing, and sometimes discussing together, which 'refueled' me after work.

Pendant les années passé, j'ai eu la chance de rencontre Vanessa, Aurélien, Irina, Yann, Lilianne, Syndie, Maud, et Mélanie, qui m'ont encouragés à apprendre le français et avec qui j'ai bien profité de la vie et des vacances en France. C'est grâce à eux que j'ai commencé à comprendre la culture française.

Finally, I would like to give my deepest thank to the beloved parents and my family in China, without whose supports this thesis may never be possible. I am grateful for their understanding and support on my decisions. Though we did not have long conversations everyday, we exchanged a few words regularly and it is already enough to learn that they will always back me up.

Contents

List of Figures

List of Tables

Acronyms

AEP	asymptotic equipartition property.
AWGN	additive white Gaussian noise.
BC	broadcast channel.
BRC	broadcast relay channel.
BS	base station.
CF	compress-and-forward.
CSI	channel state information.
CSIT	channel state information at the transmitter.
DC	distributed compression.
DCF	decode-compress-and-forward.
DF	decode-and-forward.
DoF	degrees of freedom.
EBC	erasure broadcast channel.
FME	Fourier Motzkin elimination.

GBC	Gaussian broadcast channel.
GMAT	Generalized MAT (scheme).
i.i.d.	independent and identically distributed.
IC	interference channel.
MAT	Maddah-Ali-Tse.
MIMO	multiple-input multiple-output.
MISO	multiple-input single-output.
OFDMA	orthogonal frequency-division multiple access.
pDCF	partial decode and compress forward.
pdf	probability distribution function.
pmf	probability mass function.
QMAT	Quantized MAT (scheme).
RC	relay channel.
RDC	relay-destination channel.
RIA	retrospective interference alignment.
RV	random variable.
SDC	source-destination channel.
SE DCF	simultaneous-emitted decode-compress-and-forward.
SISO	single-input single-output.
SNR	signal-to-noise ratio.
SRC	source-relay channel.

SW	Shayevitz-Wigger.
TD DCF	time-division decode-compress-and-forward.
TDMA	time-division multiple access.
w.p.	with probability.

Notations

x	constants or realizations of random variables by lowercase letters.
X	random variables by capital letters.
$\boldsymbol{x}$	vectors by boldface lower-case letters, sometimes are also used to represent sequences if it is clear.
$\boldsymbol{X}$	matrices by boldface upper-case letters.
x^n	a sequence $x_1, x_2, \cdots, x_n$.
$\mathcal{X}$	sets by calligraphic letters, ϕ is the null set.
$\mathcal{J}, \mathcal{I}$	two sets with cardinality constraints, $\lvert\mathcal{I}\rvert = i, \lvert\mathcal{J}\rvert = j$.
$\mathrm{diag}(x_1, ..., x_N)$	diagonal matrix with entries $x_1, ..., x_N$.
$\log(x)$	logarithm of x (base 2).
$\mathrm{tr}(\boldsymbol{X})$	the trace of matrix $\boldsymbol{X}$.
$\det(\boldsymbol{X})$	the determinant of matrix $\boldsymbol{X}$.
$\boldsymbol{X}^{-1}$	the inverse of matrix $\boldsymbol{X}$.
$\mathcal{X} \cup \mathcal{Y}, \mathcal{X} \setminus \mathcal{Y}$	the union of sets and the relative complement of the set $\mathcal{X}$ in $\mathcal{Y}$.
$H(\cdot), h(\cdot)$	the entropy and the differential entropy.
$I(\cdot;\cdot)$	the mutual information.
$\Pr(\cdot)$, or $P(\cdot)$	the probability measure.
$\lVert\cdot\rVert$	the Euclidean norm.

$\mathbb{E}(\cdot)$	expectation operator.
$(\cdot)^T$	the transpose of vector/matrix.
$(\cdot)^H$	the conjugate transpose of vector/matrix.
$\mathcal{T}_\delta^n(\cdots)$	the set of weak typical n-length sequences.
$\lvert\cdot\rvert$	the cardinality of the corresponding set, or the absolute value of the corresponding number.
$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	the circularly symmetric Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.
$\mathbb{R}_+$	the set of real non-negative numbers.
$[a : b]$	the set of integers between a and b with a, b included in the set.
$\mathcal{X}^c$	the complement of the set $\mathcal{X}$.

1 Introduction and Preliminaries

In this chapter, we prelude the motivation of our research, introduce several models we investigated, review some fundamental results as background material for the models studied, and then present the thesis organization and the list of publications.

1.1 MOTIVATION

One of the essential targets of researchers in communications is to provide efficient techniques accommodating the explosive demands for internet access. The reasons for this considerable connection requirements are three-fold: 1) the number of mobile devices is growing exponentially over the past years; 2) the number of wireless data application in each terminal increases dramatically; 3) the amount of data required by every application becomes immense. To accommodate such unprecedented traffic demands, enormous devices of different types are served simultaneously in a multiple base station (BS) system. Consequently, the system designers will have to face at least three fundamental issues that make the problem quite challenging, i.e., *interference*, *limited power*, and *heterogeneity* (defined in Section 1.1.3). Due to the massive user terminals in the network, the communication channel is transferred from noise-limited into interference-limited. With the same reason, the power resource allocated to the information intended for each user is strongly restricted given a total power constraint. As the wireless techniques evolve over generations, the multiple BS could be subject to different standards. As such, incorporating multiple heterogeneous BS into a heterogeneous network is indeed difficult as it may prevent them from formulating a coherent cooperation. In addition, heterogeneity does not only exist from the perspective of the BSs. For instance, when distinct mobiles want to feedback some information to the BSs to assist palliating interference, the feedback rates of some mobiles could be negligible because of the limited power budget at the device, the intense fluctuation of feedback channels, the absence of line-of-sight transmission path, e.t.c. In this case, the feedback heterogeneity takes place as it is preferable to shut down these weak feedback links for the efficiency concern. The aforementioned three aspects can be modeled and addressed separately, however, in real system, they occur mostly at the same time.

It is well known by now that the knowledge of channel state information at the transmitter (CSIT) is crucial for aforementioned multi-user wireless networking systems [1]. In particular, CSIT enables to mitigate the interference between information flows, which is the main obstacle for higher rates in multi-user

communications. Nevertheless, timely and accurate CSIT is either difficult, if possible at all, or costly to obtain. Although accurate CSIT can be eventually improved with enough resources (e.g. channel estimation, CSIT feedback, etc.), the timeliness highly depends on the nature of the physical channel. As a matter of fact, accurate channel state information (CSI) may as well become outdated when it is feedback to the transmitter in a high mobility environment where the channel varies rapidly.

In the recent work [2] on the Gaussian broadcast channel (GBC), Maddah Ali and Tse proposed a linear scheme (a.k.a Maddah-Ali-Tse (MAT) scheme) and revealed that with MAT scheme, even completely outdated CSIT can be exploited to increase substantially the throughput of the two-user multiple-input single-output (MISO) broadcast channel (BC). In this case, it achieves a sum-degrees of freedom (DoF) (identified as the pre-log factor) of $\frac{4}{3}$ as compared to 1 in the case with no CSIT. In the same work, the MAT is shown to achieve the optimal DoF for the K-user MISO case when the number of transmit antennas is not less than the number of receivers K. This conclusion holds as long as the transmitter can receive the CSI feedback (with arbitrary but finite delay) from all receivers and that the CSI error vanishes as $O(\mathsf{SNR}^{-1})$. Specifically, their scheme is built on the space-time alignment using delayed CSIT[1], and is shown to be optimal in terms of the DoF at high signal-to-noise ratio (SNR). As a matter of fact, the same principle was already used beforehand by Dueck in [3]. The optimality is straightforward from the genie-aided outer bound which provides the channel output of one user to the other and thus reduces the problem to that of the well-known *physically degraded* BC with feedback [4]. After [2], several generalizations have been carried out, such as, among others, the multiple-input multiple-output (MIMO) interference channel (IC) [5], $K-$user single-input single-output (SISO) IC/X channel [6], MIMO-BC/IC with mixed CSIT [7–9], MISO-BC with evolving CSIT [10], MISO-BC with alternating CSIT [11], certain three-user MIMO-BC [12, 13], secure communication of MIMO-BC [14], some other works [15–18] and the references therein. While these works are insightful from a signal-space point of view, i.e., space/time interference alignment, most of them focus on DoF. We may refer to the state-based interference mitigation (with the same principle) as MAT or retrospective interference alignment (RIA) (first summarized in [19]) alternatively for the sake of diversity.

1.1.1 *Delayed CSIT at finite SNR*

Although the delayed-CSIT-enabled interference mitigation has drawn remarkable attention due to its outstanding performance in high SNR regime, i.e., verified by DoF gain, whether delayed CSIT is always effective for fading GBC at *finite* SNR seems to be a more appealing question for practical wireless systems. As we know, the MAT scheme has a fixed structure built on a dimension counting argument, the DoF optimality at high SNR may not ensure its efficiency at finite SNR due to its insensitivity to such parameters. In fact, some inner bound results on the capacity of the memoryless BC have already been presented, e.g., with generalized feedback in general channels [20–22] and reference therein. In addition, there have been

[1] Delayed CSIT and state feedback are used alternatively in this thesis.

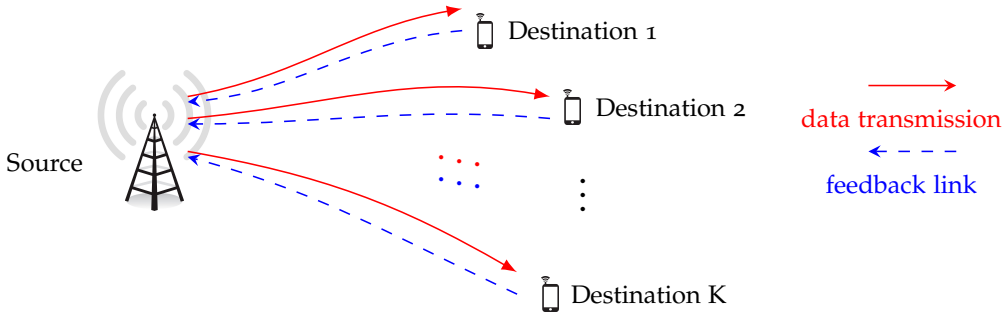

Figure 1: An example of homogeneous feedback

several other achievable rate region results that are obtained by finite-SNR oriented schemes based on the aforementioned MAT structure, e.g, in [23–26]. In [23], authors developed for K-user two precoding methods that design the precoders such that certain amount of interference at one user is allowed so that the received signal at the other is enhanced. For $K = 2, 3$, its performance gain over MAT was revealed when a specific type of decoder is selected. To reduce the penalty of multicasting in the MAT scheme, the authors of [24] included quantization on the signal that bears the auxiliary common message. For the K-user Rayleigh fading case, they demonstrated that a gap between the corresponding inner bound and a genie-aided outer bound, in terms of the symmetric rate, is upper bounded by $2\log_2(K+2)$ which scales sublinearly with K. More recently, [25] analyzed a scenario where both the CSI statistics and the delayed feedback of channel realizations are available at the transmitter. It was shown by numerical examples that statistics of CSI can enlarge the rate region for temporally correlated Rayleigh fading GBC. The authors of [26] investigated outage performance for GBC with an adapted MAT scheme. It is worth mentioning that these schemes applied either linear coding alone or linear coding with quantization to the fixed transmission length structure that is used in the MAT scheme. Despite that performance gain over MAT is possible with some of the schemes therein, their rates in *medium-to-low* SNR regime are dominated by that of the time-division multiple access (TDMA) strategy (optimal for no CSIT). Therefore, finding a scheme that exploits the potential of delayed CSIT at *finite* SNR/with *limited power* is still an attractive issue.

1.1.2 *Delayed CSIT in erasure broadcast channel (EBC) and GBC*

Across the independent works [27, 28], the capacity region of the EBC with causal state feedback was fully determined in the three-user case and partially characterized for $K > 3$ cases. The main idea behind the schemes proposed in these references is fundamentally the same: the source first sends out the uncoded packets, then generates according to the state feedback some adequate "linear combinations" of the packets that are lost by certain receivers but overheard by some others and multicast these linear combinations to certain corresponding receivers. Given that sufficient linearly independent combinations are available to each receiver, the desired packets (messages) are always ready to be recovered.

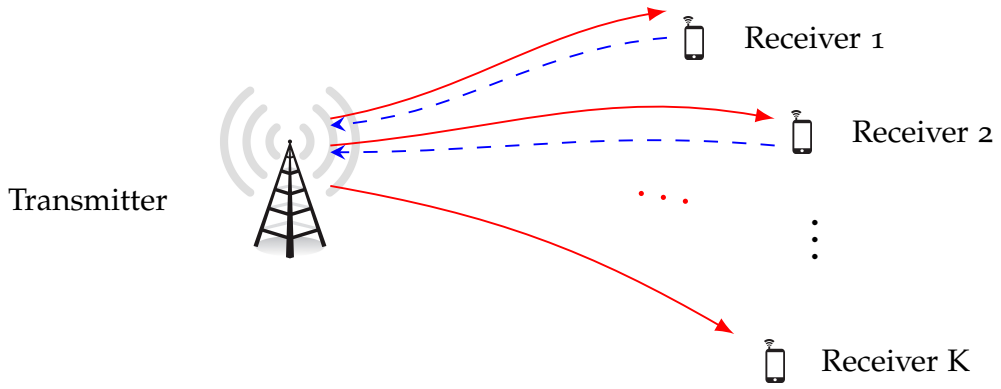

Figure 2: An example for heterogenity at the receivers' side

However, the schemes in [27, 28] only work for a packet alphabet of size 2^q with $q \in \mathbb{N}$ such that $2^q \geq K$. This is to guarantee the existence of a desired number of linearly independent vectors in the corresponding vector space. As a consequence, the capacity region is still open for other input alphabet sizes. Essentially, there is striking similarity between the MAT scheme and the schemes described in [27,28]. Specifically, the MAT scheme follows the routine as: first send out the signal uncoded, then create based on the CSI feedback and multicast the linear combinations of the overheard signals, so that the messages can be easily recovered, e.g., by Gaussian elimination, with vanishing decoding error if enough equations are available to each user. Besides, there is a resemblance between the DoF region of GBC and the capacity region of EBC. As a matter of fact, such similarity also occurred between the secure DoF/rate regions of GBC/EBC as shown in [29, 30].

1.1.3 *Heterogeneous feedback*

As claimed previously, the interference mitigation techniques might not be efficient in physical channels if we let it stand alone by excluding the concerns on limited power and/or *heterogeneity*. Motivated by this, we consider the RIA under the heterogeneous state feedback condition. For convenience, we state briefly the definition of homogeneous/heterogeneous state feedback specifically for this thesis. All the communicating nodes participate in the feedback process when we assume homogeneous feedback, e.g., as shown in Fig 1. In contrast, heterogeneous state feedback is defined as only part of the communicating nodes participates in the state feedback process; whereas homogeneous state feedback means that the delayed CSI of all the destinations are feedback to all transmitters. To be specific, we give two examples to illustrate distinct heterogeneous state feedback. The first instance is that the feedback heterogeneity lies in the receivers' side, that is, some of the receivers feedback their delayed CSI to all the senders while no feedback from others are detected, e.g., as shown in Fig 2. In this case, the source does not grasp the status of the whole system and has to design its transmission based on the partial/heterogeneous state. The second example is when the heterogeneity locates at the transmitters' side, i.e., though all the

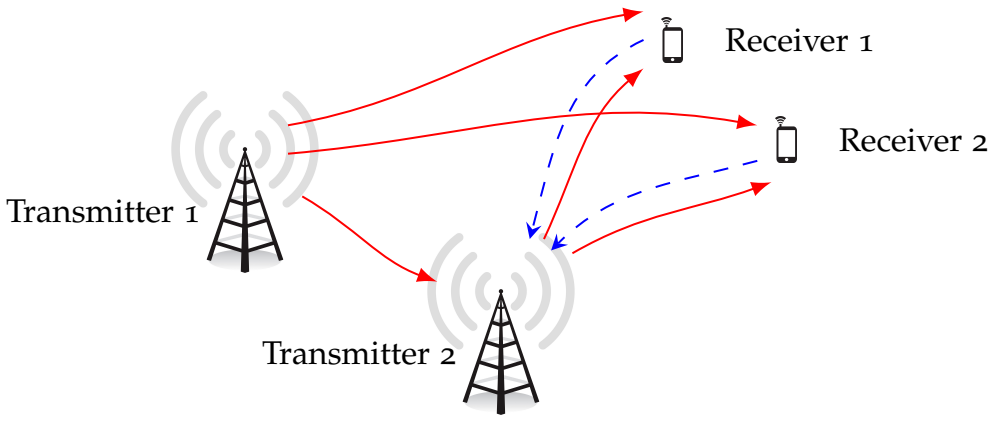

Figure 3: An example for heterogenity at the transmitters' side

users send back their states, merely few transmitters are capable of collecting the feedback information because of either the inferior quality of feedback link or the protocol settings.[2] One of the examples is shown in Fig 3. Even though heterogeneities can happen simultaneously at both ends of communication, for clarity purpose, we omit such definition.

State feedback from partial receiver(s)

Hybrid CSIT represents that states about distinct source-destination channel (SDC) available to the transmitter are of different qualities, e.g., delayed CSIT ('D'), no CSIT ('N'), perfect and instantaneous state at the source ('P'). As it is ubiquitous and general, hybrid CSIT, including heterogeneous state feedback as special cases, is of practical significance and has been extensively studied with a special emphasis on DoF [7, 11, 31–37]. Although we verified in [38–40] the performance gain with DD state at almost all applicable SNRs, the improvement is shown to vanish even in high SNR regimes and the DoF collapse to unity when the system subjects to PN state [41], let alone the DN case (two user heterogeneous feedback). In other words, from DoF's perspective, introducing a user into a point-to-point channel with perfect/delayed CSIT is not beneficial if the CSI of new user is missing. Similar results have been reported in [36] for beyond two user cases. The sum DoF of three user BC with state DDN and state PDN are limited to $\frac{4}{3}$ and $\frac{3}{2}$ which coincides with the optimal DoF of two user case with state DD [2] and PD [42], respectively.

Owing to the similarity (or we can call it uniformity) between the DoF region of GBC and the capacity of EBC in DD case, it sounds reasonable to predict a degraded capacity in EBC with DN state given the inferior DoF performance. Quite surprisingly, Lin and Wang proved in [43, 44] that the CSI feedback from a single user can still be useful when a symmetric spatially independent EBC is investigated. To see this, they proposed a multiple phase algorithm which is capable of exploiting an *opportunistic* coding gain by linear network coding. The

[2]The heterogeneity can be created deliberately (by the system designer) such that users feedback simply to few senders to address the shortage of feedback power.

essential ingredient of the approach lies in that the transmitter *XOR*s the intended but erased bits and the undesired but observed bits at one user (the one provides D state), sends them as side information, and let both user decodes wanted bits opportunistically along with the previously overheard bits. Encouraged by that, we investigated the BC with feedback heterogeneity at receivers' side. In fact, authors of [43, 44] are not the only researchers who notice the potential of heterogeneous state feedback. In particular, [41, 45, 46] took a finite state BC with additive white Gaussian noise (AWGN) as a counter example to the degraded performance conjecture for general probability mass function (pmf), i.e., a sum DoF strictly bounded away from unity is shown in [47, 48] to be achievable for DN state. Regardless of the difference on coding tools employed in [43, 44, 47, 48], the idea behind is consistent. The source encodes the side information only dependent on the partial state while the receiver knows both state so that it can decide which part to decode based on whether the corresponding state realization does conform to that in DD's scheme, e.g., receiver 2 decides to decode auxiliary bits whose states are $(S_1, S_2) = (1, 0)$ because these bits are desired by user 2 and are encoded into the side information.

State feedback to partial transmitter(s)

Given that dense BSs subject to distinct standards are deployed in the network, it is intuitive to introduce cooperation between multiple-tier BSs via wired (distributed antenna system with backhaul) or wireless (relaying) links to guarantee a coherent transmission. Thus, we consider a typical heterogeneous network setup in which a macrocell BS serves two users who are in close proximity to a smallcell (e.g., femto/pico) BS. A natural question here is whether and how smallcell BS could improve downlink performance of the macrocell network. Essentially, the smallcell BS in this setting may be considered as a relay, and the overall channel becomes a broadcast relay channel (BRC).

We revisit the main idea of MAT, which consists in two phases: 1) the broadcast phase in which signal containing new messages is sent to the receivers, and 2) the multicast phase in which side information generated based on both the past channel and the sent messages is diffused. It is therefore instinctive to think of the transmitter as two separate virtual transmitters: one takes charge of the first phase without the need of any CSIT, whereas the other is responsible for the second phase and uses the delayed CSIT and the transmitted message. The second virtual transmitter is very much like a relay that "learns" the messages and the channel coefficients and "forwards" the necessary side information to "help" the receivers decode their own messages.

With the above reasoning, in Chapter 5, we deploy an infrastructure relay that helps accomplish the RIA *a la MAT* with the heterogeneous state feedback from every mobiles to the smallcell BS (relay) instead to the macrocell BS (source). Our setting on feedback is mainly inspired by the fact that feeding back CSI accurately to the macrocell BS via uplink is costly in practice. The cost is in terms of both the uplink bandwidth and the uplink transmit power related to the feedback. Since the relay is assumed to be closer to the destinations than the source is, feedback to the relay instead may be more feasible. In other words, the feedback

communication is localized and the feedback power could be reduced thanks to lower channel attenuation.

In the next several sections, we introduce three channel models that will be studied in this thesis, i.e., BC with homogeneous state feedback, BC with heterogeneous state feedback and BRC with state feedback at the relay. With respect to the length of a dissertation, reviewing all the related works does not seem to be feasible. Therefore, we present the most relevant results in the channel models that are investigated.

1.2 BROADCAST CHANNEL WITH HOMOGENEOUS STATE FEEDBACK

As a first step to verify the potential improvement with RIA at finite SNR, we choose the possibly simplest model in the network coding problem, namely, the BC. We will study the BC with homogeneous state feedback in Chapter 2 and 3. The channel is a memoryless state-dependent K-user BC through which the source wishes to communicate with the help of state feedback from all destinations, in n channel uses, K independent messages to the K receivers, respectively. The channel can be described by the joint pmf,

$$p_{\mathbb{Y}_1,\cdots,\mathbb{Y}_K|\mathbb{X},\mathbb{S}}(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_K|\boldsymbol{x},\boldsymbol{s})p_{\mathbb{S}}(\boldsymbol{s}) = \prod_{i=1}^{n} p_{Y_1,\cdots,Y_K|X,S}(y_{1i},\ldots,y_{Ki}|x_i,s_i)p_S(s_i), \quad (1)$$

where $\boldsymbol{x} \in \mathcal{X}^n, \boldsymbol{y}_k \in \mathcal{Y}_k^n$, and $\boldsymbol{s} \in \mathcal{S}^n$ for all $k \in \mathcal{K} \triangleq \{1, 2, \cdots, K\}$ are the sequences of the channel input, the k-th channel output, and the channel state, respectively. $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{S}$ are the alphabets of the channel input, the channel output, and the state, respectively. The global channel state information is known instantaneously to all the receivers, e.g., each user has access to all the error events of EBC or all the channel coefficients of GBC immediately after the transmissions. At the transmitter's side, the channel state is known strictly causally without error, e.g., via a noiseless feedback link. Unless it is stated otherwise, the cost of the state feedback is not taken into account throughout the dissertation as we assume sufficient resource is allocated to guarantee an unbounded feedback rate. For brevity, we assume that the state information is provided at the transmitter with one slot delay and the channel itself is temporally independent and identically distributed (i.i.d.). The details of the channel model is shown in Fig. 4. As it is shown in Fig. 4, the state information of first $i-1$ channel uses, namely, s^{i-1}, is available to the transmitter only at the end of ith channel use. For any subset $\mathcal{U} \subseteq \mathcal{K}$, we define $S_{\mathcal{U}}$ and $Y_{\mathcal{U}}$ as the combinations of states and channel outputs of users in $\mathcal{U}$, respectively.

Let $M_k \in \mathcal{M}_k \triangleq \{1, 2, \cdots, 2^{nR_k}\}$ be the message for user k chosen from the message set $\mathcal{M}_k$, for $k \in \mathcal{K}$. We say that the rate tuple $(R_1, \ldots, R_K)$ is achievable if there exist

- a sequence of encoding functions $\{f_i : \mathcal{M}_1 \times \cdots \times \mathcal{M}_K \times \mathcal{S}^{i-1} \to \mathcal{X}\}_{i=1}^n$,
- K decoding functions $\{g_k : \mathcal{Y}_k^n \times \mathcal{S}^n \to \mathcal{M}_k\}_{k=1}^K$,

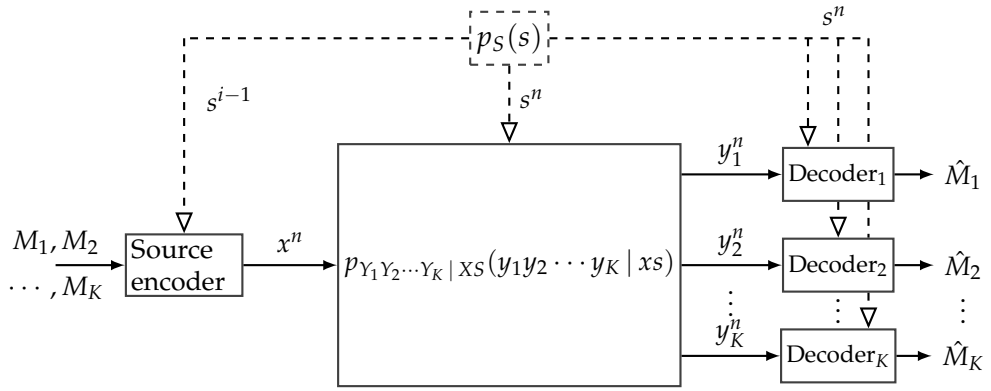

Figure 4: General system model of K-user BC with state feedback.

such that $\max_k \Pr\big(g_k(Y_k^n, S^n) \neq M_k\big) \to 0$ when $n \to \infty$. The symmetric rate R_{sym} is achievable if the rate tuple $(R_{\text{sym}}, \ldots, R_{\text{sym}})$ is achievable.

Similar to the no feedback case, the capacity of BC with feedback is fully characterized in physically degraded case [4] but is still unknown in general. Dueck illustrated in [3] that feedback can enlarge the capacity region for some classes of BC by letting the source broadcast CSI of common interest to both receivers. Ever since, some remarkable works have been reported, e.g., [20,22,49,50] and reference therein, amongst which [20] provides a most relevant region to our work. In particular, Shayevitz and Wigger studied the two-user broadcast channel with generalized feedback and proposed an achievable rate region, quoted hereafter as Shayevitz-Wigger (SW) region, based on double binning and block Markov coding techniques in a separate source-channel coding framework. Albeit the relevance, one shall note that we will focus on the state feedback whereas their work concentrated on generalized feedback.

In the next two subsections, two subclass channels of BC are defined in which we illustrate how the homogeneous state feedback is exploited in these channels.

1.2.1 *Fading Gaussian broadcast channel*

The Gaussian channel is employed frequently in research as it captures the main properties of wireless communications. The capacity region of the GBC is known when the channel state is deterministic [51,52]. When the channel is subject to fading unknown to the transmitter (no CSIT), the capacity is still open except for some particular fading processes such as the case when both users experience the exact same fading distribution [53]. For more general cases, only high SNR characterizations in terms of DoF are available in the literature.

The Fading GBC is a state-dependent AWGN channel with n_t transmitting antennas and $n_{r,k}$ receiving antennas at user k, defined by

$$\boldsymbol{y}_{k,i} = \boldsymbol{H}_{k,i}\boldsymbol{x}_i + \boldsymbol{z}_{k,i}, \quad k = 1, \ldots, K;\ i = 1, \ldots, n \tag{2}$$

where $\boldsymbol{z}_{k,i} \in \mathbb{C}^{n_{\text{r},k}\times 1}$, $\boldsymbol{x}_i \in \mathbb{C}^{n_{\text{t}}\times 1}$, $\boldsymbol{y}_{k,i} \in \mathbb{C}^{n_{\text{r},k}\times 1}$, and $\boldsymbol{H}_{k,i} \in \mathbb{C}^{n_{\text{r},k}\times n_{\text{t}}}$. Moreover, we assume the noise $\boldsymbol{z}_{k,i} \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_{n_{\text{r},k}})$ and the input $\boldsymbol{x}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{Q})$. The channel

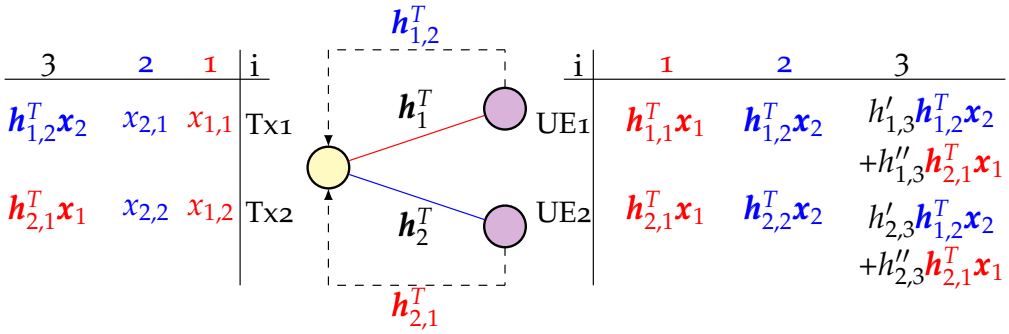

Figure 5: Example of MAT scheme.

input is subject to the power constraint $\frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{x}_i\|^2 \leq P = \mathrm{tr}(\boldsymbol{Q})$ for any input sequence $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. The subscript $i \in \{1, \cdots, n\}$ is omitted in general and will be specified when necessary. The SNR is defined as $\mathsf{SNR} \triangleq \frac{P}{n_t \sigma^2}$. Unless indicated otherwise, we assume that the channel matrices are independent across users.

We identify the state S_k with the channel matrix $\boldsymbol{H}_k$. Similarly as for $S_{\mathcal{U}}$, we use $\boldsymbol{H}_{\mathcal{U}}$ to denote a matrix from a vertical concatenation of all the matrices $\{\boldsymbol{H}_k\}_{k\in\mathcal{U}}$, same notation applies for $\boldsymbol{Y}_{\mathcal{U}}$ and $\boldsymbol{Z}_{\mathcal{U}}$. Hence, it follows that $\boldsymbol{Y}_{\mathcal{U}} = \boldsymbol{H}_{\mathcal{U}}\boldsymbol{X} + \boldsymbol{Z}_{\mathcal{U}}$.

Definition 1.1. *A Fading Gaussian BC as stated above is said to be symmetric if the channel matrices $\boldsymbol{H}_k$ follow an identical distribution for all $k \in \mathcal{K}$, i.e., $f(\boldsymbol{H}_k = \boldsymbol{H}) = f(\boldsymbol{H}_{k'} = \boldsymbol{H})$, $k, k' \in \mathcal{K}$, where $f(\cdot)$ defines the probability density function.*

MAT scheme in GBC with homogeneous state feedback

A two user MISO fading GBC is considered with $n_t = 2$ and $n_{r,1} = n_{r,2} = 1$, which leads to $\boldsymbol{H}_{k,i} = \boldsymbol{h}_{k,i}^T = [h'_{k,i}, h''_{k,i}]$. We give a concise explanation on how one variant of the MAT scheme, as illustrated in Fig. 5, can achieve a sum DoF of $\frac{4}{3}$ in the model described. $\boldsymbol{x}_k = [x_{k,1}\ x_{k,2}]^T$ contains two streams intended for user k where k equals to 1 or 2 and $x_{k,j}$ stands for the channel input symbol emitted via jth antenna. Once the transmitter sends the $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in first two slots, the destination 1 and 2 feedback to the source the fading coefficients $\boldsymbol{h}_{1,2}^T$ and $\boldsymbol{h}_{2,1}^T$, respectively. With the delayed state, the source reconstructs the two overheard streams $\boldsymbol{h}_{1,2}^T\boldsymbol{x}_2$ and $\boldsymbol{h}_{2,1}^T\boldsymbol{x}_1$ and send them independently in two antennas. Destination 1 can get two linearly independent equations $(\boldsymbol{h}_{1,1}^T\boldsymbol{x}_1, \boldsymbol{h}_{2,1}^T\boldsymbol{x}_1)$ of $\boldsymbol{x}_1$, from which $\boldsymbol{x}_1$ is attainable by sorting and performing a Gaussian elimination on the two observations. Thus, the two streams $\boldsymbol{x}_1$ is communicated to user 1 in three slots which leads to a $\frac{2}{3}$ DoF. Analogously, $\frac{2}{3}$ DoF/rate is also achievable at user 2 which renders a $\frac{4}{3}$ sum DoF.

In fact, Kim *et al.* demonstrated in [54] that in the two-user symmetric setting, the SW scheme actually includes the MAT scheme as a special case. Thus, it attains the same DoF as MAT. However, the comparison on rate regions between the two in *asymmetric* case is not straightforward and thus is unclear in general.

1.2.2 *Erasure broadcast channel*

The erasure channel by itself plays an important role in communication due to its usefulness on modeling the packet decoding error from a higher layer beyond Physical layer. An EBC is a BC whose outputs subject to possible eliminations.

The EBC is a state-dependent deterministic channel defined by

$$Y_k = \begin{cases} X, & S_k = 1, \ with\ probability(w.p.)\ 1-\delta_k \\ E, & S_k = 0, \ w.p. \quad \delta_k \end{cases}, \tag{3}$$

which is completely characterized by the set of probabilities of the state realization

$$\mathrm{P}_{\mathcal{U},\bar{\mathcal{U}}} \triangleq \Pr(S_{\mathcal{U}} = \mathbf{0},\ S_{\bar{\mathcal{U}}} = \mathbf{1}), \quad \mathcal{U} \subseteq \{1,\dots,K\}, \tag{4}$$

with $\sum_{\mathcal{U}} \mathrm{P}_{\mathcal{U},\bar{\mathcal{U}}} = 1$. Throughout the thesis, we use $S_{\mathcal{U}} = \mathbf{0}$ (*resp.* $S_{\mathcal{U}} = \mathbf{1}$) to mean that $S_k = 0$ (*resp.* $S_k = 1$), $\forall\ k \in \mathcal{U}$. For the sake of simplicity, we denote $\Pr(S_{\mathcal{F}} = 0)$ with $\delta_{\mathcal{F}}$ and $\Pr(S_{\mathcal{F}} = 0, S_{\mathcal{T}} = 1)$ with $\mathrm{P}_{\mathcal{F},\mathcal{T}}$ for any $\mathcal{F}$ and $\mathcal{T}$ that satisfy $\mathcal{F} \cap \mathcal{T} = \phi$ and $\mathcal{F} \cup \mathcal{T} \subseteq \mathcal{K}$, which correspond to the event that erasures occur to all the users in set $\mathcal{F}$ and the event that every one in set $\mathcal{F}$ encounters an erasure while each user in set $\mathcal{T}$ receives perfectly the input signal, respectively. In Chapter 3 and 4, we use $\mathcal{I}$ and $\mathcal{J}$ for some subsets of users with implicit size constraints $|\mathcal{I}| = i$ and $|\mathcal{J}| = j$, respectively. The constraints are made explicit when necessary. Moreover, $\mathcal{F}$, $\mathcal{T}$, $\mathcal{U}$ are also assumed to be the subsets of users but without particular size constraints. We recite the Lemma 10 in [28] here which shows $\mathrm{P}_{\mathcal{F},\mathcal{T}}$ and δ are connected by,

$$\mathrm{P}_{\mathcal{F},\mathcal{T}} = \sum_{\mathcal{U} \subseteq \mathcal{T}} (-1)^{|\mathcal{U}|} \delta_{\mathcal{F} \cup \mathcal{U}}, \tag{5}$$

For notational convenience, $\delta_{\{k\}}$ is written as δ_k. We further define two subclass of erasure channel, symmetric EBC and spatially independent EBC.

Definition 1.2. *An EBC is said to be symmetric if for any $\mathcal{J}$, the erasure probability $\delta_{\mathcal{J}}$ is determined only by the cardinality of the set $\mathcal{J}$, that is, $\delta_j = \delta_{\mathcal{J}} = \delta_{\mathcal{J}'}$ always holds for any sets $\mathcal{J}$ and $\mathcal{J}'$ such that $|\mathcal{J}| = |\mathcal{J}'| = j$.*

Definition 1.3. *An EBC is assumed to be spatially independent as long as the erasures across users are i.i.d.. For the spatially independent EBC, we have:*

$$\delta_{\mathcal{U}} = \prod_{k \in \mathcal{U}} \delta_k, \quad \mathcal{U} \subseteq \mathcal{K}, \tag{6}$$

which yields

$$\mathrm{P}_{\mathcal{F},\mathcal{T}} = \prod_{k \in \mathcal{F}} \delta_k \prod_{k' \in \mathcal{T}} (1-\delta_{k'}), \quad \mathcal{F} \cap \mathcal{T} = \phi,\ \mathcal{F} \cup \mathcal{T} \subseteq \mathcal{K}. \tag{7}$$

Without loss of generality, we assume that the marginal erasure probabilities fulfill $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_K$. We restate the definition of 'one-sided fairness' or 'one-sidedly fair condition' (first defined in [27, 28]) as follows.

Definition 1.4. *A rate tuple $(R_1, \cdots, R_K)$ is said to satisfy one-sidedly fair condition or*

to be one-side fair if given $\delta_i \geq \delta_j$, $R_i\delta_i \geq R_j\delta_j$ holds for any $i \leq j$.

Linear coding in EBC with homogeneous state feedback

The state feedback from both users has been demonstrated to be beneficial in EBC. To see this, we present an example in two user case.

Theorem 1.1 (Theorems in [27, 28, 55])**.** *The capacity region of a two user packet EBC with state feedback from both users at the transmitter is characterized by,*

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq L, \tag{8a}$$

$$\frac{R_1}{1-\delta_{12}} + \frac{R_2}{1-\delta_2} \leq L. \tag{8b}$$

where L is the number of bits per packet (packet length) in the packet EBC.

Sketch of proof. The converse is quite standard and is provided in several works including [27, 28, 55]. Hence, we only present achievability by a multiple phase approach that employs the linear combination of packets as follows.

Phase 1: the source broadcasts M_1 packets intended for user 1. In the case of erasure, each packet is re-transmitted until at least one of the receivers observes it. Then, with the help of state feedback, the transmitter collects in a queue Q_1 the packets destined to user 1 but seen only by user 2. Thus, the length of phase 1 is $N_1 = \frac{M_1}{1-\delta_{12}}$, the destination 1 has received $M_1\frac{1-\delta_1}{1-\delta_{12}}$ packets while destination 2 observed $m_1 = M_1\frac{\delta_1-\delta_{12}}{1-\delta_{12}}$ overheard packets which are collected at the transmitter.

Phase 2: we let the sender convey M_2 packets intended for user 2 and switch the role of user 1 and 2. Similarly, we have the length of phase 2 as $N_2 = \frac{M_2}{1-\delta_{12}}$ and the source knows the $m_2 = M_2\frac{\delta_2-\delta_{12}}{1-\delta_{12}}$ overheard packets.

Phase 3: the transmitter communicates the random linear combinations of the $m_1 + m_2$ overheard packets in Q_1 and Q_2. The combination coefficients are generated in advance so that all the communications nodes are aware of them.

We remind the readers that the following analysis holds only when the transmission length is sufficiently large. Since the information at phase 3 is a combination of overhead packets and both users have partial knowledge of the overheard packets, each one only needs to decode the remaining packets which are also desired packets. To ensure they are decodable, the length of phase 3 should be at least $N_3 = \max\{\frac{m_1}{1-\delta_1}, \frac{m_2}{1-\delta_2}\}$. Then, we compute the rate for each user

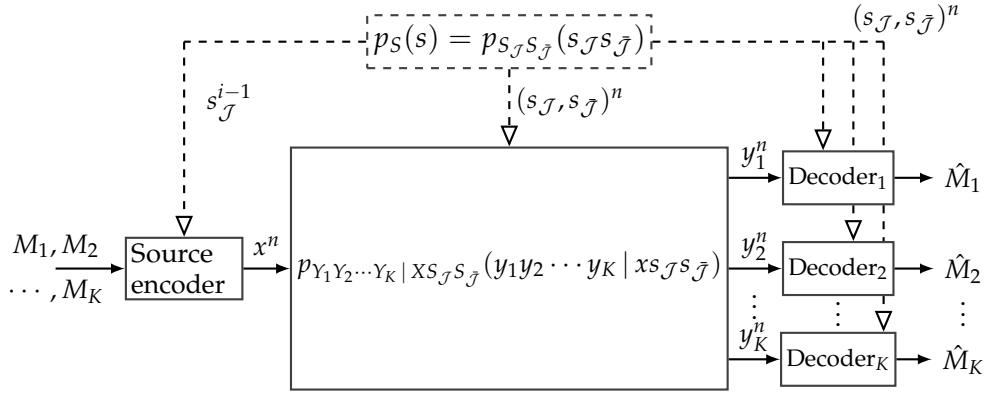

Figure 6: General system model of a multiple user BC with partial state feedback.

as,

$$R_1 = L\frac{M_1}{N_1+N_2+N_3} = L\frac{\lambda(1-\delta_{12})}{1+\max\{\lambda\frac{\delta_1-\delta_{12}}{1-\delta_1}, (1-\lambda)\frac{\delta_2-\delta_{12}}{1-\delta_2}\}},$$
$$R_2 = L\frac{M_2}{N_1+N_2+N_3} = L\frac{(1-\lambda)(1-\delta_{12})}{1+\max\{\lambda\frac{\delta_1-\delta_{12}}{1-\delta_1}, (1-\lambda)\frac{\delta_2-\delta_{12}}{1-\delta_2}\}},$$

where $\lambda = \frac{M_1}{M_1+M_2}$. We further simplify the rates and find out the corner point of (8a) and (8b)

$$(R_1, R_2) = \left(L\frac{(\delta_2-\delta_{12})(1-\delta_1)(1-\delta_{12})}{(1-\delta_{12})^2-(1-\delta_1)(1-\delta_2)}, L\frac{(\delta_1-\delta_{12})(1-\delta_2)(1-\delta_{12})}{(1-\delta_{12})^2-(1-\delta_1)(1-\delta_2)}\right)$$

is obtainable. Therefore, the vertex of capacity region is obtained while $(1-\delta_1, 0)$ and $(0, 1-\delta_2)$ are single user rates achieved by trivial strategy. The proof is then completed by using an argument on convexity of the region. ■

Remark 1.1. *The above capacity region is strictly larger than the no-feedback capacity where only single user rate is attainable.*

Remark 1.2. *The capacity results of EBC with delayed state at transmitter can be generalized to three user case and some subclass channels of EBC with K receivers by using finite field tools. In particular, these subclass channels are symmetric EBC and one-sided fair (to be defined in Chapter 3) spatially independent EBC when the channel input alphabet size 2^L is larger than K.*

1.3 BROADCAST CHANNEL WITH HETEROGENEOUS STATE FEEDBACK

We investigate the BC with heterogeneous state feedback in Chapter 4. The model we are interested in is similar to that defined in Section 1.2 in the sense that both systems consist of a single source, multiple destinations, and a BC via which K independent messages from the source are conveyed to the intended destination. The difference lies in the state information known to the source. In this channel,

we divide the state set into two parts as $S = (S_{\mathcal{J}}, S_{\bar{\mathcal{J}}})$, where $\bar{\mathcal{J}} \triangleq \mathcal{K} \setminus \mathcal{J}$ and $S_{\bar{\mathcal{J}}}$ is the complement of state set $S_{\mathcal{J}}$.

The state information are available *heterogeneously* to the distinct nodes. In particular, the destinations are aware of the global state information instantaneously. On the contrary, only *part* of the outdated CSI, namely $S_{\mathcal{J}}$, is learned by the transmitter, e.g., the users in set $\mathcal{J}$ feedback their states to the source. The detailed system model is depicted in Fig. 6.

In this channel, the encoding and decoding functions are modified as follows.

- A sequence of encoding functions assign $X_i(M_1, M_2, \cdots, M_K, S_{\mathcal{J}}^{i-1})$ to the messages and part of state sequence $S_{\mathcal{J}}^{i-1} \in {\mathcal{S}_{\mathcal{J}}}^{i-1}$ as shown,

$$\{\varphi_{s,i} : \mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_K \times {\mathcal{S}_{\mathcal{J}}}^{i-1} \to \mathcal{X}\}_{i=1}^{n}.$$

- Each of the K decoding functions assigns an index $\hat{M}_k$ based on the observed outputs $(y_k^n, s_{\mathcal{J}}^n, s_{\bar{\mathcal{J}}}^n)$.

$$\phi_k : \mathcal{Y}_k^n \times \left(\mathcal{S}_{\mathcal{J}}, \mathcal{S}_{\bar{\mathcal{J}}}\right)^n \to \mathcal{M}_k.$$

The rate tuple $(R_1, \ldots, R_K)$ is said to be achievable if there exists one series of encoding and decoding functions such that

$$\limsup_{n\to\infty} \max_k \mathbb{E}_{S_{\mathcal{J}}^n, S_{\bar{\mathcal{J}}}^n}\left[\Pr\left(\phi_k(Y_k^n, S_{\mathcal{J}}^n, S_{\bar{\mathcal{J}}}^n) \neq M_k\right)\right] \to 0.$$

Without loss of generality, we assume that the users in $\mathcal{J}$ who feedback states are indexed with $1, 2, \cdots, j$ and thus the rest who provides no state information to the source is indexed with $j+1, \cdots, K$. When $\mathcal{J} = \mathcal{K}$, it falls into the scenario studied in Chapter 3, i.e., homogeneous state feedback. For notational convenience, we identify the CSIT of various quality: the instantaneous/Perfect, Delayed, or Not known with P, D, N [11, 33]. Although our essential motif is to explore the impact of hybrid DN state feedback, PDN notations are still defined for comparison.

Definition 1.5. *In the K-user BC, the state feedback is said to be in the $D_jN_lP_{K-j-l}$ case if the delayed CSI of the users $1, \cdots, j$ are obtained by the transmitter via noiseless feedback link, while the users $j+1, \cdots, j+l$ does not inform the source of any state, whereas the source knows the perfect state of destinations $j+l+1, \cdots, K$.*[3]

1.3.1 *Linear coding in EBC with heterogeneous state feedback*

We show below the linear coding scheme of [44] in two user symmetric spatially independent EBC ($\delta_k = \delta$ and $\delta_{12} = \delta^2$). It is worth mentioning that the scheme is based on XOR operations on bits.

[3]Omitting the subscripts for the state status (P, D, N) is allowed for simplicity when no user falls into the corresponding category, e.g., $D_2N_0P_0$ is written as D_2.

Theorem 1.2 ([44, Theorem 3.3]). *For a two user EBC when $\delta_k = \delta$, $\delta_{12} = \delta^2$, and state feedback from a* single *user is available at the transmitter, an inner bound $\mathcal{R}_{LNC_1}$ is defined by the rate pair (R_1, R_2) satisfying,*

$$\frac{R_1}{\frac{1-\delta}{1-\delta^2+\delta^3}} + \frac{R_2}{1-\delta} \leq 1, \tag{9a}$$

$$\frac{R_1}{1-\delta} + \frac{R_2}{(1+\delta-\delta^2)(1-\delta)} \leq 1 \tag{9b}$$

Proof. Due again to the convex argument, we only need to prove the achievability of the corner point $(\frac{1-\delta}{1+\delta}, \frac{\delta(1-\delta)(1+\delta-\delta^2)}{1+\delta})$ for the sake of simplicity. We assume the transmission length is sufficiently large. The scheme is described below.

Phase 1: the transmitter attempts to send some uncoded bits wanted by user 1 through T_1 transmissions. Given the number of transmission tends to infinity, there will be $T_1\delta$ bits erased by destination 1 where we denote the erased sequence with $U_1^{\delta T_1}$.

Phase 2: the sender emits uncoded bits desired by user 2 during $T_2 = \frac{\delta T_1}{1-\delta}$ slots. Approximately, destination 1 observed $(1-\delta)T_2 = \delta T_1$ bits. Let us denote them with $V_1^{\delta T_1}$.

Phase 3: now that the number of bits erased at destination 1 in phase 1 equals to the bits learned at destination 1 in phase 2, the source can perform XOR between $U_1^{\delta T_1}$ and $V_1^{\delta T_1}$ bit-by-bit. Then, the XOR bits are sent in $T_3 = T_2 = \frac{\delta T_1}{1-\delta}$ transmissions which guarantees all the XOR bits are received by both users.

The destination 1 receives $(1-\delta)T_1$ uncoded bits during the course of phase 1 and δT_1 XOR bits in phase 3. With the observation on $V_1^{\delta T_1}$, the destination 1 is capable of recovering $U_1^{\delta T_1}$ and has T_1 bits in total. Hence, the rate R_1 equals to $\frac{T_1}{T_1+2\frac{\delta T_1}{1-\delta}} = \frac{1-\delta}{1+\delta}$. Receiver 2 obtains δT_1 uncoded bits in phase 2. Among δT_1 XOR bits acquired in third phase, only δ proportion is useful for user 2. By useful, we mean that those bits encoded into XOR bits are only helpful when destination 2 has not received them in phase 2. In fact, to decode these XOR bits, the receiver needs to know the corresponding U_1, among which only $(1-\delta)$ proportion is recovered at user 2. Therefore, there remain $(1-\delta)\delta^2 T_1$ such bits. In total, the receiver 2 get $(\delta+\delta^2-\delta^3)T_1$ bits which yields a rate of $R_2 = \frac{(\delta+\delta^2-\delta^3)T_1}{T_1+2\frac{\delta T_1}{1-\delta}} = \frac{\delta(1-\delta)(1+\delta-\delta^2)}{1+\delta}$. ∎

The opportunistic coding gain results from the fact that all the XOR bits are profitable for destination 1, which is assured by delayed state of user 1, and are partially useful to destination 2. In the third phase, the source codes in the hope that all the $U_1^{\delta T_1}$ bits are received by destination 2. Although this may not always be true, there should be some slots that conform to the assumption, which would be beneficial. As a matter of fact, such scheme relies heavily on the finite state

setting, as otherwise no slots will satisfy the assumption and the XOR would never be decodable at user 2.

Remark 1.3. *The sum rate of region's vertex (the corner point) is* $1-\delta+\frac{\delta^2(1-\delta)^2}{1+\delta}$ *which is bounded away from single user rate* $1-\delta$.

Remark 1.4. *The convex hull of two user EBC capacities subject to different physically degradedness can be identified as an outer bound for the EBC with heterogeneous feedback. However, it can be easily seen that the outer bound is not tight for this case. Unlike homogeneous feedback case where the channel forms a 'virtual' physically degraded channel, the feedback from one destination only assures the one-sided degradedness, which renders the outer bound by degraded argument loose.*

1.4 BROADCAST RELAY CHANNEL WITH STATE FEEDBACK AT RELAY

As several important tools were proposed in relay channel (RC), we discuss briefly the RC before going to the BRC. Although there have been remarkable progresses, since the first introduction of RC by van der Meulen [56] in 70s, in approaching the capacity for certain classes of relay channel (e.g., the degraded and reversely degraded relay channels, state-dependent semideterministic, orthogonal relay channels, etc.), the capacity of the general three-node relay network is still unknown (see [57–59] and reference therein). When the focus moves to a general multiple user relay network, the problem of characterizing the capacity remains widely open. Nevertheless, achievability results (without the proof of optimality), also known as inner bounds of the capacity region, for different relay networks have been well established based on classical relaying strategies ([57] by Cover and El Gamal) such as decode-and-forward (DF) and compress-and-forward (CF) (see, e.g., [60–62]). These coding strategies are of particular significance in that they have been widely extended and adapted in the network communication problems.

The BRC can be viewed as a combination of BC and RC whereas relay broadcast channel cannot. In fact, it is important to distinguish the BRC from relay broadcast channel initiated and studied in [63,64] to avoid confusion caused by close names. In relay broadcast channel, no infrastructure relay is deployed. Instead, direct communication between users is allowed, e.g., via device-to-device link. As such, each user could be considered as a helper for the other and functions like a real relay. In contrast, the relay in the proposed BRC model is not a conceptual relay and it subjects to its own constraint (rather than to the restriction of users). In addition, no direct link across destinations is available in BRC.

In the BRC, several works have been recorded, e.g., [60,61,65], where various schemes are extended from the fundamental strategies DF and CF in the conventional RC. Incorporating the DF and CF strategies as a hybrid scheme is not new either in the relay networks, such as, the noisy network coding scheme in [65]. The hybrid scheme lets part of the relay nodes compress what they observed but impose the others to decode completely the messages. If the network possesses a single relay, such scheme will reduce to the DF or CF strategies. This results from that normally one strategy dominates the other for certain given channel situations and thus a single strategy should be selected according to the channel

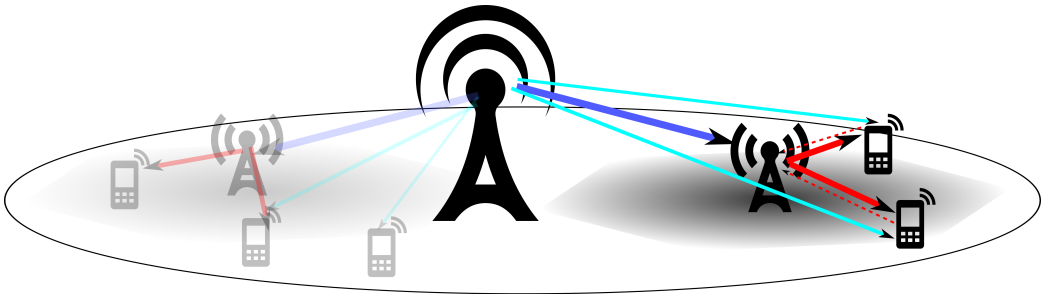

Figure 7: Broadcast relay channel in a downlink heterogeneous network.

condition at a single relay for the sake of efficiency. In Chapter 5, we will propose a counter-intuitive scheme which impel the relay to decode thoroughly the messages and then to compress the decoded information.

Recently, Han and Wang introduced in the work [66] a quite similar BRC model, however, with a stronger state condition than our setting, i.e., delayed CSI is obtained by both the source and the relay. The setting in [66] is more suitable for the scenario where sufficient resource are assigned to the feedback links. In particular, they proposed a linear network coding algorithm that is dedicated to the erasure case, which turns out to be close to the outer bound obtained by exhaustive search. In another noteworthy work [67], the authors considered a heterogeneous model consisting of two independent BC where one of the transmitters multicasts information and could as well be viewed as a relay. The essential difference between the aforementioned model and that in [67] lies in whether or not the relay could have an infinite-rate link from the source or to the users. It is shown in [67] that the completely outdated CSIT can achieve the same DoF as instantaneous CSIT, given that sufficiently high rate is available to the multicasting channel.

Some other related works are listed as follows. The impact of noisy CSI (statistic only/limited feedback) at BS/relay is investigated in [68, 69] on a downlink of orthogonal frequency-division multiple access (OFDMA) multi-relay cellular system in which a two-phase relaying transmission strategy is employed. In particular, it is observed that the average achievable rate performance with noisy CSIT and perfect CSIT are not comparable and yet noisy CSIT is promising within the practical scenario in the sense of exploiting relay-assisted diversity gain. In [70], a sum-rate degradation is reported in two-hop BRC that results from the low quality of CSI feedback with the assumption of quantized channel state feedback performed from users to the relay and then to the BS.

1.4.1 *Modelization from practical wireless system*

In this part, we describe a real system model which is attached to the BRC. This model leads to some limitations on the BRC setup for Chapter 5 We consider a single-cell downlink heterogeneous channel setup with one macro BS (source) and two mobile users (destinations) who are helped by a smallcell BS (relay), as shown in Fig. 7. We assume that all the nodes share the same spectrum and the relay

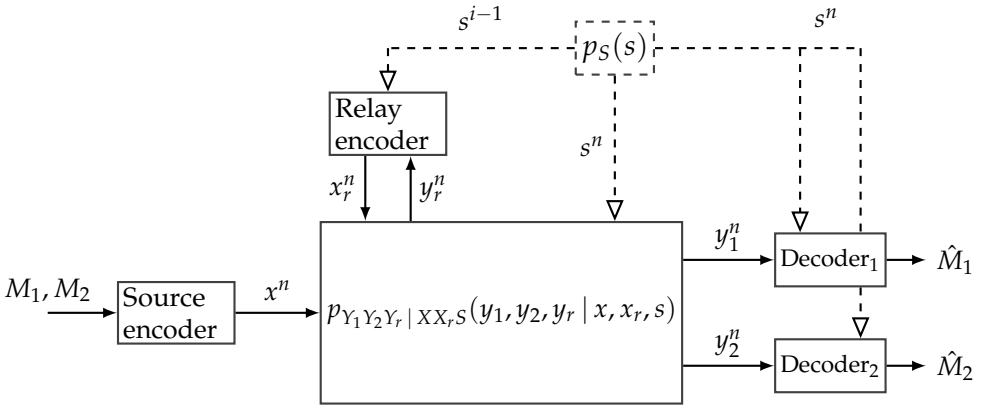

Figure 8: General system model of a two-user BRC with heterogeneous state feedback.

works in full-duplex mode, i.e., it can transmit and receive simultaneously. As such, the interference between the source signal and the relay signal, in addition to the interference generated by broadcasting, is as well a main limitation to the system. In Chapter 5, we will focus on the fast fading scenario, i.e., the realization of the channel state at each time slot varies randomly and independently according to an identical distribution.[4] Furthermore, the instantaneous state is unknown at the macro/small-cell BSs while it can be learned perfectly at mobile ends.

The particularity of our model lies in the existence of an independent feedback link from the destinations to the relay, instead of to the source. This is motivated by the geographical proximity of the relay to the destinations. At the end of each time slot, the destinations send the corresponding state information to the relay using this local communication link.

To analyze such a system in the heterogeneous settings, in Chapter 5, we resort to an information-theoretic framework in which a general class of memoryless BRCs are defined in the following. Based on the results obtained in such a framework, we will then focus on the Gaussian noise channel with fast fading.

1.4.2 *General memoryless two-user BRC*

The considered downlink channel belongs to a broader class of memoryless two-user BRC, as shown in Fig. 8, which is defined by the joint pmf

$$p(y_1^n, y_2^n, y_r^n | x^n, x_r^n, s^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}, y_{ri} | x_i, x_{ri}, s_i) \tag{10}$$

where x^n and x_r^n denote the sequence of n input symbols at the source and the relay, respectively; y_1^n, y_2^n, and y_r^n stand for the channel outputs at receiver 1, receiver 2, and the intermediate relay, respectively; s^n defines the channel state

[4]Different fading assumptions such as the block fading can also be considered, although the interplay between the feedback delay and the coherence time should be taken into account in such settings. Here, we consider the i.i.d. fast fading for simplicity.

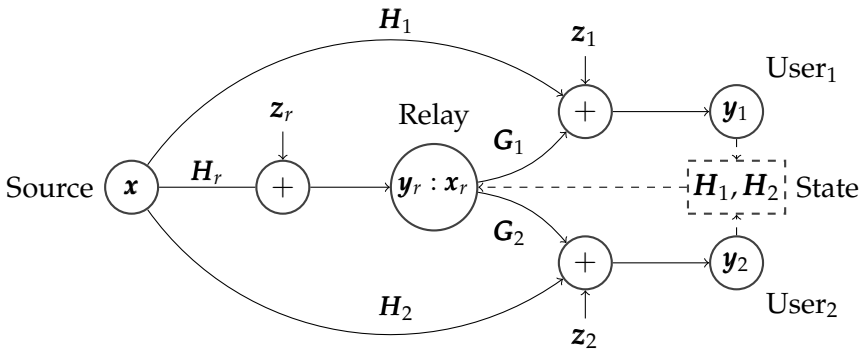

Figure 9: The fading Gaussian two-user BRC model.

which determines the channel behaviour given certain inputs x^n and x_r^n. As shown in (10), the channel is temporally i.i.d. and thus the outputs are decided exclusively by the current channel inputs and the state. At slot i, only the state of previous slots s^{i-1} is available to the relay while instantaneous state s^i is known to the decoder upon reception.

The source encoder maps each pair of messages M_1 and M_2 into a sequence of input symbols x^n, where $x \in \mathcal{X}$;[5] the relay encoder maps its past observations (y_r^{i-1}, s^{i-1}), from the source and from the destinations through feedback link, to a symbol $x_{r,i}$ at each time slot i, where $y_r^{i-1} \in \mathcal{Y}_r^{i-1}$, $s^{i-1} \in \mathcal{S}^{i-1}$, and $x_{r,i} \in \mathcal{X}_r$; each destination k decodes its own message, $\hat{M}_k$, from the received signal $y_k^n \in \mathcal{Y}_k^n$ and the state sequence s^n, for $k = 1, 2$. The encoders and decoders are defined as follows.

- Source encoder: $\varphi_s : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}^n$;
- Relay encoder: $\{\varphi_{r,i} : \mathcal{Y}_r^{i-1} \times \mathcal{S}^{i-1} \to \mathcal{X}_r\}_{i=1}^n$;
- Decoder k: $\phi_k : \mathcal{Y}_k^n \times \mathcal{S}^n \to \mathcal{M}_k$.

A rate pair (R_1, R_2) is considered as achievable if there exists a series of such encoding/decoding functions such that the maximal probability of error in decoding either message goes to zero with $n \to \infty$, as defines in Section 1.2.

1.4.3 *Fading Gaussian noise BRC*

As a special case of the memoryless BRC, a Gaussian noise BRC is defined as following, which is as well shown in Fig. 9,

$$\boldsymbol{y}_1[i] = \boldsymbol{H}_1[i]\boldsymbol{x}[i] + \boldsymbol{G}_1[i]\boldsymbol{x}_r[i] + \boldsymbol{z}_1[i], \tag{11a}$$
$$\boldsymbol{y}_2[i] = \boldsymbol{H}_2[i]\boldsymbol{x}[i] + \boldsymbol{G}_2[i]\boldsymbol{x}_r[i] + \boldsymbol{z}_2[i], \tag{11b}$$
$$\boldsymbol{y}_r[i] = \boldsymbol{H}_r[i]\boldsymbol{x}[i] + 0 \qquad + \boldsymbol{z}_r[i], \tag{11c}$$

[5]The specification of the cardinality of the channel input set should be adapted to distinct channel, e.g., infinite in the analog channel, binary in the binary erasure/symmetric channel, etc.

for time slot $i = 1, 2, \ldots, n$, where $\boldsymbol{x}[i] \in \mathbb{C}^{n_{ts}\times 1}$, $\boldsymbol{x}_r[i] \in \mathbb{C}^{n_{tr}\times 1}$ are inputs at the source and at the relay, subject to the average power constraints $\frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{x}[i]\|^2 \leq P_s$ and $\frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{x}_r[i]\|^2 \leq P_r$, respectively; $\boldsymbol{y}_1 \in \mathbb{C}^{n_{r1}\times 1}$, $\boldsymbol{y}_2 \in \mathbb{C}^{n_{r2}\times 1}$ are the channel outputs at the respective user, while $\boldsymbol{y}_r[i] \in \mathbb{C}^{n_{rr}\times 1}$ stands for the relay output y_r^n; $\boldsymbol{z}_1, \boldsymbol{z}_2$ and $\boldsymbol{z}_r$ are the AWGNs with temporarily and spatially i.i.d. entries $\mathcal{CN}(0, 1)$. We assume that the relay channel input is cancelled when observed at the relay's side, which is expressed as 0 in (11c).

The set of all channel matrices $\boldsymbol{H}_1[i] \in \mathbb{C}^{n_{r1}\times n_{ts}}$, $\boldsymbol{H}_2[i] \in \mathbb{C}^{n_{r2}\times n_{ts}}$, $\boldsymbol{H}_r[i] \in \mathbb{C}^{n_{rr}\times n_{ts}}$, $\boldsymbol{G}_1[i] \in \mathbb{C}^{n_{r1}\times n_{tr}}$, and $\boldsymbol{G}_2[i] \in \mathbb{C}^{n_{r2}\times n_{tr}}$ are identified with the channel state $s_i = (\boldsymbol{H}_1[i], \boldsymbol{H}_2[i], \boldsymbol{H}_r[i], \boldsymbol{G}_1[i], \boldsymbol{G}_2[i])$ and are temporarily i.i.d. channel matrix processes. In particular, $\boldsymbol{H}_1$, $\boldsymbol{H}_2$, and $\boldsymbol{H}_r$ are the channel matrices of SDC to the user 1, SDC to the user 2, and the source-relay channel (SRC), respectively. Similarly, $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ are matrices related to the corresponding relay-destination channels (RDCs). Specifically, each receiver is aware of all the channel matrices, $\boldsymbol{H}_1[i], \boldsymbol{H}_2[i], \boldsymbol{H}_r[i], \boldsymbol{G}_1[i], \boldsymbol{G}_2[i]$, while no channel coefficients are known to the source; to the relay end, only delayed channel matrices of SDC, $\boldsymbol{H}_1[i-1], \boldsymbol{H}_2[i-1]$, and the instantaneous channel information of SRC, $\boldsymbol{H}_r[i]$, are available. It is worth mentioning that the CSI of RDC is not obtained by the relay. Such assumption is realistic provided that the relay in the real-life system are only allowed to ask destinations to feedback limited information.

Note that n_{ts} and n_{tr} denote the transmit antenna number at the source and the relay, respectively; n_{rr}, n_{r1}, and n_{r2} are the number of the receive antennas at the relay, user 1, and user 2, respectively. Therefore, the joint distribution of the channel $p(y_1, y_2, y_r|x, x_r, s)$ in (10) can be specified as

$$\begin{aligned} &p_Z(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{x} - \boldsymbol{G}_1\boldsymbol{x}_r)p_Z(\boldsymbol{y}_2 - \boldsymbol{H}_2\boldsymbol{x} - \boldsymbol{G}_2\boldsymbol{x}_r)p_Z(\boldsymbol{y}_r - \boldsymbol{H}_r\boldsymbol{x}_r) \\ &= \frac{1}{\pi'} e^{-\|\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{x} - \boldsymbol{G}_1\boldsymbol{x}_r\|^2 - \|\boldsymbol{y}_2 - \boldsymbol{H}_2\boldsymbol{x} - \boldsymbol{G}_2\boldsymbol{x}_r\|^2 - \|\boldsymbol{y}_r - \boldsymbol{H}_r\boldsymbol{x}_r\|^2}, \end{aligned} \tag{12}$$

where $\pi' = \pi^{n_{rr}+n_{r1}+n_{r2}}$.

1.5 THESIS OUTLINE

We present the layout of this dissertation in the sequel. In Chapter 1, we stated the motivation of our research, described the problems and models we investigated, made a brief survey on some background results for the mentioned problems, and will present the list of publications in the next section.

The remaining of the thesis consists of three parts as clarified below.

The first part comprises Chapter 2 and 3, which address the interference mitigation problem in multi-user BC with feedback. In particular, the feedback is limited to homogeneous state feedback. In Chapter 2 [38], we present the inner and outer bounds for the two user general case and specify the two in Gaussian channel. Based on the numerical results, we assess the performance of delayed CSIT at finite SNR. In Chapter 3 [39, 40], we provide a unified approach in K user BC and derive the corresponding inner bound. The inner bound is then evaluated in EBC and in GBC. Simulation results are analyzed, which give us some insight

on how the employed coding tools could help exploit the state feedback at finite SNR.

The second part is formulated by Chapter 4 and 5. Both chapters deal with broadcasting problem with heterogeneous state feedback where each one studies different feedback heterogeneity, that is, heterogeneity at the transmitters' and receivers' sides, respectively. In Chapter 4 [71], we focus on two-user state-dependent BC with state feedback from part of the receivers. Several inner bounds by distinct non-linear methods are introduced and are then specified into EBC, in which comparisons with existing inner bounds in the literature are made. The analyses allow us to understand where the benefits of exploiting non-linear coding lie in and to know whether adding a user who does not feedback state will be helpful. In Chapter 5 [17,72], we insert a relay to the broadcast channel and impose a limitation that the state of both users is accessible only to the relay. We develop an inner bound whose specification in Gaussian channel is divided into several variants and modes according to how the random variables (RVs) are chosen. Along with some schemes extended from conventional relay strategies, the proposed scheme is assessed numerically.

In the third part (Chapter 6), the main outcomes of this thesis are revisited, accompanied by some general comments and comparisons across the chapters. Finally, we close the dissertation with a perspective section that suggests some interesting but unexplored directions (due to the time constraint on a Ph.D. program).

Appendices are attached at the end of each chapter.

1.6 LIST OF PUBLICATIONS

The publications during the course of this PhD are enlisted below.

Journal Articles

[71] C. He, S. Yang, and P. Piantanida, "Exploiting heterogeneous delayed CSIT in broadcast channel: linear or non-linear coding?", *in preparation*.

[40] C. He, S. Yang, and P. Piantanida, "A Non-Linear Transmission Strategy for the *K*-user BC with Delayed CSIT," *submitted to IEEE Transactions on Wireless Communications*, 2016.

[72] C. He, S. Yang, and P. Piantanida, "On the Gaussian Fading Broadcast Relay Channel with Causal State Feedback," in *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2797-2807, 2016.

Conference Papers

[39] C. He, S. Yang, and P. Piantanida, "A Non-Linear Transmission Scheme for the *K*-User Broadcast Channel with State Feedback," in *International Symposium on Information Theory and Its Applications (ISITA), to be presented*, Nov. 2016.

[17] C. He, S. Yang, and P. Piantanida, "An Achievable Rate Region of Broadcast Relay Channel with State Feedback," in *Proceedings of the IEEE International Conference on Communications (ICC)*, June. 2015, pp. 4193–4198.

[38] C. He, S. Yang, and P. Piantanida, "On the Capacity of the Fading Broadcast Channel with State Feedback," in *IEEE International Symposium on Communications, Control and Signal Processing (ISCCSP), invited*, May. 2014, pp. 182-185.

Part I

Broadcasting with Homogeneous State Feedback

2

An achievable rate region of the two-user BC with State Feedback

2.1 OVERVIEW

We recall that lacking effective algorithms on delayed CSIT for finite SNR spurs the following work. This chapter is a short chapter in which, we investigate the capacity problem in two-user state-dependent memoryless BC. The general and fading Gaussian BC models studied here can be specified from the models defined in Section 1.2 by setting $K = 2$. Since the capacity result of two user EBC has been long proved, we focus on GBC in this chapter. The main outcomes of this chapter are listed below.

- We first derive simple inner and outer bounds on the capacity of the general two-user memoryless BC with state feedback. The inner bound is based on a block-Markov scheme [73] used to communicate to both users a noisy function of all transmit codewords and state feedback from the previous block while the outer bound is derived from the standard genie-aided approach [4].
- We then specialize these rate regions to the case of the fading Gaussian MISO-BC. Moreover, we tighten the outer bound by exploiting the fact that it only depends on the marginal distribution of the outputs, which allows us to minimize over the joint probability distribution of noises. Finally, sum-rate performances are evaluated through numerical simulation.

The rest of the chapter is organized as: in Section 2.2, we present two regions: an inner bound achieved by the corresponding block Markov scheme explained in appendix and an outer bound by constructing a degraded BC. In Section 2.3, we specify both bounds in this model followed by some numerical evaluations. Finally, Section 2.4 concludes the chapter.

2.2 MAIN RESULTS

2.2.1 *Achievable rate region*

A non-linear scheme mainly built on block Markov coding and Wyner-Ziv compression techniques is proposed with the respective achievable rate region (inner

bound) shown in Theorem 2.1. With the help of block Markov coding, the transmitter is capable of broadcasting the novel independent messages and multicasting some useful side information simultaneously. During each transmission, message could be overheard by undesired user and thus is considered as "interference". In this scheme, the source encodes the messages with the state feedback to re-construct the corresponding observations at unintended decoders, compresses the encoded sequences and maps it into a signal according to the compression index (multicast message) and the new messages. Since each user has certain amount of the "interference", the cost of recovering the side information becomes relatively low and thus the rate region is enlarged as compared to the TDMA region where no interference is exploited.[1]

Theorem 2.1 (Inner bound). *An achievable rate region for the memoryless state-dependent BC with state feedback is given by the set of all rate pairs* (R_1, R_2) *satisfying:*

$$R_1 \leq \min\left\{I(V_1; Y_1\hat{Y} \mid V_{12}SQ), I(V_1V_{12}; Y_1|SQ) - I(\hat{Y}; V_2X|V_1Y_1V_{12}SQ)\right\}, \tag{13a}$$

$$R_2 \leq \min\left\{I(V_2; Y_2\hat{Y} \mid V_{12}SQ), I(V_2V_{12}; Y_2|SQ) - I(\hat{Y}; V_1X|V_2Y_2V_{12}SQ)\right\}, \tag{13b}$$

for all probability distributions (PD) that verify

$$\begin{aligned} p(y_1, y_2, x, v_{12}, v_1, v_2, s, q) = p(y_1, y_2|x, s)p(\hat{y} \mid x, v_{12}, v_1, v_2, s) \\ \times p(x|v_{12}, v_1, v_2, q)p(v_1|q)p(v_2|q)p(v_{12}|q)p(s)p(q). \end{aligned} \tag{14}$$

Sketch of proof. The V_k signal carries the independent message M_k, $\hat{Y}$ is the *common* side information available at each destination while V_{12} is linked with side-information bearing message. The condition $R_k \leq I(V_k; Y_k\hat{Y} \mid V_{12}SQ)$ determines the rate that the user k can decode the intended message at the presence of side information if available. The side information is communicated without error to the user k as long as $R_k \leq I(V_kV_{12}; Y_k|SQ) - I(\hat{Y}; V_mX|V_kY_kV_{12}SQ)$ is satisfied for $m \neq k$ and $m \in \{1, 2\}$. The detailed coding scheme, which follows the approach in [73], is provided in the Appendix 2.A. To better distinguish the scheme from the other algorithms proposed in the thesis, we refer to this scheme as block Markov scheme (for two users). ■

Remark 2.1. *Note that the above rate region is only defined by the single rate bounds and no sum rate constraint is involved. This arises from the fact that we exclude Marton's coding and rate splitting in the scheme. These two powerful but complex tools are removed as otherwise it will yield a less tractable region in the concerned Gaussian channel.*

Remark 2.2. *Inner bound regions for a more general feedback setting have previously been reported in [20, 22]. As a matter of fact, a direct comparison between the inner bound in Theorem 2.1 and those regions for general pmf does not appear to be straightforward. As it was stated in [22], it is not trivial to compare their region with a region that employs the backward decoding technique to exploit the side information, which is the case for our*

[1] By setting $\hat{Y} = V_{12} = 0$, the region in Theorem 2.1 attains the TDMA rate.

block Markov scheme. Moreover, it seems difficult as well to determine whether SW region [20] includes our region in general.

We compare a specific case of the SW region with the proposed region in the sequel. In particular, we first generalize the SW region to the version with channel state, implement a coded time sharing with Q, and then abandon the binning in Marton's coding (let the signals intended for each user be independent, i.e., $I(U_1; U_2|U_0SQ) = 0$). The simplified SW region on $(R_1,\ R_2)$ defined by the rate constraints (15a), (15b), and (15c) is obtained if we adapt the RVs in [20], say, let U_0, U_k, V_0, V_1, and V_2 in [20] equal to V_{12}, V_k, $\hat{Y}$, $\hat{Y}$, and $\hat{Y}$, respectively. Such adaptation conforms to the design of SW scheme: U_0 is the update information based on the feedback and wanted by both users, U_k denotes the signal that bears independent message for user k, $V_0 = V_1 = V_2 = \hat{Y}$ represents the side information.

$$R_1 \leq I(V_1; Y_1\hat{Y}|V_{12}SQ) + I(V_{12}; Y_1\hat{Y}|SQ) - I(\hat{Y}; V_1V_2V_{12}|Y_1SQ), \tag{15a}$$

$$R_2 \leq I(V_2; Y_2\hat{Y}|V_{12}SQ) + I(V_{12}; Y_2\hat{Y}|SQ) - I(\hat{Y}; V_1V_2V_{12}|Y_2SQ), \tag{15b}$$

$$\begin{aligned} R_1 + R_2 \leq\ & I(V_1; Y_1\hat{Y}|V_{12}SQ) + I(V_2; Y_2\hat{Y}|V_{12}SQ) \\ & + \min_{i\in\{1,2\}} I(V_{12}; Y_i\hat{Y}|SQ) - \max_{j\in\{1,2\}} I(\hat{Y}; V_1V_2V_{12}|Y_jSQ). \end{aligned} \tag{15c}$$

The proposed region can be rewritten as

$$R_k \leq I(V_k; Y_k\hat{Y} \mid V_{12}SQ) + I(V_{12}; Y_k\hat{Y}|SQ) - I(\hat{Y}; V_1V_2V_{12}|Y_kSQ), \tag{16a}$$

$$R_m \leq I(V_m; Y_m\hat{Y} \mid V_{12}SQ). \tag{16b}$$

where $k, m \in \{1, 2\}$ and $k \neq m$. From the above regions, we notice two facts: 1) (16a) coincides with the conditions (15a) and (15b); 2) Summing up (16a) and (16b) for different k yields two conditions for SW region, namely, the constraints (15c) when $i = j \in \{1, 2\}$. In fact, the SW scheme performs a separate source and channel coding in a multiple-block framework: the Lossy Gray Wyner encoding (source coding for auxiliary refinement messages) and Marton's coding (channel coding for original messages and refinement messages) are applied sequentially at the transmitter's side; at each destination, the Gray Wyner decoder is used to retrieve side information given some relevant descriptions, which is followed by a Marton's decoder that exploits the side information to find the desired messages and to decode compression description of the side information for previous blocks. The condition $\max_{j\in\{1,2\}} I(\hat{Y}; V_1V_2V_{12}|Y_jSQ) = I(\hat{Y}; V_1V_2V_{12}|SQ) - \min_{j\in\{1,2\}} I(\hat{Y}; Y_j|SQ)$ is given by the source encoding/decoding that hinges on Gray Wyner coding, where the rate of side information is determined as the minimum encoding rate such that source codeword can be reconstructed at both destinations. On the other hand, $\min_{i\in\{1,2\}} I(V_{12}; Y_i\hat{Y}|SQ)$ is provided by the analysis on channel encoding/decoding that encompasses Marton's coding in conventional block Markov structure, which guarantees the channel coding codeword is decodable at two receivers. Such separation on the coding leads to the mismatched conditions ($i \neq j$ in (15c)), where the respective conditions cannot be found in our region. Therefore, it is not clear whether one region includes the other.

It can be, however, easily shown that the *symmetric* rate with SW scheme is not

smaller than our rate in *symmetric* channel. The SW symmetric rate is defined by

$$R_{sym} \leq I(V_k; Y_k\hat{Y}|V_{12}SQ) + \min\{\frac{1}{2}I_{0,k}, I_{0,k}\}. \tag{17}$$

where $I_{0,k} \triangleq I(V_{12}; Y_k\hat{Y}|SQ) - I(\hat{Y}; V_1V_2V_{12}|Y_kSQ)$. The symmetric rate by the proposed scheme can be written as

$$R_{sym} \leq I(V_k; Y_k\hat{Y} \,|\, V_{12}SQ) + \min\{0, I_{0,k}\}, \tag{18}$$

Then, we discuss as: 1) when $I_{0,k} \geq 0$, our region is dominated by the condition $I(V_k; Y_k\hat{Y} \,|\, V_{12}SQ)$ which is strictly smaller than SW's symmetric rate; 2) when $I_{0,k} < 0$, both SW and our rates are limited by the same condition $I(V_k; Y_k\hat{Y}|V_{12}SQ) + I_{0,k}$. Due to the time limitation, we will leave the comparison between our region and the region introduced in [21] for future given that their region could be even more general than SW region.

Remark 2.3. *With a proper design of the auxiliary RVs, the proposed region coincides with the capacity region of erasure BC with delayed CSIT, which is given in* (8)*. In a two-user erasure BC, the capacity is achievable if we set*

$$X = \begin{cases} V_{12}, & Q = 0 \ w.p. \ \alpha_0, \\ V_1, & Q = 1 \ w.p. \ \alpha_1, \\ V_2, & Q = 2 \ w.p. \ \alpha_2, \end{cases}$$

$$\hat{Y} = \begin{cases} V_1, & Q = 1, Y_2 = X, Y_1 = E, \\ V_2, & Q = 2, Y_1 = X, Y_2 = E, \\ 0, & otherwise. \end{cases}$$

where $\alpha_0 + \alpha_1 + \alpha_2 = 1$*. The converse is provided in [27, 28].*

2.2.2 *Outer bound*

Theorem 2.2 (Outer bound)**.** *Any achievable rate pair* (R_1, R_2) *for the broadcast channel with state feedback must satisfy*

$$R_1 \leq \min\{I(U_1; Y_1 \,|\, S), \ I(X; Y_1, Y_2 \,|\, U_2, S)\} \tag{19a}$$
$$R_2 \leq \min\{I(U_2; Y_2 \,|\, S), \ I(X; Y_1, Y_2 \,|\, U_1, S)\} \tag{19b}$$

for some PD $p(y_1, y_2, s \,|\, x)p(x, u_1, u_2)$*.*

Sketch of proof. In order to prove the outer bound, we establish a new *physically degraded* BC by providing user 1's output y_1 to the user 2 so that the observation $Y_1^{'}$ at user 1, i.e., $Y_1^{'} = (Y_1, Y_2)$, is a degraded version of the information Y_2 at the user 2. As a matter of fact, we can ignore the state feedback since feedback does not enlarge the capacity region of a physically degraded BC [4]. The capacity region of the physically degraded BC without feedback is achievable with a simple superposition coding [74]. Therefore, the capacity region of this physically

degraded BC without feedback serves as an outer bound for the original BC with feedback as defined in previous Chapter. With a similar reasoning, we can obtain another outer bound of the interested BC by letting user 2 give its observation to user 1 for free and ignore the feedback. Since both outer bounds for the BC with state feedback should hold simultaneously, a tighter outer bound as the intersection of the two outer bounds is obtained. Finally, the outer bound in Theorem 2.2 can then be achieved with standard procedures by taking into account the memoryless nature of the channel [4]. ■

2.3 GAUSSIAN MISO BC AND NUMERICAL RESULTS

In the sequel, we consider the two-antenna Gaussian MISO BC with fast fading. We let the variance of noise entry be $\sigma^2 = 1$ and thus $\mathsf{SNR} = \frac{P}{2}$. The channel matrices $\boldsymbol{H}_k$ becomes channel vectors $\boldsymbol{h}_k^T$. For brevity, we define $\boldsymbol{H} = [\boldsymbol{h}_1\ \boldsymbol{h}_2]^T$.

2.3.1 *Inner bound in Gaussian case*

We let $\tilde{X}$ be $\boldsymbol{v}_1 + \boldsymbol{v}_2$, X be $\boldsymbol{v}_0 + \boldsymbol{v}_1 + \boldsymbol{v}_2$, and finally $\hat{Y}$ be $(\boldsymbol{h}_1^T\boldsymbol{v}_2 + b_1, \boldsymbol{h}_2^T\boldsymbol{v}_1 + b_2)$. Furthermore, we choose, for simplicity, $\boldsymbol{v}_j \sim \mathcal{CN}(0, P_j\mathbf{I}_2)$, $j = 0, 1, 2$, with $P_0 + P_1 + P_2 = \frac{P}{2}$; and the compression noise $(b_1, b_2) \sim \mathcal{CN}(0, N\mathbf{I}_2)$. The power constraints are met asymptotically with the above RV choices as the block length of the transmission tends to infinity. We follow the same principles of the achievability proofs of Theorem 9.1.1 in [75] and of Theorem 3 in [74] to adapt the block Markov scheme/region intended for discrete case to the case where the RVs are continuous. For instance, we quantize the continuous RV V into $[V]_j \in \{-j\Delta, -(j-1)\Delta, \cdots, -\Delta, 0, \Delta, \cdots, (j-1)\Delta, j\Delta\}$, $\Delta = \frac{1}{\sqrt{j}}$, for every $j = 1, 2, \cdots$. Thus, we have $2j + 1$ levels of quantized value. The quantization is obtained by mapping V to the closest quantization point such that $|[V]_j| \leq |V|$. For a given j and a given tuple of $[V]_j$'s, $Y_{k,j}$ is the corresponding output at receiver k given the quantized input and auxiliary RVs $[V]_j$'s. Then, we quantize the $Y_{k,j}$ into $[Y_{k,j}]_l$. One can show that, when j's and l's tend to infinity, all the mutual information in the region in Theorem 2.1 are arbitrarily close to their counterpart if we substitute the $V's, X, Y's$ with their quantized version. Therefore, the power constraint is always fulfilled. We have

$$I(V_1; \hat{Y}Y_1 \mid V_{12}SQ) = \mathbb{E}_{\boldsymbol{H}} \log\det\left(\mathbf{I} + P_1 \begin{bmatrix} (1+\frac{\|\boldsymbol{h}_1\|^2 P_2 N}{\|\boldsymbol{h}_1\|^2 P_2 + N})^{-1} & \\ & N^{-1} \end{bmatrix} \boldsymbol{H}\boldsymbol{H}^H\right), \quad (20a)$$

$$I(V_1V_{12}; Y_1 \mid SQ) = \mathbb{E}_{\boldsymbol{h}_1} \log\det(1 + \frac{P}{2}\|\boldsymbol{h}_1\|^2) - \mathbb{E}_{\boldsymbol{h}_1} \log(1 + P_2\|\boldsymbol{h}_1\|^2), \quad (20b)$$

$$I(\hat{Y}; V_2X \mid Y_1V_{12}V_1SQ) = \mathbb{E}_{\boldsymbol{h}_1} \log\left(1 + \frac{N^{-1}P_2\|\boldsymbol{h}_1\|^2}{1 + P_2\|\boldsymbol{h}_1\|^2}\right). \quad (20c)$$

Corollary 2.3 (Inner bound). *A rate pair (R_1, R_2) for the Gaussian BC with*

state feedback is achievable if

$$R_1 \le \min\left\{ C_{h_1}(\tfrac{1}{2}P) - C_{h_1}((1+N^{-1})P_2), C_H\Big(P_1 \begin{bmatrix} (1+\frac{\|\boldsymbol{h}_1\|^2 P_2 N}{\|\boldsymbol{h}_1\|^2 P_2 + N})^{-1} & \\ & N^{-1} \end{bmatrix}\Big)\right\},$$
$$R_2 \le \min\left\{ C_{h_2}(\tfrac{1}{2}P) - C_{h_2}((1+N^{-1})P_1), C_H\Big(P_2 \begin{bmatrix} (1+\frac{\|\boldsymbol{h}_2\|^2 P_1 N}{\|\boldsymbol{h}_2\|^2 P_1 + N})^{-1} & \\ & N^{-1} \end{bmatrix}\Big)\right\},$$

for some $P_1, P_2, N \ge 0$, $P_1 + P_2 \le \frac{P}{2}$, $N < \infty$, *where we define*

$$C_H(\boldsymbol{M}) \triangleq \mathbb{E}_H \log\det(\mathbf{I} + \boldsymbol{M}\boldsymbol{H}\boldsymbol{H}^H), \tag{22a}$$
$$C_{h_k}(m) \triangleq \mathbb{E}_{h_k} \log(1 + m\|\boldsymbol{h}_k\|^2), \quad k = 1, 2. \tag{22b}$$

Remark 2.4. *Though the general optimal choice of* $(P_k,\ N)$ *is unknown, the parameters that attain the optimal sum DoF* $\frac{4}{3}$ *are readily seen from above region, as* $P_k = P^{\frac{1}{3}}$ *and* $N = O(P)$.

2.3.2 *Outer bound in Gaussian case*

Corollary 2.4 (Outer bound in Gaussian BC)**.** *Any achievable rate pair* (R_1, R_2) *for the Gaussian BC with state feedback must satisfy*

$$\mu_1 R_1 + R_2 \le r_1(\mu_1), \tag{23a}$$
$$\mu_2 R_2 + R_1 \le r_2(\mu_2), \tag{23b}$$

for any $\mu_1, \mu_2 \ge 1$ *with*

$$r_k(\mu) \triangleq \max_{\boldsymbol{Q}' : tr(\boldsymbol{Q}') \le P} \mathbb{E}_{\boldsymbol{H}} \log \frac{\det\big(\mathbf{I} + \boldsymbol{H}\boldsymbol{Q}'\boldsymbol{H}^H\big)}{(1 + \boldsymbol{h}_k^T \boldsymbol{Q}' \boldsymbol{h}_k^*)^{\mu}} + \mu\, \mathbb{E}_{\boldsymbol{h}_k} \log\big(1 + \frac{P}{2}\|\boldsymbol{h}_k\|^2\big) \tag{24}$$

where $\boldsymbol{h}_k^T$ *is* k*th row of the matrix* $\boldsymbol{H}$.

Sketch of proof. First, from Theorem 2.2, we have

$$\mu_1 R_1 + R_2 \le I(X; Y_1, Y_2 \,|\, U_1, S) + \mu_1 I(U_1; Y_1 \,|\, S) \tag{25a}$$
$$= h(Y_1, Y_2 \,|\, U_1, S) - \mu_1 h(Y_1 \,|\, U_1, S) + \mu_1 h(Y_1 \,|\, S) - h(Y_1, Y_2 \,|\, U_1, S, X). \tag{25b}$$

As adapted in the vector BC, we assume $Cov(X) = \boldsymbol{Q}$ and $Cov(X|U_1) \preceq \boldsymbol{Q}'$, where the invocation of the latter is related to the optimization of first two terms in (25b). Then, we obtain (23a) using the facts that

1. As the optimal value of $p(x|u_1)$ on the first line of (25b) is a concave function of $\boldsymbol{Q}'$, $h(Y_1, Y_2 \,|\, U_1, S) - \mu_1 h(Y_1 \,|\, U_1, S)$ is maximized when the conditional input fulfills $X \,|\, U_1 \sim \mathcal{CN}(0, \boldsymbol{Q}')$ for some $\boldsymbol{Q}'$ and any real number $\mu_1 \ge 1$ due to the extremal inequality in [76, Theorem 8];

2. $h(Y_1 \mid S)$ is maximized when $X \sim \mathcal{CN}(0, \boldsymbol{Q})$, and that $\boldsymbol{Q} \preceq P\mathbf{I}_2$ if $\boldsymbol{Q} \succeq 0$ with $\mathrm{tr}(\boldsymbol{Q}) \leq P$ with standard entropy maximization tool, see [75, Chapter 12];

3. $h(Y_1, Y_2 \mid U_1, S, X) = 2\log(\pi e)$.

Due to the symmetry, (23b) holds with the same reasoning. ■

Corollary 2.5 (Sum-rate upper bound in symmetric case)**.** *For a symmetric Gaussian BC with state feedback, the sum-rate is upper-bounded by*

$$R_1 + R_2 \leq \min_{\mu \in [1,2]} \frac{2\, r(\mu)}{1+\mu}. \tag{26}$$

where $r \triangleq r_1 = r_2$ as defined in (24) *due to the symmetry.*

Proof. By setting $\mu_1 = \mu_2 = \mu$ in (23a) and (23b), we obtain the objective function of the above minimization. By assumption, we have $\mu \geq 1$. It remains to show that the optimal value of μ must not be larger than 2. Let us suppose that $\mu > 2$, then we can show that the first term in (24) within the maximization is always upper-bounded by 0, as

$$\begin{aligned} h(Y_1, Y_2 \mid U_1, S) &\leq h(Y_1 \mid U_1, S) + h(Y_2 \mid U_1, S), \\ &= 2h(Y_k \mid U_1, S), \\ &\leq \mu\, h(Y_k \mid U_1, S). \end{aligned}$$

where the first inequality holds since conditioning reduces entropy while the second line follows the symmetry by assumption. The last inequality holds because

$$h(Y_k \mid U_1, S) \geq h(Z_k \mid U_1, S, X) = h(Z_k) \geq 0$$

and $\mu > 2$. Further, the second term in (24) is increasing with μ. Therefore, (24) is increasing with μ, when $\mu > 2$, which implies that the minimum of the objective function lies within $\mu \in [1, 2]$. ■

2.3.3 *Numerical results*

In Fig. 10, we provide a numerical example in which a symmetric Gaussian MISO-BC with i.i.d. Rayleigh fading is considered. We compare the achievable sum-rate of the proposed scheme derived from Corollary 2.3 to two cases. The first case is the sum-rate upper bound given by (26). The second case is the sum capacity of Gaussian BC when no CSIT is available. It is readily shown that the sum-rate in this case cannot be better than the single user rate. From Fano's

inequality and giving M_1 to user 2, we have

$$\begin{aligned} & n(R_1+R_2) \\ & \quad \le I(M_1;Y_1^n,S^n)+I(M_2;Y_2^n,M_1,S^n)+n\epsilon_n && \text{(27a)} \\ & \quad = h(Y_2^n \mid M_1,S^n)-h(Y_1^n \mid M_1,S^n) \\ & \qquad + h(Y_1^n \mid S^n)-h(Y_2^n \mid M_1,M_2,X^n,S^n)+n\epsilon_n \\ & \quad = I(X^n;Y_k^n \mid S^n)+n\epsilon_n, \quad k=1,2 && \text{(27b)} \end{aligned}$$

where we used the mutually independence between M_1, M_2, and S^n, the fact that X^n is a function of M_1 and M_2, and applied the symmetry, that is,

$$\begin{aligned} & h(Y_1^n \mid S^n)=h(Y_2^n \mid S^n), \\ & h(Y_1^n \mid M_1,S^n)=h(Y_2^n \mid M_1,S^n), \\ & h(Y_1^n \mid M_1,M_2,X^n,S^n)=h(Y_2^n \mid M_1,M_2,X^n,S^n)=h(Y_k^n \mid X^n,S^n). \end{aligned}$$

The comparison at finite SNR with some MAT-related schemes, e.g., [23, 24], is deferred to Chapter 3 for beyond two user case. We remind the readers that the region by the scheme proposed in [20] indeed includes our region in the symmetric scenario, which is the case in the numerical results. Thus, the comparison with SW region [20] is omitted. In addition, comparing our region with the region in [22] is difficult as the two schemes make use of the side information in completely distinct way at the decoders. To be specific, our scheme utilizes the backward decoding whereas the scheme in [22] exploits the block-wise decoding, where the messages are decoded at the end of each block. Furthermore, the complication of the scheme in [22] prevents us from obtaining a straightforward optimization on the RV choices in Gaussian case and hence the scheme itself is excluded from the simulation results here (since not even a numerical optimization can be claimed on such scheme).

From Fig. 10, the numerical evidence shows that delayed CSIT is beneficial even with moderate SNR (from less than 10 dB). In the medium-to-high SNR regime, remarkable rate gain is observed. Towards the high SNR regime, the slope of the sum-rate also identified as the DoF, of the proposed scheme is clearly higher than the one without CSIT and appears to be close to the one of the upper bound. This is consistent with the DoF results in [2] showing that the optimal sum DoF with delayed CSIT is 4/3 as compared to a DoF of 1 with no CSIT. Although it is not clearly shown in the plot, the proposed scheme may need higher SNR to achieve the optimal DoF provided that the gap against the outer bound increases slightly with the SNR. In fact, the gap between the sum-rate of the proposed scheme and the sum-rate upper bound is notable in all the SNR concerned. This may be due to the looseness of the outer bound in which the observation of one user is provided for free to the other user. An analytical characterization of this gap is an interesting direction of our future investigation.

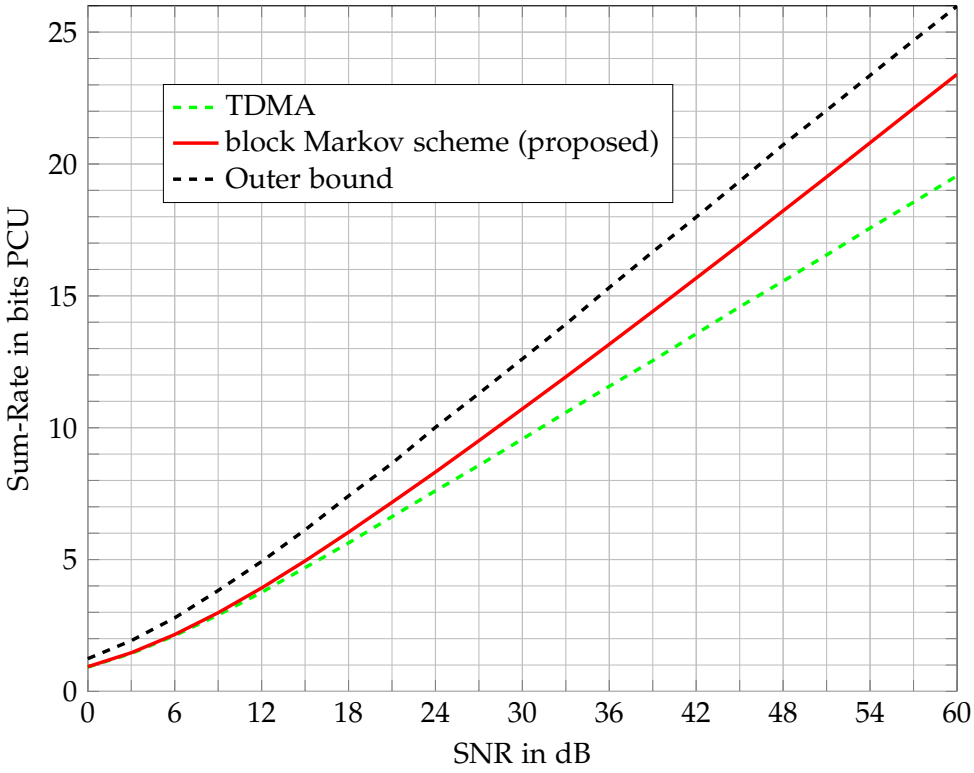

Figure 10: Sum-rate performance of the block Markov scheme against the corresponding upper bound and TDMA strategy.

2.4 CLOSING REMARKS

In this chapter, we studied the two-user BC with delayed CSIT. To be specific, we derived an inner bound by a scheme based on block Markov coding and an outer bound by the conventional genie-aided technique. The proposed scheme achieves the capacity of erasure BC and the optimal sum-DoF of Gaussian BC. In addition, we specified the inner and outer bounds in Gaussian channel where outer bound is tightened with some information theoretical inequalities. The parameters of both bounds are optimized numerically. In the simulation, remarkable gain by the proposed scheme over TDMA scheme is noted, specially in high SNR regime. Nevertheless, it is not trivial to optimize the parameters $(P_k,\ N)$ in the region of Corollary 2.3 since they locate inside the expectation. To that end, in Chapter 3, we propose a scheme whose optimization is relatively simple in not only the Gaussian channel, but also the general memoryless BC.

2.A PROOF OF THEOREM 2.1

Throughout the thesis, we use weak typical sequences in the achievability proofs. The definition and the main properties of weak typical sequences are provided in the Appendix A. We show the detailed proof for Theorem 2.1 in this Sub-appendix as an example on how we exploit weak typical sequences. Therefore, some details in the proofs that are similar to the technique we use here, unless necessary, will be omitted in the the remaining of the thesis.

We present below the scheme that obtains the region in Theorem 2.1. The transmission is divided into $B + L$ blocks with block length being n. Although the channel model is only defined for n transmission, it is reasonable to extend it to $n(B + L)$ transmissions since n tends to infinity. In addition, we assume that $\lim_{L,B\to\infty} \frac{L}{B} \to 0$ and $\lim_{n,L\to\infty} \frac{L}{n} \to 0$. The messages M_k are divided into B independent sub-messages, namely, $M_k = (m_{k[1]}, \ldots, m_{k[B]})$, $k = 1, 2$, with $m_{k[b]} \in \{1, \ldots, 2^{nR_k}\}$, for $b = [1 : B]$. Decoding at the destination is done *backwardly*. In particular, the last auxiliary index is first decoded from the observations obtained from last $L - 1$ blocks and then all compression indices and messages are jointly decoded in a reverse order. The region in Theorem 2.1 holds only when n, B, L tend to infinity. The RV V_k represents the encoded signal of the private message intended for user k and V_{12} is assigned to the encoded signal of the messages that are desired by both receivers.

2.A.1 *Codebook generation*

For an arbitrary number $\epsilon > 0$, we fix a pmf and function $x(v_{12}, v_1, v_2)$ that attains the region in Theorem 2.3, which reads as (14),

$$p(\hat{y} \,|\, x, v_{12}, v_1, v_2, s)p(x|v_{12}, v_1, v_2)p(v_1)p(v_2)p(v_{12}),$$

we proceed as follows.

1. Randomly and independently generate $2^{n\hat{R}}$ sequences v_{12}^n drawn i.i.d. from $p_{V_{12}^n}(v_{12}^n) = \prod_{i=1}^n p_{V_{12}}(v_{12i})$ and index them as $v_{12}^n(l)$ with $l \in \{1, \ldots, 2^{n\hat{R}}\}$[2] with
$$\hat{R} \triangleq I(\hat{Y}; V_1V_2X|V_{12}S) + \epsilon\,. \tag{28}$$
$\hat{R}$ is the rate for auxiliary message l.

2. Randomly and independently generate 2^{nR_k} sequences v_k^n drawn i.i.d. from $p_{V_k^n}(v_k^n) = \prod_{i=1}^n p_{V_k}(v_{ki})$ and index them as $v_k^n(m_k)$ with $m_k \in \{1, \ldots, 2^{nR_k}\}$, for each message $k = 1, 2$. R_k is the rate for private message intended for user k.

3. For each conceived sequence $v_{12}^n(l)$, randomly generate $2^{n\hat{R}}$ conditionally

[2] In fact, the size of alphabet of l should be the largest integer that is smaller than $2^{n\hat{R}}$. Throughout the thesis, the alphabets' sizes are always integers and for simplicity, we assume that 2^{nR} for any R is integer as n tends to infinity. This can be justified by the fact that removing a finite number from 2^{nR} results in negligible loss on the rate.

independent sequences $\hat{y}^n$ drawn from $p_{\hat{Y}|V_{12}}(\hat{y}^n|v_{12}^n) = \prod_{i=1}^n p_{\hat{Y}|V_{12}}(\hat{y}_i|v_{12i})$ and index them as $\hat{y}^n(l', l)$ with $l' \in \{1, \ldots, 2^{n\hat{R}}\}, l \in \{1, \ldots, 2^{n\hat{R}}\}$.

4. For each tuple of (m_1, m_2, l), the source generate x^n as

$$x_i = x(v_{12i}(l), v_{1i}(m_1), v_{2i}(m_2)).$$

5. Provide the corresponding codebooks to the decoders.

2.A.2 *Encoding*

The Table 1 is provided to illustrate the encoding and decoding of this block Markov scheme.

- For every block $b = [1 : B]$, the source sends messages $(m_{1[b]}, m_{2[b]})$ based on $\left(v_1^n(m_{1b}), v_2^n(m_{2b})\right)$. Moreover, for $b = [B+1 : B+L]$, the source sends dummy messages $m_{k[b]} = 1$ assumed to be known to both destinations.
- For every block $b = [1 : B+1]$ and after receiving the state sequence $s^n(b-1)$ via noiseless feedback links, the encoder searches for at least one index l_b such that

$$\begin{aligned}\left(v_{12}^n(l_{b-1}), v_1^n(m_{1[b-1]}), v_2^n(m_{2[b-1]}), s^n(b-1), x^n(b-1), \hat{y}^n(l_{b-1}, l_b)\right) \\ \in \mathcal{T}_\delta^n(V_{12}V_1V_2SX\hat{Y})\end{aligned}$$

  are *jointly typical*. By convention, we let $l_0 = 1$. To avoid degenerate case, even though $m_{k[0]}$ and $s^n(0)$ should not exist, we assume that $m_{k[0]} = 1$ and $s^n(0) = \{1, 1, \cdots, 1\}$.

  Without loss of generality, we assume that L_{b-1}, $M_{1[b-1]}$, and $M_{2[b-1]}$ are chosen in block $b-1$. We define the error event for this encoding as

$$\begin{aligned}\mathcal{E}_1(b) \triangleq \{(v_{12}^n(L_{b-1}), v_1^n(M_{1[b-1]}), v_2^n(M_{2[b-1]}), s^n(b-1), x^n(b-1), \hat{y}^n(L_{b-1}, l_b)) \\ \notin \mathcal{T}_\delta^n(V_{12}V_1V_2SX\hat{Y}) \text{ for all } l_b \in [1 : 2^{n\hat{R}}]\}.\end{aligned}$$

  The probability that $\mathcal{E}_1(b)$ occurs goes to zero as n tends to infinity, provided by expression (28). The upper bounding the error probability (by using union bounds) and the analysis of all error events concerning the proposed block Markov scheme are deferred to the end of this appendix for readability. We only define the error events in the encoding and decoding sections for each encoding/decoding operation and give the rate constraints, whose explanations are provided in the Section 2.A.4.

- For $b = [1 : B+1]$, the source is assumed to know the index l_{b-1} generated based on the information of previous block, and it finds at the beginning of the block b the sequence $v_{12}^n(l_{b-1})$. For $b = [B+2 : B+L]$, the source let $l_b = l_{B+1}$.
- Finally, at block b the encoder sends $x^n(b)$ that can be learned from the outputs of the function $x_i(b) = x(v_{12i}(l_{b-1}), v_{1i}(m_{1[b]}), v_{2i}(m_{2[b]}))$.

Block	$b=1$	$b=2$	$\cdots$	$b=B$	$b=B+1$	$b=B+2$	$\cdots$	$b=B+L$
V_k	$v_k^n(m_{k1})$	$v_k^n(m_{k2})$	$\cdots$	$v_k^n(m_{kB})$	$v_k^n(1)$	$v_k^n(1)$	$\cdots$	$v_k^n(1)$
$\hat{Y}$	$\hat{y}^n(1,l_1)$	$\hat{y}^n(l_1,l_2)$	$\cdots$	$\hat{y}^n(l_{B-1},l_B)$	$\hat{y}^n(l_B,l_{B+1})$	———		
V_{12}	$v_{12}^n(1)$	$v_{12}^n(l_1)$	$\cdots$	$v_{12}^n(l_{B-1})$	$v_{12}^n(l_B)$	$v_{12}^n(l_{B+1})$	$\cdots$	$v_{12}^n(l_{B+1})$
Y_k	$\hat{l}_1, m_{k1}$	$\hat{l}_2, m_{k2}$	$\cdots$	$\hat{l}_{B-1}, m_{k(B-1)}$	$\hat{l}_B, m_{kB}$	$\hat{l}_{B+1}$		

Table 1: Encoding and decoding of the block Markov scheme for Theorem 2.3

2.A.3 *Decoding and error events*

1. We shall assume the next condition holds:

$$I(\hat{Y}; XV_{\bar{k}}|V_{12}V_kY_kS) \quad \leq \quad I(V_{12}; Y_k|V_kS)\ , \tag{29}$$

for $k, \bar{k} = 1, 2$ and $k \neq \bar{k}$. This is to guarantee the right-hand side of (37) is strictly non-negative, which ensures the rate R_1 is bounded away from zero. However, if it was not the case, multicasting useful common message l via V_{12} to both user is too costly as compared to the gain provided by the side information $\hat{Y}$. Hence, in such a case, the encoder does not generate common messages l's and obtain x^n from the function $x(v_{12}, v_1, v_2)$ by letting l be dummy message 1 (which is known by all the destinations). The decoders would proceed as follows. Destination k –with the assumption that $v_k^n(m_{k[b]})$ has been correctly decoded– looks for a unique sequence $v_k^n(m_{k[b-1]})$ such that

$$\big(v_k^n(m_{k[b-1]}), y_k^n(b-1), s^n(b-1)\big) \in \mathcal{T}_\delta^n(V_kY_kS).$$

Here the decoding is performed in a backward manner, however, backward decoding does not provide any gain and is used to give a simple example of backward decoding so that the readers are prepared for the backward decoding used in the coming parts. In other words, backward decoding is used only to make the proof look coherent. We define the error events of this decoding as

$$\begin{aligned}
\mathcal{E}_2(b-1) &= \{(v_k^n(M_{k[b-1]}), y_k^n(b-1), s^n(b-1)) \notin \mathcal{T}_\delta^n(V_kY_kS)\},\\
\mathcal{E}_3(b-1) &= \{(v_k^n(\hat{M}_{k[b-1]}), y_k^n(b-1), s^n(b-1)) \in \mathcal{T}_\delta^n(V_kY_kS)\\
&\qquad\qquad \text{for some } \hat{M}_{k[b-1]} \neq M_{k[b-1]}\},
\end{aligned}$$

where $\mathcal{E}_2(b-1)$ denotes the event that the sequence by correct message is not jointly typical with the output and the state sequence whereas $\mathcal{E}_2(b-1)$ stands for the case that the sequence based on a wrong message is jointly typical with the output and the state sequence. The probability of such error event goes to zero provided that

$$R_k < I(V_k; Y_k|S)\ . \tag{30}$$

2. Otherwise, we assume condition (29) holds. In this case, both destinations wait until the last block and uses *backward decoding*. Indeed, receiver k searches for the unique index l_{B+1} such that

$$\left(v_{12}^n(l_{B+1}), v_k^n(1), y_k^n(b), s^n(b)\right) \in \mathcal{T}_\delta^n(V_{12}V_kY_kS) \ ,$$

holds for all $b \in \{B+2, \ldots, B+L\}$ simultaneously. We define the error events of this decoding as

$$\begin{aligned}
\mathcal{E}^{'} &= \{L_{B+1} \text{ is not correctly decoded}\}, \\
\mathcal{E}_4(b) &= \{(v_{12}^n(L_{B+1}), v_k^n(1), y_k^n(b), s^n(b)) \notin \mathcal{T}_\delta^n(V_{12}V_kY_kS)\}, \\
\mathcal{E}_5(b) &= \{(v_{12}^n(\hat{L}_{B+1}), v_k^n(1), y_k^n(b), s^n(b)) \in \mathcal{T}_\delta^n(V_{12}V_kY_kS) \\
&\qquad \text{for some } \hat{L}_{B+1} \neq L_{B+1}\},
\end{aligned}$$

Then, we have

$$\Pr(\mathcal{E}^{'}) \leq \prod_{b=B+2}^{B+L} (\Pr(\mathcal{E}_4(b)) + \Pr(\mathcal{E}_5(b)))$$

The probability of decoding error goes to zero provided, for $k = 1, 2$,

$$I(\hat{Y}; V_1V_2X|V_{12}S) + \epsilon' = \hat{R} \leq (L-1)I(V_{12}; V_kY_k|S) \ . \tag{31}$$

3. After finding l_{B+1} and since $m_{1[B+1]} = m_{2[B+1]} = 1$, each destination decodes jointly the indices $(m_{1[b]}, l_b)$ and $(m_{2[b]}, l_b)$, respectively, where again decoding is performed "backwardly" with the assumptions that $(m_{1[b+1]}, l_{b+1})$ and $(m_{2[b+1]}, l_{b+1})$ have already been decoded correctly at the corresponding destination.

   Consider the decoder 1, that finds the unique pair of indices $(m_{1[b]} = \hat{M}_{1[b]}, l_b = \hat{L}_b)$ (estimated messages) such that

$$\left(v_{12}^n(l_b), v_1^n(m_{1[b]}), \hat{y}^n(l_b, L_{b+1}), y_1^n(b), s^n(b)\right) \in \mathcal{T}_\delta^n(V_{12}V_1\hat{Y}Y_1S) \ . \tag{32}$$

   Several error events have to be considered in order to guarantee that the error probability of (32) goes to zero:

   - An error occurs if some of the transmit and received codewords in (32) are not joint typical. The error event is defined as

     $$\mathcal{E}_6(b) = \{(v_{12}^n(L_b), v_1^n(M_{1[b]}), \hat{y}^n(L_b, L_{b+1}), y_1^n(b), s^n(b)) \notin \mathcal{T}_\delta^n(V_{12}V_1\hat{Y}Y_1S)\}.$$

     However, by the way in which the codebooks are generated and encoding is done, this has an arbitrary small probability.

   - An error occurs if $(l_b \neq L_b, m_{1[b]} = M_{1[b]})$. Such error event is defined as

     $$\begin{aligned}
     \mathcal{E}_7(b) = \{&(v_{12}^n(\hat{L}_b), v_1^n(M_{1[b]}), \hat{y}^n(\hat{L}_b, L_{b+1}), y_1^n(b), s^n(b)) \in \mathcal{T}_\delta^n(V_{12}V_1\hat{Y}Y_1S) \\
     &\qquad \text{for some } \hat{L}_b \neq L_b\}.
     \end{aligned}$$

The probability of this event goes to zero provided that

$$\hat{R} < I(V_{12}\hat{Y}; V_1 Y_1 S) \ . \tag{33}$$

- An error occurs if $(l_b = L_b, m_{1[b]} \neq M_{1[b]})$. Such error event is defined as

$$\mathcal{E}_8(b) = \{(v_{12}^n(L_b), v_1^n(\hat{M}_{1[b]}), \hat{y}^n(L_b, L_{b+1}), y_1^n(b), s^n(b)) \in \mathcal{T}_\delta^n(V_{12}V_1\hat{Y}Y_1S) \text{ for some } \hat{M}_{1[b]} \neq M_{1[b]}\}.$$

Again the probability of this event goes to zero provided that

$$R_1 < I(V_1; Y_1\hat{Y}|V_{12}S) \ . \tag{34}$$

- An error occurs if $(l_b \neq L_b, m_{1[b]} \neq M_{1[b]})$. Such error event is defined as

$$\mathcal{E}_9(b) = \{(v_{12}^n(\hat{L}_b), v_1^n(\hat{M}_{1[b]}), \hat{y}^n(\hat{L}_b, L_{b+1}), y_1^n(b), s^n(b)) \in \mathcal{T}_\delta^n(V_{12}V_1\hat{Y}Y_1S) \text{ for some } \hat{M}_{1[b]} \neq M_{1[b]}, \hat{L}_b \neq L_b\}.$$

The probability of this event goes to zero provided that [73]

$$R_1 + \hat{R} - I(\hat{Y}; V_1V_2XS|V_{12}) < I(V_1V_{12}; Y_1|S) - I(\hat{Y}; XV_2|V_{12}V_1Y_1S) \ . \tag{35}$$

4. We first remark that by using some algebra the combination of expressions (28) and (33) yields the condition stated in (29), which is satisfied by assumption (see [73] for further details). Thus, we only need to focus on the other expressions. By letting (L, B, n) tend to infinity such that $\frac{L}{B}$ goes to zero as $n \to \infty$, we obtain from expressions (34), (28) and (35) the following conditions:

$$\begin{aligned} R_1 &< I(V_1; Y_1\hat{Y}|V_{12}S) \ , & (36)\\ R_1 &< I(V_{12}V_1; Y_1|S) - I(\hat{Y}; XV_2|V_{12}V_1Y_1S) \ . & (37) \end{aligned}$$

Finally, by exchanging the indices in (36) and (37) we can derive the complementary expressions for receiver 2. The proof concludes from this observation by adding a time-sharing random variable Q.

2.A.4 *Error events analysis*

The probability of error throughout the communication is upper bounded as

$$\begin{aligned} P_e \leq & \Pr(\mathcal{E}_1(b)) + \Pr(\mathcal{E}_2(b)) + \Pr(\mathcal{E}_3(b)) + \Pr(\mathcal{E}') + \Pr(\mathcal{E}_6(b)) \\ & + \Pr(\mathcal{E}_7(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}'^c) + \Pr(\mathcal{E}_8(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}'^c) + \Pr(\mathcal{E}_9(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}'^c) \end{aligned}$$

where P_e denotes the probability that at least one error occurs during the communication.

By the independence of codebooks and the covering lemma for weak typicality (Lemma A.7) [77], as long as $\hat{R} \geq I(\hat{Y}; V_1 V_2 X | V_{12} S)$ holds, the sequences

$$(v_{12}^n(l_{b-1}), v_1^n(m_{1[b-1]}), v_2^n(m_{2[b-1]}), s^n(b-1), x^n(b-1), \hat{y}^n(l_{b-1}, l_b))$$

are jointly typical. This is due to the fact that $\hat{y}^n$ is conditionally independent of (v_1^n, v_2^n, s^n, x^n) given v_{12}^n and the sequences $(v_1^n, v_2^n, s^n, x^n, v_{12}^n)$ are jointly typical. As such, the encoder can always find a l_b to ensure that $\Pr(\mathcal{E}_1(b))$ tends to zero as n goes to infinity. Therefore, we can let $\hat{R}$ be $I(\hat{Y}; V_1 V_2 X | V_{12} S) + \epsilon$.

Given that the RVs used in the encoding and decoding methods when $\mathcal{E}_2(b)$ and $\mathcal{E}_3(b)$ may occur are irrelevant to the side information $\hat{Y}$ and the common-message-bearing signal V_{12}, we use $\Pr(\mathcal{E}_2(b))$ instead of $\Pr(\mathcal{E}_2(b) \cap \mathcal{E}_1^c(b))$ in the upper bound of P_e. $\Pr(\mathcal{E}_2(b))$ tends to zero as n goes to infinity by considering the independence of the codebooks and a straightforward extension of the conditional typicality lemma (Lemma A.4) to a set of sequences $(v_k^n,\ y_k^n,\ s^n)$ that concern three RVs (V_k, Y_k, S). Since v_k^n and s^n are independent, they are jointly typical by definition. In addition, y_k^n follows a distribution $p(y_k^n | v_k^n, s^n) = \prod_{i=1}^n p_{Y_k|V_k,S}(y_{k,i} | v_{k,i}, s)$. Thus, the conditional typicality lemma still holds when an extra conditioning on the s^n is included.

We associate the v_k^n, ϕ, and (y_k^n, s^n) in the error event $\mathcal{E}_3(b)$ with X^n, $\tilde{U}^n$, and $\tilde{Y}^n$ in the packing lemma (Lemma A.6), respectively. By the same independence and the packing lemma, $\Pr(\mathcal{E}_3(b))$ tends to zero as $n \to \infty$ if $R_k < I(V_k; Y_k S | \phi) = I(V_k; Y_k | S)$ is satisfied, where last equality holds since V_k is independent of S.

We know that $\Pr(\mathcal{E}') \leq \prod_{b=B+2}^{B+L} (\Pr(\mathcal{E}_4(b)) + \Pr(\mathcal{E}_5(b)))$. With a similar argument, $\Pr(\mathcal{E}_4(b))$ tends to zero as n goes to infinity due to the independence of the codebooks and a straightforward extension of the conditional typicality lemma (Lemma A.4) to four RVs case. Note that $v_k^n(1)$ that known to both ends of the communication is independent of all the other sequence. Then, $(v_{12}^n(L_{B+1}), v_k^n(1), s^n(b)) \in \mathcal{T}_\delta^n(V_{12} V_k S)$ is true since they are mutual independent. In addition, y_k^n follows the probability $P(y_k^n | v_{12}^n, v_k^n, s^n) = \prod_{i=1}^n P_{Y_k|V_{12}V_kS}(y_{ki} | v_{12i} v_{ki} s_i)$. As such, the conditional typicality lemma can be applied. Then, we can update the error probability as

$$\begin{aligned}
\Pr(\mathcal{E}') &\leq \prod_{b=B+2}^{B+L} (\Pr(\mathcal{E}_4(b)) + \Pr(\mathcal{E}_5(b))) \\
&\leq \prod_{b=B+2}^{B+L} \left(\epsilon' + \Pr(\mathcal{E}_5(b))\right) \\
&\leq \prod_{b=B+2}^{B+L} \Pr(\mathcal{E}_5(b)) \\
&= (L-1)\Pr(\mathcal{E}_5)
\end{aligned}$$

where ϵ' tends to zero at the same rate as n, where leads to a negligible term in the second inequality since we assumed that $\lim_{n,L\to\infty} \frac{L}{n} \to 0$. The last equality follows that the probability of error event $\mathcal{E}_5(b)$ in block b is identical for $b = B+2, \cdots, B+L$.By the same independence and an extended version of the packing lemma (Lemma A.6), $(L-1)\Pr(\mathcal{E}_5(b))$ tends to zero as $n \to \infty$ if $\hat{R} \leq$

$(L-1)I(V_{12};V_kY_k|S)$ is satisfied. To apply the packing lemma to this error event, we associate v_{12}^n, ϕ, and (v_k^n, y_k^n, s^n) in the error event $\mathcal{E}_5(b)$ with X^n, $\tilde{U}^n$, and $\tilde{Y}^n$ in the packing lemma. Since L tends to infinity, such a constraint always holds as long as $\hat{R}$ is finite.

Upper bounding $\Pr(\mathcal{E}_6(b))$ follows a similar argument as bounding $\Pr(\mathcal{E}_4(b))$, i.e., $\Pr(\mathcal{E}_6(b))$ tends to zero as n goes to infinity due to the independence of the codebooks and an extension of the conditional typicality lemma (Lemma A.4) to five RVs case. This can be justified as follows. Given that L_{b+1} is known correctly, the pair of $(v_{12}^n(L_b), \hat{y}^n(L_b, L_{b+1}))$, $v_1^n(M_{1[b]})$, and $s^n(b)$ are mutual independent and thus jointly typical. In addition, y_1^n follows the $p(y_1^n|v_{12}^n, \hat{y}^n, v_1^n, s^n) = \prod_{i=1}^n p_{Y_1|V_{12}\hat{Y}V_1S}(y_{1i}|v_{12i}, \hat{y}_i, v_{1i}, s_i)$.

Suppose that events $\mathcal{E}_1(b)$ and $\mathcal{E}^{'}$ do not occur, then $\Pr(\mathcal{E}_7(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}^{'c})$ goes to zero as $n \to \infty$ if $\hat{R} < I(V_{12}\hat{Y}; V_1Y_1S)$ is fulfilled. The vanishing error probability of $\mathcal{E}_7(b)$ is justified by the independence of codebooks along with the packing lemma (Lemma A.6) extended to five RVs case. To apply the packing lemma to this error event, we associate $(v_{12}^n, \hat{y}^n)$, ϕ, and (v_1^n, y_1^n, s^n) in the error event $(\mathcal{E}_7(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}^{'c})$ with X^n, $\tilde{U}^n$, and $\tilde{Y}^n$ in the packing lemma. With a standard argument, $\Pr(\mathcal{E}_8(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}^{'c})$ tends to zero as n approaches infinity if $R_1 < I(V_1; Y_1\hat{Y}V_{12}S) = I(V_1; Y_1\hat{Y}|V_{12}S)$, which can be proved given the independence of codebooks, the packing lemma (Lemma A.6), and the independence between V_1 and (V_{12}, S). To apply the packing lemma to this error event, we associate v_1^n, ϕ, and $(v_{12}^n, \hat{y}^n, y_1^n, s^n)$ in the error event $(\mathcal{E}_8(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}^{'c})$ with X^n, $\tilde{U}^n$, and $\tilde{Y}^n$ in the packing lemma.

Finally, the error probability of $(\mathcal{E}_9(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}^{'c})$ tends to zero as $n \to \infty$ if

$$\begin{gathered} R_1 + \hat{R} < I(V_1; Y_1S) + I(\hat{Y}V_{12}; V_1Y_1S) \\ \Updownarrow \\ R_1 + \hat{R} - I(\hat{Y}; V_1V_2XS|V_{12}) < I(V_1V_{12}; Y_1|S) - I(\hat{Y}; XV_2|V_{12}V_1Y_1S) \end{gathered}$$

This can be demonstrated by applying twice the joint typicality lemma (Lemma A.5) given that the codebooks are independent.

$$\begin{aligned}
&\Pr((\mathcal{E}_9(b) \cap \mathcal{E}_1^c(b) \cap \mathcal{E}^{'c})) \\
\leq &\sum_{\hat{M}_{1[b]} \neq M_{1[b]}, \hat{L}_b \neq L_b} P(\{v_1^n, y_1^n, s^n, v_{12}^n, \hat{y}^n\} \in \mathcal{T}_\delta^n(V_1Y_1SV_{12}\hat{Y})) \\
= &\sum_{\hat{M}_{1[b]} \neq M_{1[b]}, \hat{L}_b \neq L_b} \sum_{\{v_1^n, y_1^n, s^n, v_{12}^n, \hat{y}^n\} \in \mathcal{T}_\delta}^n p(v_1^n, y_1^n, s^n, v_{12}^n, \hat{y}^n) \\
= &\sum_{\hat{M}_{1[b]} \neq M_{1[b]}, \hat{L}_b \neq L_b} \sum_{\{v_1^n, y_1^n, s^n, v_{12}^n, \hat{y}^n\} \in \mathcal{T}_\delta}^n p(v_1^n) p(y_1^n, s^n) p(v_{12}^n, \hat{y}^n) \\
\leq &\sum_{\hat{M}_{1[b]} \neq M_{1[b]}, \hat{L}_b \neq L_b} |\mathcal{T}_\delta^n(V_1Y_1SV_{12}\hat{Y})| 2^{-n(H(V_1)-\delta H(V_1))} 2^{-n(H(Y_1,S)-\delta H(Y_1,S))} 2^{-n(H(V_{12},\hat{Y})-\delta H(V_{12},\hat{Y}))} \\
\leq &\sum_{\hat{M}_{1[b]} \neq M_{1[b]}, \hat{L}_b \neq L_b} 2^{n(H(V_1,Y_1,S,V_{12},\hat{Y})+\delta H(V_1,Y_1,S,V_{12},\hat{Y}))} 2^{-nH(V_1)} 2^{-nH(Y_1,S)} 2^{-nH(V_{12},\hat{Y})} 2^{n\delta'} \\
\leq &2^{nR_1} 2^{n\hat{R}} 2^{-n(I(V_1;Y_1S)+I(V_{12}\hat{Y};V_1Y_1S)-\epsilon(\delta))}
\end{aligned}$$

where there always exists $\epsilon(\delta) > 0$ tends to zero as δ goes to zero.

3

K user BC with State Feedback: Distributed Compression method

3.1 OVERVIEW

Apart from the uniformity across the delayed CSIT schemes in EBC and GBC mentioned in Chapter 1, the positive results in Chapter 2 [38] at finite SNR also serves as an inspiration for the work of this chapter. In the previous chapter, the numerical results showed that by carefully designing the coding RVs in the block Markov scheme, the region of delayed CSIT can outperform no-feedback one in a medium-to-low SNR. However, there are two main shortcomings: 1) the optimal choice of scheme parameters for the proposed inner bound cannot be easily obtained even with numerical optimizer in a two user GBC; 2) the considerable gap against outer bound reveals the possibilities that the coding tools and/or scheme structure may not suit the delayed CSIT. In order to simplify the region as well as the optimization, we seek a simple tool that can address efficiently and uniformly the K user BC (both EBC and GBC) with homogeneous state feedback. To be concrete on the performance gain in GBC, we will involve several other inner bounds that are related to MAT scheme.

In particular, we focus on the K-user (mainly $K > 2$) case and investigate the state-dependent BCs with homogeneous state feedback which includes both the EBC and the Fading GBC as special cases. The channel models are presented in Section 1.2. The main contributions are enclosed as follows.

- We propose a non-linear scheme whose main ingredient is a joint source-channel coding that includes distributed lossy coding (distributed compression) and coded time sharing. We quote the scheme by distributed compression (DC) scheme.[1] The corresponding rate region is derived in the general BC. It is worth mentioning that our goal here is not to derive a general rate region that includes all the known regions (e.g., the two-user Shayevitz-Wigger region [20]). Instead, we are interested in schemes with relatively simple structure in order to have a numerically tractable rate region. To that end, we make some reasonable choices such as excluding block Markov coding and binning at the transmitter. Since the instantaneous state is unknown, we know that binning has limited benefit in GBC. Therefore, the proposed region is simple to evaluate.

[1] To avoid confusion, we clarify that the scheme name is invoked in order to emphasize the difference against some conventional schemes that use joint compression, e.g., Wyner-Ziv compression in [38], rather than to imply a specific source coding problem.

- Instead of using an input alphabet of limited size ($2^L \geq K$) [27, 28], we show that the proposed scheme achieves, for *any* input alphabet size, capacity of a symmetric EBC and a spatially independent EBC that fulfills one-sided fairness. Further, for the Fading symmetric GBC, we derive an achievable symmetric rate as a function of SNR and compression noises. Thanks to the symmetry and coded time sharing structure, only $K-1$ quantization covariances are involved in the optimization, which then becomes more feasible. Analytical results show that the proposed scheme achieves the optimal DoF under the same setting as in [2], while a numerical example of the two/three-user symmetric rate reveals a superior performance, for a wide range of SNR, over several existing schemes in the literature.

The remaining of the chapter is organized as: we provide the main results in Section 3.2. The novel scheme is presented in Section 3.3. An example on the Fading Gaussian BC is given in Section 3.4. Finally, Section 3.5 summarizes the chapter. The detailed proofs are relegated to the appendices.

3.2 MAIN RESULTS

In this section, we first introduce some RVs that help comprehend the proposed scheme of this chapter and the corresponding rate region as claimed in Theorem 3.1. Then, the main result of this paper is presented as an achievable rate region in the general BC, followed by its application in EBC and Fading GBC. Specifically, the rate region achieves capacity under certain condition in EBC and obtains optimal DoF in Fading GBC. The details of the scheme that achieves the rate region in Theorem 3.1 is provided in Section 3.3. In the proposed scheme, all the transmission slots are divided into K part with each part is named a phase, where phase j is dedicated to multicast different common information to the respective user set $\mathcal{J}$ that has $|\mathcal{J}| = j$ users.

3.2.1 *RVs and general region*

The RVs are determined as follows.

- $X^{(j)}$, $\{Y_k^{(j)}\}_{k\in\mathcal{K}}$, $S^{(j)}$, and $Q^{(j)}$ are the channel input, the channel output at user k, the channel state, and time sharing RVs, respectively, for phases $j = 1, \ldots, K$, with pmf

$$\prod_{j=1}^{K} p(y_1^{(j)}, \ldots, y_K^{(j)} \mid x^{(j)}, s^{(j)}) p(x^{(j)} \mid s^{(j)}, q^{(j)}) p(s^{(j)}) p(q^{(j)}). \tag{38}$$

- $(\hat{Y}_{i\to\mathcal{J}}, V_{i\to\mathcal{J}})$, $i < j$ and $|\mathcal{J}| = j$, are the side information intended for all the users in $\mathcal{J}$ from phase i, and the signal that carries such information,

respectively, with pmf

$$\prod_{j=1}^{K}\prod_{i=0}^{j-1}\prod_{\mathcal{J}} p(v_{i\to\mathcal{J}})p(\hat{y}_{i\to\mathcal{J}} \mid \{v_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, s^{(i)}, q^{(i)}), \tag{39}$$

where we define $v_{\mathcal{U}} \triangleq \{v_{k\to\mathcal{U}} : k < |\mathcal{U}|\}$ and $V_{\mathcal{U}} \triangleq \{V_{k\to\mathcal{U}} : k < |\mathcal{U}|\}$ for simplicity in the scenario where the origin phase k of $v_{k\to\mathcal{U}}$ and $V_{k\to\mathcal{U}}$ does not matter in the analysis. $V_{i\to\mathcal{J}}$ is transmitted in phase j and the corresponding message is wanted by users in subset $\mathcal{J}$. The messages that are sent in phase i and are desired by users in subset $\mathcal{I} \subset \mathcal{J}$ can be decoded with the help of $\hat{Y}_{i\to\mathcal{J}}$. Note that from pmf's perspective, the side information $\hat{y}_{i\to\mathcal{J}}$ is not built on $v_{i\to\mathcal{J}}$, which differs from the block Markov scheme proposed in Chapter 2.

- $X^{(j)}$ is the channel input at phase j, $j = 1, \ldots, K$, which is generated with the pmf

$$\prod_{j=1}^{K} p(x^{(j)} \mid \{v_{\mathcal{J}}\}_{\mathcal{J}}, q^{(j)}). \tag{40}$$

The achievable region in general case is stated below.

Theorem 3.1 (DC Inner bound). *A rate tuple $(R_1, \ldots, R_K)$ is achievable with the proposed scheme, that will be described in Section 3.3, for the K-user BC with causal state feedback provided that*

$$R_k \le \alpha_1 I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \mid S^{(1)}, Q^{(1)}), \tag{41a}$$

$$0 \le \min_{\substack{i,j,k,\mathcal{J}:\\ i<j,k\in\mathcal{J}}} \Big\{ \alpha_j I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \mid S^{(j)}, Q^{(j)}) - \alpha_i I(\{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \mid Y_k^{(i)}, S^{(i)}, Q^{(i)}) \Big\}, \tag{41b}$$

for some K-tuple $(\alpha_1, \ldots, \alpha_K) \in \mathbb{R}_+^K$ with $\sum_k \alpha_k = 1$, and some pmf as described in (39) *and* (40).

Proof. The detailed proof is relegated to Section 3.3, where a K phase scheme is described. Briefly, the scheme can be recognized as: K unicast messages are sent in phase 1 while the successive $K-1$ phases are designed to multicast side information that would be useful for certain subset of users. The ith phase is allocated to the messages that are intended for any subset containing i users. In the phase i, the source transmits the messages $M_{\mathcal{I}}$ at rates that are larger than those can be recovered with the phase i's observation only. Although the messages are only wanted by users in $\mathcal{I}$, they are as well observed by users in $\mathcal{K} \setminus \mathcal{I}$ due to the wireless nature of the channel. Then, at the end of phase i, the encoder selects some subsets $\mathcal{J} \supset \mathcal{I}$ based on the state feedback and generates the refinement information $\hat{Y}$ intended for all users in $\mathcal{J}$. The side information $\hat{Y}_{i\to\mathcal{J}}$ that reduces the ambiguity in decoding $M_{\mathcal{I}}$ is associated with $M_{i\to\mathcal{J}}$ which will be sent out in phase j. At each destination k, backward decoding is involved to make sure that

the messages $M_{i\to\mathcal{J}}$ are decoded with the knowledge from subsequent phases $\hat{Y}_{j\to\mathcal{U}}$, for $k \in \mathcal{J} \subset \mathcal{U}$.

Therefore, the condition (41a) is interpreted as the rate of phase 1's message that can be reliably decoded at destination k with all the side information user k collected during the rest of the phases. (41b) guarantees that the side information $\hat{Y}_{i\to\mathcal{J}}$ encoded into $M_{i\to\mathcal{J}}$ can be perfectly recovered at user k from Y_k and the respective upgraded side information, i.e., $\{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}$. ■

Remark 3.1. *As mentioned in Section 2.2.1, the scheme in [22] utilizes a decoding method differs from backward decoding which makes it rather difficult to compare its region with the one in Theorem 3.1 for general pmf's. Moreover, we recall that the pmf that attains the DC region contains $p(\hat{y}_{i\to\mathcal{J}} \mid \{v_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, s^{(i)}, q^{(i)})$ which differs from the one in SW region [20] in that the quantized side information $\hat{y}_{i\to\mathcal{J}}$ is independent of the corresponding signal $v_{i\to\mathcal{J}}$. Although both SW and DC schemes employ distributed lossy source coding, the proposed scheme performs the distributed compression within a joint source-channel coding structure, with the help of which the number of rate constraints is reduced (see Section 3.3). In addition, the SW region is governed by some mismatched constraints, as described in Section 2.2.1. Given that, the optimal pmf for SW could be completely different to the one for DC region. Thus, the comparison between SW and DC regions is not trivial. Nevertheless, thanks to the symmetry which releases the mismatched conditions, the SW region does achieve the symmetric capacity of EBC as well as the proposed region. It is worth mentioning that the region in Theorem 2.1 may not contain the one in Theorem 3.1 due to the distinctions on pmf.*

3.2.2 *Achievable region in EBC*

We apply the rate region in Theorem 3.1 to EBC and obtain the following corollary.

Corollary 3.2. *For any input alphabet size $|\mathcal{X}|$, the rate $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ is achievable in EBC with state feedback by the proposed scheme if the sequel is satisfied,*

$$R_k \le \alpha_1 P(Q_2^{(1)} = k)(1-\delta_{\mathcal{K}}) \log|\mathcal{X}|, \tag{42a}$$

$$0 \le \min_{\substack{j,k,\mathcal{J}:\\ k\in\mathcal{J}}} \Big\{ \alpha_j P(Q_2^{(j)} = \mathcal{J})(1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}}) - \sum_{i=1}^{j-1} \alpha_i \sum_{\mathcal{I}\subset\mathcal{J},\mathcal{I}\ni k} P(Q_2^{(i)} = \mathcal{I}) \mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k\},\mathcal{J}\setminus\mathcal{I}} \Big\}, \tag{42b}$$

for some K-tuple $(\alpha_1, \ldots, \alpha_K) \in \mathbb{R}_+^K$ with $\sum_k \alpha_k = 1$.

Proof. To prove the achievability, we apply the following auxiliary RVs to the achievable rate region in Theorem 3.1.

- Time-sharing RVs $Q^{(j)} \triangleq (Q_1^{(j)}, Q_2^{(j)}) \in \mathcal{Q}^{(j)}$ where

$$\mathcal{Q}^{(j)} \triangleq \begin{cases} \{(q_1, q_2): \quad q_1 = 0, q_2 \in [1:K]\}, & j = 1 \\ \{(q_1, q_2): \quad q_1 \in [1:j-1], q_2 = \mathcal{J}\}, & j \geq 2 \end{cases} \tag{43}$$

  with $|\mathcal{Q}^{(1)}| = K$ and $|\mathcal{Q}^{(j)}| = (j-1)\binom{K}{j}$, for $j \geq 2$. We let

$$p(q^{(j)}) = p(q_1^{(j)})p(q_2^{(j)}), \quad \text{if } q^{(j)} \in \mathcal{Q}^{(j)} \tag{44}$$

  and 0 otherwise. $q^{(j)}$ at jth phase is divided in two, $q_1^{(j)}$ for origin of the information and $q_2^{(j)}$ for the set of the dedicated users. In phase 1, the signal that carries the initial message M_k is transmitted when $q_2^{(1)} = k$; in phase j ($j > 2$), the message $M_{i\to\mathcal{J}}$ is conveyed via $V_{i\to\mathcal{J}}$ while $q_1^{(j)} = i$ and $q_2^{(j)} = \mathcal{J}$, as shown in (46).

- Uniform V's in $\mathcal{X}$: for any $i < j$ and $\mathcal{J}$, we set $\mathcal{V}_{i\to\mathcal{J}} = \mathcal{X}$, and

$$p(v_{i\to\mathcal{J}}) = \frac{1}{|\mathcal{X}|}, \quad \text{if } v_{i\to\mathcal{J}} \in \mathcal{X} \tag{45}$$

  and 0 otherwise. This ensures $H(V_{i\to\mathcal{J}}) = \log|\mathcal{X}|$.

- Input as deterministic function of V and Q:

$$X^{(j)} = V_{Q_1^{(j)}\to Q_2^{(j)}}. \tag{46}$$

- Side information $\hat{Y}$ as deterministic function of (X, S, Q): for any $i < j$ and $\mathcal{J}$,

$$\hat{Y}_{i\to\mathcal{J}} = X^{(i)}, \tag{47}$$

  if there exists $\mathcal{I}$ such that $Q_2^{(i)} = \mathcal{I}$, $S_{\mathcal{J}\setminus\mathcal{I}} = \mathbf{1}$, $S_{\mathcal{K}\setminus\mathcal{J}} = \mathbf{0}$, and $S_{\mathcal{I}} \neq \mathbf{1}$, and $\hat{Y}_{i\to\mathcal{J}} = 0$ otherwise. The restriction can be interpreted as: side information is barely needed if $S_{\mathcal{I}} = \mathbf{1}$, the set $\mathcal{J}$ is selected rather than any other $\mathcal{U} \supset \mathcal{I}$ as the novel multicast set for $V_{l\to\mathcal{I}}$ only when the signal $V_{\mathcal{I}}$ is completely received by users in $\mathcal{J} \setminus \mathcal{I}$ whereas erased by destinations in $\mathcal{K} \setminus \mathcal{J}$.

A sketch of proof is provided in the Appendix 3.A. ■

The intuition behind the setting of the RV $\hat{Y}_{i\to\mathcal{J}}$ is the following. When a signal $X^{(i)}$ intended for a user group $\mathcal{I}$ is sent, if some of the users in $\mathcal{I}$ do not receive the signal[2] *and* some other users (denoted by $\mathcal{J} \setminus \mathcal{I}$) overhear it, then this signal becomes a side information for user group $\mathcal{J}$ and will be compressed and transmitted in phase j. This idea is inspired by the schemes in [27, 28]. The main difference is that here we only care about *what* side information to be shared among users in *which* user group, but do not deal with *how*. With the linear

[2]In erasure BC, the state feedback provides same amount of information to the transmitter as the output feedback does.

schemes in [27, 28], however, one has to decide how to combine different side information based on what each user already has and needs. In our scheme, these details are treated with the tools from distributed compression.

With the region in Corollary 3.2, the one-sidedly fair condition can be rewritten as $P(Q_2^{(1)} = i)\delta_i \geq P(Q_2^{(1)} = j)\delta_j$ if $i \leq j$. For simplicity, we define

$$\alpha_{\{k\}} = \alpha_1 P(Q_2^{(1)} = k),$$
$$\alpha_{\mathcal{J}} = \sum_{i=1}^{j-1} \alpha_j P(Q_1^{(j)} = i) P(Q_2^{(j)} = \mathcal{J}) = \alpha_j P(Q_2^{(j)} = \mathcal{J}).$$

Corollary 3.3. *For symmetric EBC or spatially independent EBC that satisfies the one-sidedly fair condition, the capacity region on the EBC defined by* (48) *can always be achieved with the proposed scheme for arbitrary input alphabet size* $|\mathcal{X}|$*,*

$$\mathcal{C}_{EBC} = \left\{ \begin{array}{c} \forall \pi(\mathcal{K}), (R_{\{\pi(1)\}}, \ldots, R_{\{\pi(K)\}}) \in \mathbb{R}_+^K, \\ \sum_{k=1}^K \frac{R_{\pi(k)}}{1 - \delta_{\{\pi(1),\cdots,\pi(k)\}}} \leq \log |\mathcal{X}|. \end{array} \right\}, \tag{48}$$

where $\pi(\mathcal{K}) = \pi(1), \cdots, \pi(K)$ *denotes a permutation of the elements in* $\mathcal{K}$*. In the definition of one-sidedly fair condition, we can substitute* $\pi(k)$ *for* k*.*

Proof. The converse can be found in [27, 28] with the standard outer bound techniques by creating an artificial degraded BC. A detailed proof on the achievability is provided in Appendix 3.B. ■

Remark 3.2. *Similar to [27, 28], we do not propose a uniform scheme as strong as to achieve three user capacity. In their works, a specific scheme was provided where the encoding incorporates the messages at distinct phases, e.g.,* M_1 *in first phase and* $M_{1\to\{2,3\}}$ *in second phase. Although the scheme balances carefully the length of subphases by merging the messages from different phases and obtains three user capacity, such a mixture across phases does not provide enough insights to be generalized to* $K > 3$ *case and complicates the coding procedure. Thus, we exclude the cross phase mixture in that the goal here is to find a general scheme that works for EBC and GBC and at the same time benefits from relatively simple evaluation both analytically and numerically.*

3.2.3 *Symmetric rate and DoF of fading GBC*

We apply the rate region in Theorem 3.1 to a Fading symmetric GBC and obtain the corollary below.

Corollary 3.4. *For the Symmetric Fading GBC, the proposed scheme achieves*

the following symmetric rate:

$$R_{sym} = \max_{\beta_i \geq 0} \left(K + \sum_{j=2}^{K} \binom{K}{j} \prod_{t=2}^{j} \frac{\sum_{l \leq t} b_{l,t}}{a_t} \right)^{-1} a_1 \tag{49}$$

where, for $t = 1, \ldots, K$,

$$a_t \triangleq \mathbb{E}_{\boldsymbol{H}_{\mathcal{T}}} \log\det\left(\boldsymbol{I} + \mathsf{SNR}\, \boldsymbol{H}_{\mathcal{T}}^H \boldsymbol{\Lambda}_t \boldsymbol{H}_{\mathcal{T}} \right) \tag{50a}$$

$$b_{l,t} \triangleq \mathbb{E}_{\boldsymbol{H}_1, \boldsymbol{H}_l} \log\det\left(\boldsymbol{I} + \frac{\mathsf{SNR}}{\beta_t} \boldsymbol{H}_l (\boldsymbol{I} + \mathsf{SNR}\, \boldsymbol{H}_1^H \boldsymbol{H}_1)^{-1} \boldsymbol{H}_l^H \right) \tag{50b}$$

with $\mathcal{T} \triangleq \{1\} \cup \{t+1, \ldots, K\}$, $\boldsymbol{\Lambda}_t \triangleq \mathrm{diag}\big(\boldsymbol{I}_{n_r}, \beta_t^{-1} \boldsymbol{I}_{(K-t)n_r}\big)$.

Proof. We use the following auxiliary RVs.

- Time sharing RVs $Q^{(j)} \triangleq (Q_1^{(j)}, Q_2^{(j)}) \in \mathcal{Q}^{(j)}$ with

$$p(q^{(j)}) = \binom{K}{j}^{-1}, \quad \text{if } q_1^{(j)} = j-1,\ q_2^{(j)} = \mathcal{J} \tag{51}$$

  and 0 otherwise. Owing to the channel symmetry, $Q^{(j)}$ is uniformly distributed.

- Gaussian distributed V's: for any $i = j-1$ and $\mathcal{J}$, we let $V_{i\to\mathcal{J}} \sim \mathcal{CN}(\boldsymbol{0}, \frac{P}{n_t} \boldsymbol{I}_{n_t})$. Note that in Gaussian channel, each user always has an observation regardless of the channel quality ($\mathrm{P}(\boldsymbol{H}_k = \boldsymbol{0}) = 0$). As such, the encoder can add whichever user to the multicast set in the next phase. In order to further simplify the region assessment, only $i = j-1$ is allowed.

- Input as deterministic function of V and Q:

$$\boldsymbol{X}^{(j)} = V_{Q_1^{(j)} \to Q_2^{(j)}}. \tag{52}$$

- Side information $\hat{Y}$ as noisy function of (X, S, Q), for any $i = j-1$ and $\mathcal{J}$,

$$\hat{\boldsymbol{Y}}_{i\to\mathcal{J}} = \boldsymbol{H}_{\mathcal{J}\setminus\mathcal{I}} \boldsymbol{X}^{(i)} + \hat{\boldsymbol{Z}}_{\mathcal{J}\setminus\mathcal{I}}, \tag{53}$$

  if $Q_2^{(i)} = \mathcal{I}$, and $\hat{\boldsymbol{Y}}_{i\to\mathcal{J}} = 0$ otherwise. The compression noise follows $\hat{\boldsymbol{Z}}_{\mathcal{J}\setminus\mathcal{I}} \sim \mathcal{CN}(\boldsymbol{0}, \beta_i \boldsymbol{I})$.

It is easily seen from (49) that for a given group of β's, the optimal choice of phase 1's normalized length is determined by

$$\alpha_1^* = \left(1 + \sum_{j=2}^{K} \frac{\binom{K}{j}}{K} \prod_{t=2}^{j} \frac{\sum_{l \leq t} b_{l,t}}{a_t} \right)^{-1} \tag{54}$$

Recall that the main difficulty on the optimization for Corollary 2.3 lies in lacking close form expressions on the allocated powers when P's are always

inside the expectation, which results from the choice of inputs. To palliate the complexity of optimizing numerically the symmetric rate in Section 3.4, we discard the superposed-style input $x = v_{12} + v_1 + v_2$ and let $x^{(j)} = v_{\mathcal{J}}$ behave in a time sharing manner with an identical power restriction. In place of tuning the SNR, the information in each message could be adjusted alternatively with phase duration.

A sketch of proof of (49) and (54) is provided in the Appendix 3.C. ■

As in the EBC case, we explain the intuition behind the choice of the side information $\hat{Y}_{i\to\mathcal{J}}$. When a signal $\boldsymbol{X}^{(i)}$ that is intended for some user group $\mathcal{I}$ is transmitted, user k, outside of the group, overhears the noisy version of the signal $\boldsymbol{H}_k\boldsymbol{X}^{(i)}$ which is useful to all users in group $\mathcal{I}$. Therefore, $\boldsymbol{H}_k\boldsymbol{X}^{(i)}$ becomes side information to be shared among the users in the group $\mathcal{J} \triangleq \{k\} \cup \mathcal{I}$. Since $\boldsymbol{H}_k\boldsymbol{X}^{(i)}$ can be costly to send, only a noisy (compressed) version, namely, $\boldsymbol{H}_k\boldsymbol{X}^{(i)} + \hat{\boldsymbol{Z}}_k$ is used instead. Again, as mentioned earlier, we do not need to explicitly deal with *how* to share the side information.

By setting $\beta_i = 1, \forall\, i$, in (49) and letting $\mathsf{SNR} \to \infty$, we can verify that the following optimal symmetric DoF can be achieved in the MISO case [2],

$$\mathsf{DoF}_{\mathrm{sym}} = \left(\sum_{k=1}^{K} \frac{1}{k}\right)^{-1}, \tag{55}$$

when $n_{\mathrm{t}} \geq K$ and $n_{\mathrm{r},1} = \cdots = n_{\mathrm{r},K} = 1$. Details are provided in the Appendix 3.D.

3.3 DESCRIPTION OF THE PROPOSED SCHEME

In the sequel, we present the detailed scheme that attains the region in Theorem 3.1. We divide the n-slot transmission into K phases with phase j of length n_j and define the normalized length of phase j as the proportion, $\alpha_j \triangleq \frac{n_j}{n}$ which is subject to $\sum_{j=1}^{K} \alpha_j = 1$. The detailed explanation of the role of auxiliary RVs are given in Section 3.2.

In phase 1, K independent source messages M_k are encoded and sent. In phase j ($j \geq 2$), upon receiving the channel state feedback of the previous phase, the encoder generates auxiliary multicast messages, $M_{i\to\mathcal{J}} \in \mathcal{M}_{i\to\mathcal{J}} \triangleq [1 : 2^{n_j R_{i\to\mathcal{J}}}]$. The message $M_{i\to\mathcal{J}}$ carries some useful side information on the original messages if $i = 1$ or on the auxiliary messages sent in the past phase i if $i \neq 1$ and intends for destinations in $\mathcal{J}$, where the choice of $\mathcal{J}$ is determined by the outdated CSI in phase i. At phase j, encoded signal $V_{i\to\mathcal{J}}$ is spread via X. Joint source-channel coding is employed.

Codebook generation

Fix the pmf as described in (39) and (40) and functions $x^{(j)}(\{v_{i\to\mathcal{J}}\}_{\mathcal{J}\subseteq\mathcal{K},|\mathcal{J}|=j}, q^{(j)})$ for all $j = 1, 2, \cdots, K$ that attain the region.

1. Before the beginning of phase $j \in \mathcal{K}$, randomly generate the time sharing

sequence according to $p(q^{j,n_j}) = \prod_{t=1}^{n_j} p(q_t^{(j)})$, where q^{j,n_j} is the sequence of time sharing RV of phase j with length n_j.

2. At the start of phase 1, randomly generate 2^{nR_k} independent sequences $\boldsymbol{v}_k$ which is drawn i.i.d. from

$$p(\boldsymbol{v}_k) = \prod_{i=1}^{n_1} p(v_{k,i}),$$

and index them as $\boldsymbol{v}_k(m_k)$ with $m_k \in [1 : 2^{nR_k}]$, $k \in \mathcal{K}$.

3. At the end of phase $i \in \{1, \ldots, K-1\}$, for each $j > i$ and each $\mathcal{J} \subseteq \mathcal{K}$ with $|\mathcal{J}| = j$, randomly generate $2^{n_i R_{i\to\mathcal{J}}}$ independent sequences $\hat{\boldsymbol{y}}_{i\to\mathcal{J}}$ and $2^{n_i R_{i\to\mathcal{J}}}$ independent sequences $\boldsymbol{v}_{i\to\mathcal{J}}$, drawn i.i.d. from

$$p(\hat{\boldsymbol{y}}_{i\to\mathcal{J}} \,|\, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)}) = \prod_{t=1}^{n_i} p(\hat{y}_{i\to\mathcal{J}}[t] \,|\, s^{(i)}[t], q^{(i)}[t]),$$
$$p(\boldsymbol{v}_{i\to\mathcal{J}}) = \prod_{t=1}^{n_j} p(v_{i\to\mathcal{J}}[t]),$$

respectively, and then index them as $\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(m_{i\to\mathcal{J}})$ and as $\boldsymbol{v}_{i\to\mathcal{J}}(m_{i\to\mathcal{J}})$, respectively, with $m_{i\to\mathcal{J}} \in [1 : 2^{n_i R_{i\to\mathcal{J}}}]$. The choice that the sequences $\hat{\boldsymbol{y}}_{i\to\mathcal{J}}$ and $\boldsymbol{v}_{i\to\mathcal{J}}$ are indexed with the same messages $M_{i\to\mathcal{J}}$ enables the joint source-channel encoding and the simultaneous decoding explained later.

Encoding

- At the start of phase 1, the encoder obtains the sequence $x^{(1)}$ by using the function that maps the broadcast messages $(\boldsymbol{v}_1(M_1), \cdots, \boldsymbol{v}_K(M_K))$ and the time-sharing RV $\boldsymbol{q}^{(1)}$ and then sends it out.

- At the end of phase i, $i = 1, \ldots, K-1$, and for each $j > i$ and each $\mathcal{J}$ with $|\mathcal{J}| = j$, given the causal state feedback of all the previous phases, the source searches for some auxiliary messages $M_{i\to\mathcal{J}}$ such that

$$(\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(M_{i\to\mathcal{J}}), \{\boldsymbol{v}_{l\to\mathcal{I}}\}_{l<i,\mathcal{I}\subset\mathcal{J}}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)}) \in \mathcal{T}_{\delta'}^{n_i}(\hat{Y}_{i\to\mathcal{J}}, \{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, S, Q)$$

are jointly typical. The error event of such encoding is the case that for any $m \in \{1, 2, \cdots, 2^{n_i R_{i\to\mathcal{J}}}\}$, the corresponding $\hat{y}$ sequence is not jointly typical with $(\{\boldsymbol{v}_{l\to\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)})$, which is given by

$$\mathcal{E}_1(i, \mathcal{J}) = \{(\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(m), \{\boldsymbol{v}_{l\to\mathcal{I}}\}_{l<i,\mathcal{I}\subset\mathcal{J}}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)}) \notin \mathcal{T}_{\delta'}^{n_i}(\hat{Y}_{i\to\mathcal{J}}, \{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, S, Q) \text{ for any } m \in \{1, 2, \cdots, 2^{n_i R_{i\to\mathcal{J}}}\}\}$$

It is worth clarifying that the error events defined in different chapter are specifically for the corresponding scheme in that chapter and we shall not mix them.

The source is capable of finding such sequence in the codebook with high

probability if

$$R_{i\to\mathcal{J}} \geq I(\hat{Y}_{i\to\mathcal{J}}; \{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}} \mid S^{(i)}, Q^{(i)}) + \epsilon_{n_i}. \tag{56}$$

The proposed scheme relies on the identical message $M_{i\to\mathcal{J}}$ to connect $V_{i\to\mathcal{J}}$ and $\hat{Y}_{i\to\mathcal{J}}$, which involve same amount of information, through joint source-channel encoding. In phase j, $j = 2, \ldots, K$, after learning the messages $M_{i\to\mathcal{J}}$, the source obtains the $\{\boldsymbol{v}_{i\to\mathcal{J}}(M_{i\to\mathcal{J}})\}_{i<j,\mathcal{J}}$ and then map the sequences obtained according to $\boldsymbol{q}^{(j)}$ into the signal sequence $\boldsymbol{x}^{(j)}$ and finally sends it.

Remark 3.3. *We emphasize that the $\hat{y}$'s are doing a distributed compression of correlated sources. However, it is not needed to guarantee the joint typicality of all descriptions together since decoding is performed separately. This is different from the standard distributed source coding framework, which excludes the use of generalized version of Markov lemma that is not possible when weak typicality tools are exploited.*

Decoding

We focus on the decoding procedure at user k while the constraints for the rest of the users follow similarly. At the end of phase K, a backward decoding is performed for the desired messages $\{M_{\mathcal{U}}\}_{\mathcal{U}\ni k}$ at destination k, which operates in a reverse phase order, that is, $M_{\mathcal{J}}$ are recovered before $M_{\mathcal{I}}$ whenever $i < j$. We assume that $M_{i\to\mathcal{J}}$ is the message chosen at the encoder, while the estimated message is $\hat{M}_{i\to\mathcal{J}}$. In the error events definition, we use $M_{i\to\mathcal{J}}$ to represent estimated message when $\hat{M}_{i\to\mathcal{J}} = M_{i\to\mathcal{J}}$.

1. For each set $\mathcal{J}$ that includes k and $i < j$ ($j = K, \ldots, 2$), given all the auxiliary messages $\hat{M}_{j\to\mathcal{U}}$ already decoded by then, the decoder looks for a unique $\hat{M}_{i\to\mathcal{J}}$ such that

$$\begin{aligned}(\boldsymbol{v}_{i\to\mathcal{J}}(\hat{M}_{i\to\mathcal{J}}), \boldsymbol{y}_k^{(j)}, \{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}(\hat{M}_{j\to\mathcal{U}})\}_{\mathcal{U}\supset\mathcal{J}}, \boldsymbol{s}^{(j)}, \boldsymbol{q}^{(j)}) \\ \in \mathcal{T}_{\delta'}^{n_j}(V_{i\to\mathcal{J}}, Y_k, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}, S, Q)\end{aligned}$$

   are jointly typical and that

$$(\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(\hat{M}_{i\to\mathcal{J}}), \boldsymbol{y}_k^{(i)}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)}) \in \mathcal{T}_{\delta'}^{n_i}(\hat{Y}_{i\to\mathcal{J}}, Y_k, S, Q)$$

   are jointly typical, *simultaneously*. For $j = K$, $\hat{Y}_{j\to\mathcal{U}}$ equals to zero by convention.

We define the error events for this decoding as follows.

$$\begin{aligned}
\mathcal{E}_2(i,\mathcal{J},k) &= \{(\boldsymbol{v}_{i\to\mathcal{J}}(M_{i\to\mathcal{J}}),\boldsymbol{y}_k^{(j)},\{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}},\boldsymbol{s}^{(j)},\boldsymbol{q}^{(j)}) \\
&\qquad \notin \mathcal{T}_{\delta'}^{n_j}(V_{i\to\mathcal{J}},Y_k,\{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}},S,Q)\} \\
\mathcal{E}_3(i,\mathcal{J},k) &= \{(\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(M_{i\to\mathcal{J}}),\boldsymbol{y}_k^{(i)},\boldsymbol{s}^{(i)},\boldsymbol{q}^{(i)}) \notin \mathcal{T}_{\delta'}^{n_i}(\hat{Y}_{i\to\mathcal{J}},Y_k,S,Q)\} \\
\mathcal{E}_4(i,\mathcal{J},k) &= \{(\boldsymbol{v}_{i\to\mathcal{J}}(\hat{M}_{i\to\mathcal{J}}),\boldsymbol{y}_k^{(j)},\{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}},\boldsymbol{s}^{(j)},\boldsymbol{q}^{(j)}) \\
&\qquad \in \mathcal{T}_{\delta'}^{n_j}(V_{i\to\mathcal{J}},Y_k,\{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}},S,Q) \text{ for some } \hat{M}_{i\to\mathcal{J}}\neq M_{i\to\mathcal{J}}\} \\
\mathcal{E}_5(i,\mathcal{J},k) &= \{(\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(\hat{M}_{i\to\mathcal{J}}),\boldsymbol{y}_k^{(i)},\boldsymbol{s}^{(i)},\boldsymbol{q}^{(i)}) \\
&\qquad \in \mathcal{T}_{\delta'}^{n_i}(\hat{Y}_{i\to\mathcal{J}},Y_k,S,Q) \text{ for some } \hat{M}_{i\to\mathcal{J}}\neq M_{i\to\mathcal{J}}\}
\end{aligned}$$

where $k\in\mathcal{J}$ and $|\mathcal{J}|=j$. $\mathcal{E}_2(i,\mathcal{J},k)$ stands for the event that the sequence $\boldsymbol{v}_{i\to\mathcal{J}}$ corresponding to the correct message $M_{i\to\mathcal{J}}$ is not jointly typical with $(\boldsymbol{y}_k^{(j)},\{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}},\boldsymbol{s}^{(j)},\boldsymbol{q}^{(j)})$. $\mathcal{E}_3(i,\mathcal{J},k)$ stands for the event that the sequence $\hat{\boldsymbol{y}}_{i\to\mathcal{J}}$ corresponding to the correct message $M_{i\to\mathcal{J}}$ is not jointly typical with $(\boldsymbol{y}_k^{(i)},\boldsymbol{s}^{(i)},\boldsymbol{q}^{(i)})$. $\mathcal{E}_4(i,\mathcal{J},k)$ represents the error event that for a message $\hat{M}_{i\to\mathcal{J}}$ other than the correct one ($M_{i\to\mathcal{J}}$), the corresponding sequence $\boldsymbol{v}_{i\to\mathcal{J}}$ is still joint typical with the information (about the message $M_{i\to\mathcal{J}}$ in phase j) that is available to user k, i.e., $(\boldsymbol{y}_k^{(j)},\{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}},\boldsymbol{s}^{(j)},\boldsymbol{q}^{(j)})$. $\mathcal{E}_5(i,\mathcal{J},k)$ represents the error event that for a wrong message $\hat{M}_{i\to\mathcal{J}}$, the corresponding sequence $\hat{\boldsymbol{y}}_{i\to\mathcal{J}}$ is still joint typical with $(\boldsymbol{y}_k^{(i)},\boldsymbol{s}^{(i)},\boldsymbol{q}^{(i)})$.

It can be shown that such $\hat{M}_{i\to\mathcal{J}}$ can be found and $\hat{M}_{i\to\mathcal{J}}=M_{i\to\mathcal{J}}$ with arbitrarily small error probability provided that

$$\begin{aligned}
n_iR_{i\to\mathcal{J}} \leq & n_jI(V_{i\to\mathcal{J}};Y_k^{(j)},\{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}\mid S^{(j)},Q^{(j)}) \\
& + n_iI(\hat{Y}_{i\to\mathcal{J}};Y_k^{(i)}\mid S^{(i)},Q^{(i)}) - n_i\epsilon_{n_i}. \qquad (57)
\end{aligned}$$

The error events analysis on the simultaneous decoding can be found in [78] which investigates the source coding problem in a distributed compression manner with a similar joint source-channel coding scheme.

2. Finally, the decoder recovers $M_{1\to\mathcal{U}}$ for all $\mathcal{U}\subset\mathcal{K}$ and $|\mathcal{U}|\geq 2$. Given these auxiliary messages, the decoder searches for a unique $\hat{M}_k$ such that

$$\begin{aligned}
&(\boldsymbol{v}_k(\hat{M}_k),\boldsymbol{y}_k^{(1)},\{\hat{\boldsymbol{y}}_{1\to\mathcal{U}}(M_{1\to\mathcal{U}})\}_{\mathcal{U}\ni k},\boldsymbol{s}^{(1)},\boldsymbol{q}^{(1)}) \\
&\qquad \in \mathcal{T}_{\delta'}^{n_1}(V_k,Y_k,\{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k},S,Q)
\end{aligned}$$

are jointly typical.

The corresponding error events are defined as

$$\begin{aligned}
\mathcal{E}_6(\mathcal{U},k) =& \{(\boldsymbol{v}_k(M_k),\boldsymbol{y}_k^{(1)},\{\hat{\boldsymbol{y}}_{1\to\mathcal{U}}(M_{1\to\mathcal{U}})\}_{\mathcal{U}\ni k},\boldsymbol{s}^{(1)},\boldsymbol{q}^{(1)}) \\
&\qquad \notin \mathcal{T}_{\delta'}^{n_1}(V_k,Y_k,\{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k},S,Q)\} \\
\mathcal{E}_7(\mathcal{U},k) =& \{(\boldsymbol{v}_k(\hat{M}_k),\boldsymbol{y}_k^{(1)},\{\hat{\boldsymbol{y}}_{1\to\mathcal{U}}(M_{1\to\mathcal{U}})\}_{\mathcal{U}\ni k},\boldsymbol{s}^{(1)},\boldsymbol{q}^{(1)}) \\
&\qquad \in \mathcal{T}_{\delta'}^{n_1}(V_k,Y_k,\{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k},S,Q) \text{ for some } \hat{M}_k\neq M_k\}
\end{aligned}$$

Then, $\hat{M}_k = M_k$ can be found with arbitrarily small error probability if

$$nR_k \leq n_1 I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \mid S^{(1)}, Q^{(1)}) - n_1\epsilon_{n_1}. \tag{58}$$

By letting the transmission length of each phase go to infinity ($n_i \to \infty$, $\forall i \in \mathcal{K}$), and applying the Fourier-Motzkin elimination to all the constraints mentioned in the Encoding and Decoding part, we obtain the achievable rate region given in Theorem 3.1.

Error events analysis

Before we start the analysis, we first define the event that none of $\mathcal{E}_1(i,\mathcal{J})$ occurs as

$$\mathcal{E}' = \bigcap_{i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K}} \mathcal{E}_1^c(i,\mathcal{J})$$
$$\mathcal{E}'' = \bigcap_{d\in\{2,3,4,5\},k\in\mathcal{K},i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K}} \mathcal{E}_d^c(i,\mathcal{J},k)$$

The error probability of the communication P_e is upper bounded by

$$\begin{aligned}
P_e \leq & \sum_{i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K}} Pr(\mathcal{E}_1(i,\mathcal{J})) + \sum_{k\in\mathcal{K},i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K},|\mathcal{J}|\neq 1} Pr(\mathcal{E}_2(i,\mathcal{J},k)\cap\mathcal{E}') \\
& + \sum_{k\in\mathcal{K},i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K},|\mathcal{J}|\neq 1} Pr(\mathcal{E}_3(i,\mathcal{J},k)\cap\mathcal{E}') \\
& + \sum_{k\in\mathcal{K},i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K},|\mathcal{J}|\neq 1} \sum_{\hat{M}_{i\to\mathcal{J}}=1}^{2^{n_i R_{i\to\mathcal{J}}}} Pr(\mathcal{E}_4(i,\mathcal{J},k)\cap\mathcal{E}') Pr(\mathcal{E}_5(i,\mathcal{J},k)\cap\mathcal{E}') \\
& + \sum_{k\in\mathcal{K},\mathcal{U}\ni k} Pr(\mathcal{E}_6(\mathcal{U},k)\cap\mathcal{E}''\cap\mathcal{E}') + \sum_{k\in\mathcal{K},\mathcal{U}\ni k} Pr(\mathcal{E}_7(\mathcal{U},k)\cap\mathcal{E}''\cap\mathcal{E}')
\end{aligned}$$

Due to the fact that $(\{\boldsymbol{v}_{l\to\mathcal{I}}\}_{l<i,\mathcal{I}\subset\mathcal{J}}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)})$ are joint typical as their codebook are generated independently, these sequences are jointly typical and thus the covering lemma (weak typical version, Lemma A.7) can be applied. Therefore, $Pr(\mathcal{E}_1(i,\mathcal{J}))$ tends to zero as n goes to infinity if $R_{i\to\mathcal{J}} > I(\hat{Y}_{i\to\mathcal{J}}; \{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}} \mid S^{(i)}, Q^{(i)})$. Since the total number of possible $(i,\mathcal{J})$ is finite, the term $\sum_{i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K}} Pr(\mathcal{E}_1(i,\mathcal{J}))$ tends to zero as $n \to \infty$ if $R_{i\to\mathcal{J}} > I(\hat{Y}_{i\to\mathcal{J}}; \{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}} \mid S^{(i)}, Q^{(i)})$ hold for all $i \in \mathcal{K}, \mathcal{J} \subseteq \mathcal{K}$.

For error events $\mathcal{E}_2(i,\mathcal{J},k)\cap\mathcal{E}'$, since $\hat{\boldsymbol{y}}_{j\to\mathcal{U}}$ is created at the encoder in such a way that it is correlated with $(\boldsymbol{s}^{(j)}, \boldsymbol{q}^{(j)}, \{\boldsymbol{v}_{i\to\mathcal{J}}\}_{\mathcal{J}\subset\mathcal{U}})$, then, $(\{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}, \boldsymbol{y}_k^{(i)}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)})$ follow the joint distribution conditioning on $\boldsymbol{v}_{i\to\mathcal{J}}$, that is,

$$p(\{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}, \boldsymbol{y}_k^{(i)}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)} \mid \boldsymbol{v}_{i\to\mathcal{J}}).$$

In this case, by applying the multivariate version of conditional typicality lemma (Lemma A.4), $P(\mathcal{E}_2(i,\mathcal{J},k))$ tends to zero as n goes to infinity.

For error events $\mathcal{E}_3(i,\mathcal{J},k)\cap\mathcal{E}^{'}$, $\boldsymbol{y}_k^{(i)}$ follows

$$p(\boldsymbol{y}_k^{(i)}|\hat{\boldsymbol{y}}_{i\to\mathcal{J}}\boldsymbol{s}^{(i)}\boldsymbol{q}^{(i)})=\prod_{t=1}^{n}p(\boldsymbol{y}_{k,t}^{(i)}|\hat{\boldsymbol{y}}_{i\to\mathcal{J},t}\boldsymbol{s}_t^{(i)}\boldsymbol{q}_t^{(i)}),$$

and then the conditional typicality lemma (Lemma A.4) yields that $P(\mathcal{E}_3(i,\mathcal{J},k))$ tends to zero as $n\to\infty$.

We associate the $\boldsymbol{v}_{i\to\mathcal{J}}$, ϕ, and $(\boldsymbol{y}_k^{(j)},\{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}},\boldsymbol{s}^{(j)},\boldsymbol{q}^{(j)})$ in the error event $\mathcal{E}_4(i,\mathcal{J},k)\cap\mathcal{E}^{'}$ with X^n, $\tilde{U}^n$, and $\tilde{Y}^n$ in the packing lemma (Lemma A.6), respectively. Then, along with the independence of codebook and following the same principle of the proof of packing lemma, we can obtain

$$Pr(\mathcal{E}_4(i,\mathcal{J},k)\cap\mathcal{E}^{'})\leq 2^{-n_j(I(V_{i\to\mathcal{J}};Y_k^{(j)},\{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}\,|\,S^{(j)},Q^{(j)})-\delta^{'})}.$$

With a similar reasoning, we can obtain that

$$Pr(\mathcal{E}_5(i,\mathcal{J},k)\cap\mathcal{E}^{'})\leq 2^{-n_i(I(\hat{Y}_{i\to\mathcal{J}};Y_k^{(i)}\,|\,S^{(i)},Q^{(i)})-\delta^{'})}.$$

Thus, we have

$$\begin{aligned}&\sum_{k\in\mathcal{K},i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K},|\mathcal{J}|\neq 1}\sum_{\hat{M}_{i\to\mathcal{J}}=1}^{2^{nR_{i\to\mathcal{J}}}}Pr(\mathcal{E}_4(i,\mathcal{J},k)\cap\mathcal{E}^{'})Pr(\mathcal{E}_5(i,\mathcal{J},k)\cap\mathcal{E}^{'})\\ \leq&\sum_{k\in\mathcal{K},i\in\mathcal{K},\mathcal{J}\subseteq\mathcal{K},|\mathcal{J}|\neq 1}2^{n_iR_{i\to\mathcal{J}}-n_jI(V_{i\to\mathcal{J}};Y_k^{(j)},\{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}\,|\,S^{(j)},Q^{(j)})-n_iI(\hat{Y}_{i\to\mathcal{J}};Y_k^{(i)}\,|\,S^{(i)},Q^{(i)})+\delta^{''}}\end{aligned}$$

Since the total number of $\mathcal{K}$ is finite, the above error probability tends to zero as long as

$$n_iR_{i\to\mathcal{J}}\leq n_jI(V_{i\to\mathcal{J}};Y_k^{(j)},\{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}}\,|\,S^{(j)},Q^{(j)})+n_iI(\hat{Y}_{i\to\mathcal{J}};Y_k^{(i)}\,|\,S^{(i)},Q^{(i)})-\delta^{''}$$

The error events $\mathcal{E}_6(\mathcal{U},k)\cap\mathcal{E}^{''}\cap\mathcal{E}^{'}$ and $\mathcal{E}_7(\mathcal{U},k)\cap\mathcal{E}^{''}\cap\mathcal{E}^{'}$ can be evaluated in the standard way, which leads to the constraint

$$nR_k\leq n_1I(V_k;Y_k^{(1)},\{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k}\,|\,S^{(1)},Q^{(1)})-n_1\epsilon_{n_1}.$$

Remark 3.4. *In the communication problem with lossy distributed source coding, block Markov schemes with Wyner-Ziv compression technique are often used. In these schemes, the source and channel codewords are linked via binning and joint typicality inside/across the block(s) with either joint source-channel coding, e.g., the scheme in Chapter 2, or with separate coding, e.g., the SW scheme [20]. When $K=2$, such schemes are efficient and have the advantage of being general. However, for beyond two users case, complexity augments dramatically as a large number of rate constraints are invoked for individual auxiliary indices. This rise is due to that both the bin and sub-bin indices are recognized as common messages and are required to be decodeable at certain intended users. Furthermore, jointly decoding all the messages in the block Markov structure will yields numerous conditions and hence loads an overwhelming burden on numerical evaluation. On the other hand, the successive decoding, though decreases analysis difficulty, might not exploit*

completely the auxiliary information. Therefore, we avoid using explicit binning and block Markov coding in the proposed approach. Rather, V and $\hat{Y}$ in distinct phases are attached directly by auxiliary message itself. Surprisingly, we observe that the source codeword $\hat{Y}$ and the corresponding channel codeword V are separate in (57)*. Nevertheless, from information's perspective, the two are not stand-alone [78].*

Remark 3.5. *It is worth mentioning that Marton coding is not used in our scheme, which makes the achievable region less general. As a matter of fact, the choice has been made due to our setup of K users. Introducing Marton coding would lead to a rate region intractable and hard to evaluate.*

3.4 NUMERICAL EXAMPLE

3.4.1 *Setup and references*

In this section, we consider the two-user and three-user Gaussian MISO channel with i.i.d. Rayleigh fading. We let $n_{\mathrm{t}} = K$ equal to 2 and 3 and evaluate the symmetric rate (49). The maximization over $\{\beta_i\}$ is done numerically while the maximization over $\{\alpha_j\}$ for each given set of $\{\beta_i\}$ is performed analytically and given by (54).

We add several reference schemes/scenarios as means of comparison:

1. the TDMA scheme that is optimal for no CSIT case following the reasoning in Chapter 2. The rate is given by

$$R_{\mathrm{TDMA}} \triangleq \mathbb{E}\log(1+\frac{P}{n_t}\boldsymbol{H}_1\boldsymbol{H}_1^H); \tag{59}$$

2. the generalized MAT (GMAT) scheme in [23]. In fact, the MAT symmetric rate can be attained by letting the precoder $\boldsymbol{w}$ equal to respective channel matrix $\boldsymbol{h}$ to reconstruct the overhead observations. In GMAT, the precoder is chosen to balance the alignment of interference and the enhancement of each signal. See [23] for details;

3. the analog scheme proposed by Maddah-Ali and Tse [2];

4. the modified MAT scheme with quantization in [18]. The symmetric rate is defined in the following proposition1.

Proposition 3.5 (Quantized Linear MAT)**.** *The symmetric rate tuple $(R,\cdots,R)$ is achievable in K-user MISO BC with state feedback if* (60) *is satisfied,*

$$R \leq DoF_{sym}K^{-1}\mathbb{E}\log\det(\boldsymbol{I}+\frac{P}{n_t}\boldsymbol{H}_{\mathcal{K}}\boldsymbol{H}_{\mathcal{K}}^H\tilde{\boldsymbol{N}}) \tag{60}$$

where $\tilde{\boldsymbol{N}} = \text{diag}\{1, 5, \cdots, 1+(j-1)(j+2), \cdots, 1+(K-1)(K+2)\}$ *and* DoF_{sym} *is given by* (55).

The Quantized MAT (QMAT) scheme transmits the quantization of linearly encoded signal in multiple phases. $1+(j-1)(j+2)$ determines the quantization noise variance at jth phase. In fact, $1+(j-1)(j+2)$ is the minimum compression noise variance so that the multicasting message for $\mathcal{J}$ can be recovered at each user through a $j \times 1$ channel. To this end, it corresponds to β_{j-1} in our scheme but with extra restriction.

5. the outer bound by applying the genie-aided argument as in [2, 18, 79]. The upper bound on the symmetric capacity is expressed by

$$C_{sym} \leq \mathsf{DoF}_{\text{sym}} R_{\text{TDMA}}. \tag{61}$$

For the sake of completeness, we attach the proof for the upper bound here.

Sketch of proof. By providing the output Y_k to users $l \in \{k+1, \cdots, K\}$, we generate a degraded BC whose capacity region cannot be enlarged by feedback. The capacity region of no CSIT defined by (62) can be achieved by superposition coding.

$$R_k \leq I(U_k; Y_1, \cdots, Y_k | \{U_l\}_{l<k}), \forall k \in \mathcal{K}, U_K = X. \tag{62}$$

We sum up the weighted rate as

$$\begin{aligned}
\sum_{k=1}^{K} \frac{R_k}{k} &\leq I(U_1; Y_1) + \cdots + \frac{1}{K} I(X; Y_1, \cdots, Y_K | \{U_l\}_{l<K}), \\
&\leq h(Y_1) + \cdots + \Bigg(-\frac{1}{k} h(Y_1, \cdots, Y_k | \{U_l\}_{l<k}) \\
&+ \frac{1}{k+1} h(Y_1, \cdots, Y_{k+1} | \{U_l\}_{l<k}) \Bigg) + \cdots - \frac{1}{K} h(Y_1, \cdots, Y_K | X), \\
&\leq I(X; Y_1),
\end{aligned} \tag{63}$$

where last inequality follows Lemma 2 in [18] and the fact that conditioning reduces entropy. Then, the upper bound is obtained by imposing $R_k = C_{sym}$. ■

The scheme in [25] is not taken into account as their rate is dominated by GMAT rate in i.i.d. Fading channel. In both MAT and GMAT schemes, we apply the precoding so that the power constraint is respected in each time slot.

3.4.2 *Numerical evaluation*

In Fig. 11 and Fig. 12, we show the comparison in two user case while three user BC is evaluated in Fig. 13 and Fig. 14. When $K = 2$, we include the achievable sum rate by the block Markov scheme from Chapter 2. The cyan curves in Fig 12 and Fig 14 (the third curve shown in the legends)denote the variants of the

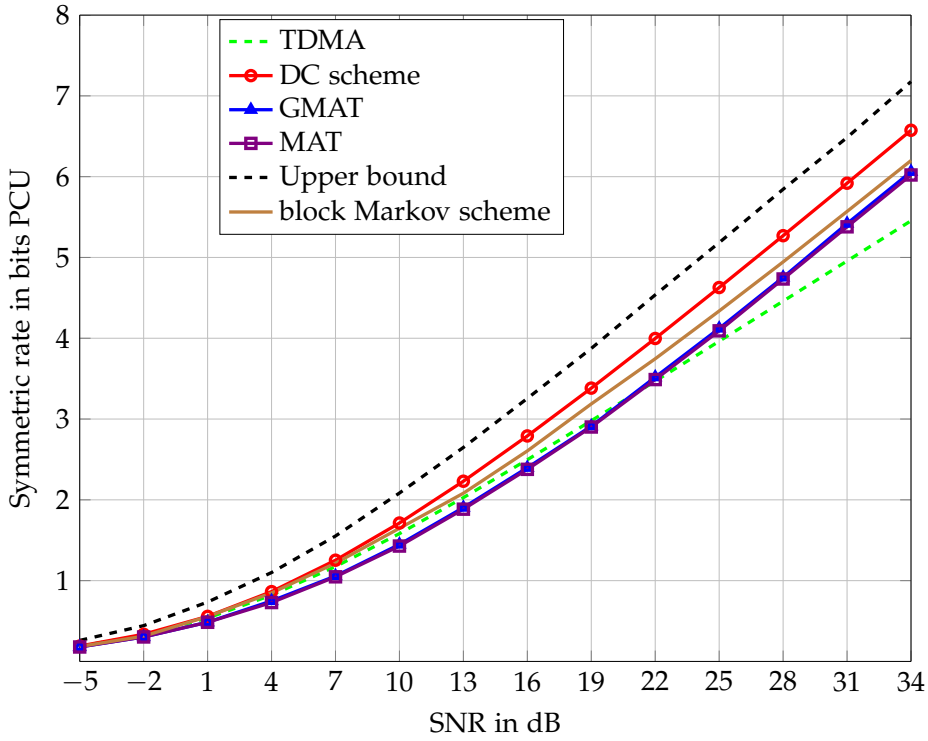

Figure 11: Symmetric rate in the symmetric two-user BC: proposed scheme versus MAT, GMAT.

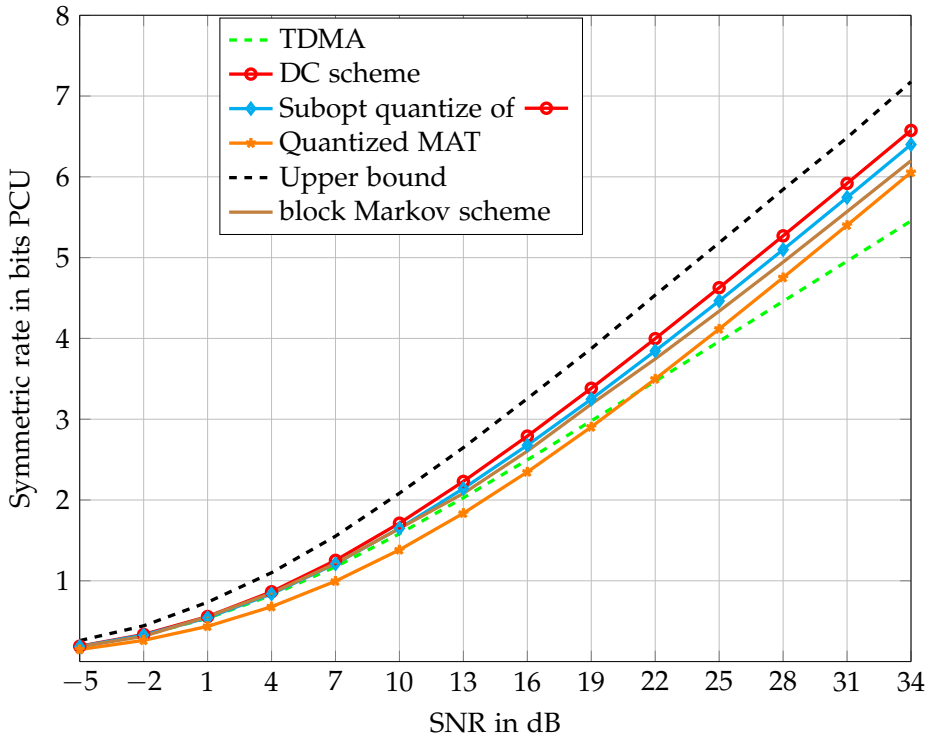

Figure 12: Symmetric rate in the symmetric two-user BC: proposed scheme, its non-optimal variant versus Quantized MAT.

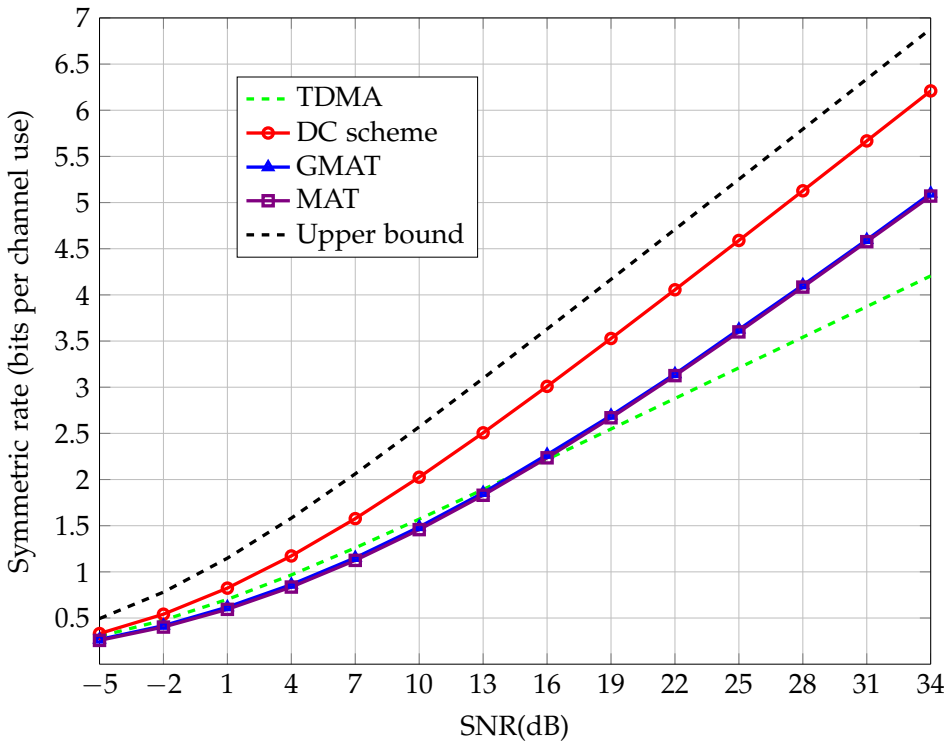

Figure 13: Symmetric rate in the symmetric three-user BC: proposed scheme versus MAT, GMAT.

proposed DC scheme for two/three-user case where the quantization noises β_j's are not optimized. Instead, we apply the same compression noise variance used in the QMAT scheme, that is, $\beta_{j-1} = 1 + (j-1)(j+2)$. These variants are included to make a fair comparison with the QMAT scheme given that QMAT scheme also uses the quantization, which would allow us to find the essential differences between our DC scheme and the QMAT scheme as shown in the following discussion.

We comment on the plots in the sequel.

- In [23], the authors investigated the achievable rate with a sub-optimal receiver/decoder and illustrated GMAT's gain over MAT. However, it is not true when we focus on the maximum achievable rate. From the plots, we observe that GMAT and MAT curves almost collapse in all SNR regime for $K = 2, 3$. As such, the performance improvement by designing carefully a precoder but still keeping the linear MAT structure is considered marginal. Despite their wish to balance between aligning interference and reinforcing signal, there does not seem to be a much better way to play with the precoder other than in MAT style. Aside from precoder, quantization is indeed an interesting direction to go with MAT scheme. Though not compared directly in the same plot, it can be readily seen that the quantization (QMAT) does ameliorate the MAT scheme especially when $K = 3$ in medium-to-high SNR regime. We observe that appreciable improvement occurs only in high SNR. This is due to the fact that QMAT scheme requires a minimum compression noise variance which strongly amplifies the ambiguity of the

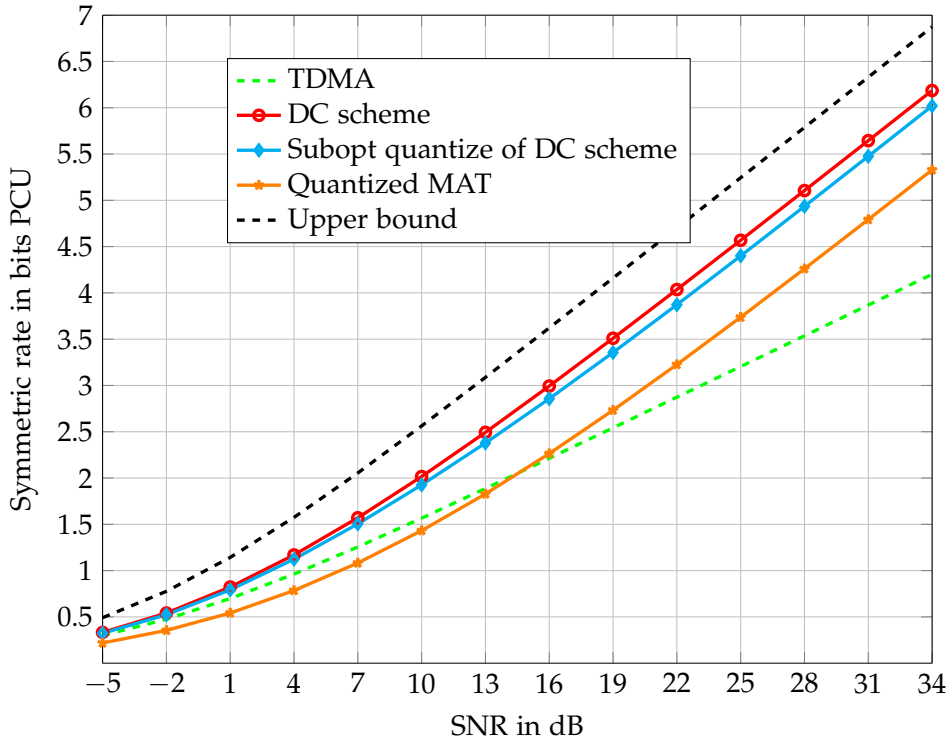

Figure 14: Symmetric rate in the symmetric three-user BC: proposed scheme, its non-optimal variant versus Quantized MAT.

common message in low SNR.

- Surprisingly, the MAT/GMAT/QMAT symmetric rates are dominated in low-to-medium SNR by the TDMA who cannot benefit from state feedback. This might result from that the coding is designed in an inefficient way for low SNR so that most of the interference is hardly mitigated. In the high SNR region, all these MAT-relevant schemes outperform the TDMA scheme, thanks to a higher DoF.

- When SNR is larger than 30dB, the slopes of the MAT, GMAT, QMAT and the proposed scheme approach the optimal one of outer bound, except for the MAT/GMAT strategies in three-user scenario as depicted in Fig. 13. Though it is not shown in the plot, MAT/GMAT obtain optimal slope in the regime with incredibly high SNR. Given the large gap of DoF-attaining SNR between $K = 2$ and 3 scenarios, a natural guess will be that the SNR thresholds for MAT/GMAT may grow with the network size K.

- In all SNR regions, the proposed scheme outperforms four reference schemes (TDMA, MAT, GMAT, QMAT) and has a non-negligible power gain over the MAT-related schemes, shown by the gap between curves. This gain becomes more appealing in the medium-to-low SNR region, in which the MAT/GMAT/QMAT schemes are not even better than the simple TDMA scheme, whereas our scheme can still take advantage of the CSI feedback to achieve a better performance. This is mainly due to the flexibility over the duration of each phase (time-slot) and the compression parameters as a

function of the SNR, which is not possible with the MAT/GMAT schemes as they are described in [23]. However, this may only account for part of the gain. To see this, we recall that the quantizations have already been allowed in QMAT scheme and the optimal duration of each phase for the proposed and the QMAT schemes in high SNR are fixed as in MAT. Provided that $\beta_{j-1} = 1 + (j-1)(j+2)$ in sub-optimal variant of the proposed scheme (of course with degraded performance), the gap between the variant (cyan curve that is the variant of DC scheme with sub-optimal quantization noise variance) and the QMAT strategy in medium-to-high SNR reveals that rate is enhanced actually with the coding gain by non-linear method (random coding).

3.5 SUMMARY AND CLOSING REMARKS

We proposed a general scheme for the state-dependent K-user BC with state feedback. The corresponding achievable region in general BC is derived followed by its application in erasure channel and symmetric Gaussian channel. Our scheme, with a proper choice of auxiliary RVs, achieves the capacity in the symmetric EBC and the spatially independent EBC satisfying one-sided fairness, and the optimal symmetric DoF in a Fading GBC. Different from existing schemes in the same setting, the main ingredient of the proposed non-linear scheme is distributed compression of side information. We demonstrated with numerical examples that a non-negligible rate gain can be obtained with the proposed scheme as compared to the TDMA scheme and the MAT-related schemes. Owing to its simplicity in K user case, the proposed region may be applied to other channels in a rather straightforward way. For instance, the adapted region can be evaluated in the multi-user BC with heterogeneous CSI feedback, say, the transmitter accesses to part of users' CSI, which is investigated in the Chapter 4.

3.A SKETCH OF PROOF OF COROLLARY 3.2

In the sequel, we first apply the RVs selected in the Section 3.2.2 and evaluate the quantities in (41a) and (41b), that is:

$$
\begin{aligned}
&I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \mid S^{(1)}, Q^{(1)}) \\
&\quad = P(Q_2^{(1)} = k)(1-\delta_{\mathcal{K}}) \log|\mathcal{X}|, && \text{(64a)} \\
&I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \mid S^{(j)}, Q^{(j)}) \\
&\quad = P(Q^{(j)} = (i,\mathcal{J}))(1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}}) \log|\mathcal{X}|, && \text{(64b)} \\
&\quad = P(Q_1^{(j)} = i) P(Q_2^{(j)} = \mathcal{J})(1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}}) \log|\mathcal{X}| && \text{(64c)} \\
&I(\{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \mid Y_k^{(i)}, S^{(i)}, Q^{(i)}) \\
&\quad = \sum_{\mathcal{I}\subset\mathcal{J}} P(Q_2^{(i)} = \mathcal{I}) I(V_{\mathcal{I}}; \hat{Y}_{i\to\mathcal{J}} \mid Y_k^{(i)}, S^{(i)}, Q^{(i)}) \\
&\quad = \sum_{\mathcal{I}\subset\mathcal{J}, \mathcal{I}\ni k} P(Q_2^{(i)} = \mathcal{I}, \hat{Y}_{i\to\mathcal{J}} \neq 0, Y_k^{(i)} = 0) H(V_{\mathcal{I}}) && \text{(64d)} \\
&\quad = \sum_{\mathcal{I}\subset\mathcal{J}, \mathcal{I}\ni k} P(Q_2^{(i)} = \mathcal{I}) \mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}, \mathcal{J}\setminus\mathcal{I}} \log|\mathcal{X}|. && \text{(64e)}
\end{aligned}
$$

where (64a) is interpret as, there are always some helpers unless none of the destinations has an observation on M_k; (64b), and (64e) are due to the choice of $\hat{Y}_{i\to\mathcal{J}}$ is a deterministic function of S, namely, $S_{\mathcal{J}\setminus\mathcal{I}} = \mathbf{1}$, $S_{\mathcal{K}\setminus\mathcal{J}} = \mathbf{0}$, and $S_{\mathcal{I}} \neq \mathbf{1}$; (64d) validates by the contradiction on S_k when $k \in \mathcal{J} \setminus \mathcal{I}$, i.e., $S_k = 0$ to ensure the mutual information $I(V_{\mathcal{I}}; \hat{Y}_{i\to\mathcal{J}} \mid Y_k^{(i)}, S^{(i)}, Q^{(i)})$ is not zero, and $S_{\mathcal{J}\setminus\mathcal{I}} = \mathbf{1}$ to guarantee side information $\hat{Y}_{i\to\mathcal{J}}$ is not null.

Using $\sum_{i=1}^{j-1} P(Q_1^{(j)} = i) = 1$ and applying (41b), we can eliminate $P(Q_1^{(j)} = i)$ with Fourier Motzkin elimination (FME) tool to obtain $K-1$ constraints on the α's, namely,

$$
\begin{aligned}
0 \le \alpha_j P(Q_2^{(j)} = \mathcal{J})(1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}}) \\
- \sum_{i=1}^{j-1} \alpha_i \sum_{\mathcal{I}\subset\mathcal{J}, \mathcal{I}\ni k} P(Q_2^{(i)} = \mathcal{I}) \mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}, \mathcal{J}\setminus\mathcal{I}}, && \text{(65)}
\end{aligned}
$$

for $j = 2, \ldots, K$, which fulfill as well $\sum_{j=1}^{K} \alpha_j = 1$. The α's constraints, together with (42a) achieved by (64a) and (41a), provide an achievable region in EBC as shown in Corollary 3.2.

3.B PROOF OF PROPOSITION 3.3

We recall some assumptions. $\delta_1 \ge \delta_2 \ge \cdots \ge \delta_K$ is assumed without loss of generality. For purpose of brevity, we use the notation $\alpha_{\mathcal{J}} = \alpha_j P(Q_2^{(j)} = \mathcal{J})$, $\alpha_{\{k\}} = \alpha_1 P(Q_2^{(1)} = k)$. With these notations defined, we can simplify the rate

region in Corollary 3.2 as

$$R_k \leq \alpha_{\{k\}}(1-\delta_{\mathcal{K}})\log|\mathcal{X}|, \tag{66a}$$

$$\alpha_{\mathcal{J}} \geq \max_{k\in\mathcal{J}}\left\{\sum_{k\in\mathcal{I}\subset\mathcal{J}}\frac{\mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k\},\mathcal{J}\setminus\mathcal{I}}}{1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}}}\alpha_{\mathcal{I}}\right\}, \tag{66b}$$

$$1 = \sum_{\mathcal{J}\subseteq\mathcal{K}}\alpha_{\mathcal{J}}. \tag{66c}$$

Note that with the structure of (66a), the optimal R_k is achievable only when the tuple $\{\alpha_{\{k\}}\}_{k\in\mathcal{K}}$ are maximized simultaneously. To maximize $\alpha_{\{k\}}$, we introduce a lemma below which provides optimal choice of $\alpha_{\mathcal{J}}$ for $\mathcal{J}\subseteq\mathcal{K}$.

Lemma C.6. *In the EBC, the $\alpha_{\mathcal{J}}$ achieves their optimal when the equalities of* (66b) *hold, as shown in* (67).

$$\alpha_{\mathcal{J}} = \max_{k\in\mathcal{J}}\left\{\sum_{k\in\mathcal{I}\subset\mathcal{J}}\frac{\mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k\},\mathcal{J}\setminus\mathcal{I}}}{1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}}}\alpha_{\mathcal{I}}\right\} \tag{67}$$

Proof. The proof is done by the following reasoning. We denote a tuple of selected α's with $A=\{\alpha_{\mathcal{J}},\forall\mathcal{J}\subset\mathcal{K}\}$ such that (66c) and (67) are satisfied except for one $\alpha_{\mathcal{J}'}$, where strict inequality of (66b) holds, as for a $\mathcal{J}'\neq\mathcal{J}$,

$$\alpha_{\mathcal{J}'} > \max_{k\in\mathcal{J}'}\left\{\sum_{k\in\mathcal{I}\subset\mathcal{J}'}\frac{\mathrm{P}_{\mathcal{K}\setminus\mathcal{J}'\cup\{k\},\mathcal{J}'\setminus\mathcal{I}}}{1-\delta_{\mathcal{K}\setminus\mathcal{J}'\cup\{k\}}}\alpha_{\mathcal{I}}\right\}.$$

By reducing the $\alpha_{\mathcal{J}'}$ to meet the equality of (66b) and adjusting all the other α's to make sure equalities still hold, we will have $\sum_{j=1}\alpha_{\mathcal{J}}=c<1$ for the reason that at least one α decreases. Then, we construct a new tuple of α's by multiplying each α in A by $\frac{1}{c}$, that is, $A^*=\{\frac{\alpha_{\mathcal{J}}}{c},\forall\alpha_{\mathcal{J}}\in A\}$. Therefore, every α in the new tuple A^* satisfies (66c) and (67). Since $c<1$, we have $\frac{\alpha_{\{k\}}}{c}>\alpha_{\{k\}}$ which assures a larger rate R_k. Thus, we define the optimal choice as $\alpha^*_{\{k\}}=\frac{\alpha_{\{k\}}}{c}$. The proof completes by applying many times the same reasoning, dealing with one inequality each time, to the case where arbitrary number of strict inequalities are activated. ■

With Lemma C.6, the optimal α's can be determined recursively. In order to further simplify the expression of optimal $\alpha_{\mathcal{J}}$, we consider two special class of EBC: the symmetric case and the spatially independent case that satisfies one-sided fairness. In such a setup, we demonstrate in Lemma C.7 that maximizing $k\in\mathcal{J}$ for $\alpha_{\mathcal{J}}$ is equivalent to let k equal to the minimum k inside $\mathcal{J}$.

Lemma C.7. *For a symmetric EBC or a spatially independent EBC that fulfills one-sided fairness, the optimal $\alpha_{\mathcal{J}}$ has the property described by,*

$$\alpha_{\mathcal{J}} = \alpha_{k^*_{\mathcal{J}},\mathcal{J}} = \sum_{k^*_{\mathcal{J}}\in\mathcal{I}\subset\mathcal{J}}\frac{\mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\},\mathcal{J}\setminus\mathcal{I}}}{1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}}}\alpha_{k^*_{\mathcal{I}},\mathcal{I}} \tag{68}$$

where $k^*_{\mathcal{J}}\triangleq\min_{k\in\mathcal{J}}k$. In addition, we define two sets $\mathcal{J}_1$ and $\mathcal{J}_2$ which verify $|\mathcal{J}_1|=|\mathcal{J}_2|=j\geq 2$ and $\mathcal{J}_1\setminus\{k^*_{\mathcal{J}_1}\}=\mathcal{J}_2\setminus\{k^*_{\mathcal{J}_2}\}$. For any pair of such sets and if $k^*_{\mathcal{J}_1}\leq k^*_{\mathcal{J}_2}$,

we have $\alpha_{\mathcal{J}_1} \geq \alpha_{\mathcal{J}_2}$ for symmetric case and $\delta_{k^*_{\mathcal{J}_1}} \alpha_{\mathcal{J}_1} \geq \delta_{k^*_{\mathcal{J}_2}} \alpha_{\mathcal{J}_2}$ for spatially independent channel with one-sided fairness, respectively.

Proof. Proof is done by the induction on j. Indeed, we can denote $\alpha_{k^*_{\mathcal{I}},\mathcal{I}}$ with $\alpha_{k^*_{\mathcal{J}},\mathcal{I}}$ in (68) without provoking any confusion provided that the summation is over all the $\mathcal{I}$ that satisfy $k^*_{\mathcal{J}} \in \mathcal{I} \subset \mathcal{J}$. In this case, the $k^*_{\mathcal{J}}$ is actually the minimum k in the concerned $\mathcal{I}$ as well. Note that $\alpha_{\mathcal{J}_1} \geq \alpha_{\mathcal{J}_2}$ and $\delta_{k^*_{\mathcal{J}_1}} \alpha_{\mathcal{J}_1} \geq \delta_{k^*_{\mathcal{J}_2}} \alpha_{\mathcal{J}_2}$ in two subclass channels are auxiliary results to prove (68) and will not be used beyond this lemma. To initial induction, we assume that $\alpha_{\mathcal{I}}$ achieves its maximum with $k^*_{\mathcal{I}}$ and the properties $\alpha_{\mathcal{I}_1} \geq \alpha_{\mathcal{I}_2}$ and $\delta_{k^*_{\mathcal{I}_1}} \alpha_{\mathcal{I}_1} \geq \delta_{k^*_{\mathcal{I}_2}} \alpha_{\mathcal{I}_2}$ are validated with the analogously defined $\mathcal{I}_1, \mathcal{I}_2, k^*_{\mathcal{I}_1}, k^*_{\mathcal{I}_2}$. We first present the proof in the symmetric channel and then in the spatially independent case with one-sided fairness.

In the symmetric channel, we assume without loss of generality that $\alpha_{\{1\}} \geq \alpha_{\{2\}} \geq \cdots \geq \alpha_{\{K\}}$ due to the symmetry in error probability. Since it always holds in the symmetric channel that $\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}} = \delta_{K-j+1}$ and $\mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\},\mathcal{J}\setminus\mathcal{I}} = \mathrm{P}_{K-j+1,j-i}$ where the error probability only concerns the corresponding set cardinality, (68) can be further simplified to

$$\alpha_{k^*_{\mathcal{J}},\mathcal{J}} = \frac{1}{1-\delta_{K-j+1}} \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subset \mathcal{J}} \alpha_{k^*_{\mathcal{I}},\mathcal{I}} \mathrm{P}_{K-j+1,j-i}. \tag{69}$$

As Lemma C.7 focus on $j \geq 2$ case, we start by validate $j = 2$ from the assumption $\alpha_{\{1\}} \geq \alpha_{\{2\}} \geq \cdots \geq \alpha_{\{K\}}$. Let us suppose that $\alpha_{\mathcal{J}} = \alpha_{\mathcal{J}_1} = \alpha_{\{t_1,t_2\}}$, $\alpha_{\mathcal{J}_2} = \alpha_{\{t_3,t_2\}}$ where $t_1 \leq t_3 \leq t_2$. Then, we can verify (68) and $\alpha_{\mathcal{J}_1} \geq \alpha_{\mathcal{J}_2}$.

$$\alpha_{\mathcal{J}} = \alpha_{\{t_1,t_2\}} = \frac{\mathrm{P}_{K-1,1}}{1-\delta_{K-1}} \max_{k\in\{t_1,t_2\}} \alpha_{\{k\}} = \frac{\mathrm{P}_{K-1,1}}{1-\delta_{K-1}} \alpha_{\{t_1\}} \tag{70a}$$

$$\alpha_{\mathcal{J}_1} = \alpha_{\{t_1,t_2\}} = \frac{\mathrm{P}_{K-1,1}}{1-\delta_{K-1}} \alpha_{\{t_1\}} \geq \frac{\mathrm{P}_{K-1,1}}{1-\delta_{K-1}} \alpha_{\{t_3\}} = \alpha_{\{t_3,t_2\}} = \alpha_{\mathcal{J}_2} \tag{70b}$$

We now assume that (68) and $\alpha_{\mathcal{J}_1} \geq \alpha_{\mathcal{J}_2}$ hold for any $\mathcal{J}, \mathcal{J}_1, \mathcal{J}_2 \subset \mathcal{K}$ with $|\mathcal{J}| = |\mathcal{J}_1| = |\mathcal{J}_2| = j \leq l-1$ $(3 \leq l \leq K)$, $k^*_{\mathcal{J}_1} \leq k^*_{\mathcal{J}_2}$, and $\mathcal{J}_1 \setminus \{k^*_{\mathcal{J}_1}\} = \mathcal{J}_2 \setminus \{k^*_{\mathcal{J}_2}\}$ and show that (68) and $\alpha_{\mathcal{L}_1} \geq \alpha_{\mathcal{L}_2}$ also hold for any $\mathcal{L}, \mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{K}$ with $|\mathcal{L}| = |\mathcal{L}_1| = |\mathcal{L}_2| = l$, $k^*_{\mathcal{L}_1} \leq k^*_{\mathcal{L}_2}$, and $\mathcal{L}_1 - \{k^*_{\mathcal{L}_1}\} = \mathcal{L}_2 - \{k^*_{\mathcal{L}_2}\}$.

Therefore, $\alpha_{\mathcal{L}}$ can be written as

$$\alpha_{\mathcal{L}} = \frac{\mathrm{P}_{K-l+1,l-j}}{1-\delta_{K-l+1}} \max_{k\in\mathcal{L}} \sum_{k\in\mathcal{J}\subset\mathcal{L}} \alpha_{\mathcal{J}} \tag{71a}$$

$$= \frac{\mathrm{P}_{K-l+1,l-j}}{1-\delta_{K-l+1}} \max \left\{ \sum_{k^*_{\mathcal{L}}\in\mathcal{J}\subset\mathcal{L}} \alpha_{\mathcal{J}}, \left\{ \sum_{k'\in\mathcal{J}'\subset\mathcal{L}} \alpha_{\mathcal{J}'} \right\}_{k'\neq k^*_{\mathcal{L}}} \right\} \tag{71b}$$

$$= \frac{\mathrm{P}_{K-l+1,l-j}}{1-\delta_{K-l+1}} \max_{\substack{k'\neq k^*_{\mathcal{L}'} \\ k'\in\mathcal{L}}} \left\{ \max \left\{ \sum_{k^*_{\mathcal{L}}\in\mathcal{J}\subset\mathcal{L}} \alpha_{\mathcal{J}}, \sum_{k'\in\mathcal{J}'\subset\mathcal{L}} \alpha_{\mathcal{J}'} \right\} \right\} \tag{71c}$$

$$= \frac{\mathrm{P}_{K-l+1,l-j}}{1-\delta_{K-l+1}} \sum_{k^*_{\mathcal{L}}\in\mathcal{J}\subset\mathcal{L}} \alpha_{\mathcal{J}} \tag{71d}$$

To prove (71d) is true, we consider four types of subsets of $\mathcal{L}$ depending on whether the subset contains $k^*_{\mathcal{L}}$ and k'. In particular, a subset that contains both $k^*_{\mathcal{L}}$ and k' appears in both terms inside the maximization of (71c) which yields $\mathcal{J} = \mathcal{J}'$, while a subset that contains neither $k^*_{\mathcal{L}}$ nor k' does not appear inside the inner maximization of (71c). We note that the rest of the subsets have either $k^*_{\mathcal{L}}$ or k' and there is always a mapping that projects a subset $\mathcal{J}$ including $k^*_{\mathcal{L}}$ into another specific subset $\mathcal{J}'$ including k' by substituting $k^*_{\mathcal{L}}$ for k', which is equivalent to $\mathcal{J} - \{k^*_{\mathcal{L}}\} = \mathcal{J}' - \{k'\}$. Thus, $\sum_{k' \in \mathcal{J}' \subset \mathcal{L}} \alpha_{\mathcal{J}'} \leq \sum_{k^*_{\mathcal{L}} \in \mathcal{J} \subset \mathcal{L}} \alpha_{\mathcal{J}}$ holds for any $k' \neq k^*_{\mathcal{L}}$ given $k^*_{\mathcal{L}} \leq k'$ and the property $\alpha_{\mathcal{J}_1} \geq \alpha_{\mathcal{J}_2}$ for $k^*_{\mathcal{J}_1} \leq k^*_{\mathcal{J}_2}$. Therefore, (71d) holds. The proof for the symmetric channel completes by

$$\sum_{\mathcal{I} \subset \mathcal{L}_1 \setminus \{k^*_{\mathcal{L}_1}\}} \alpha_{\mathcal{I} \cup \{k^*_{\mathcal{L}_1}\}} \geq \sum_{\mathcal{I} \subset \mathcal{L}_2 \setminus \{k^*_{\mathcal{L}_2}\}} \alpha_{\mathcal{I} \cup \{k^*_{\mathcal{L}_2}\}} \iff \alpha_{\mathcal{L}_1} \geq \alpha_{\mathcal{L}_2}. \tag{72}$$

where the inequality follows by identifying $\mathcal{J}_1 = \mathcal{I} \cup \{k^*_{\mathcal{L}_1}$, $\mathcal{J}_2 = \mathcal{I} \cup \{k^*_{\mathcal{L}_2}$ and the property by assumption: $\alpha_{\mathcal{J}_1} \geq \alpha_{\mathcal{J}_2}$ when $k^*_{\mathcal{J}_1} = k^*_{\mathcal{L}_1} \leq k^*_{\mathcal{L}_2} = k^*_{\mathcal{J}_2}$ and $\mathcal{J}_1 \setminus \{k^*_{\mathcal{J}_1}\} = \mathcal{I} = \mathcal{J}_2 \setminus \{k^*_{\mathcal{J}_2}\}$.

In a spatially independent channel, we assume that one-sided fairness $\alpha_{\{1\}}\delta_1 \geq \cdots \geq \alpha_{\{K\}}\delta_K$ is satisfied. In this channel, (68) is specified into

$$\alpha_{\mathcal{J}} = \alpha_{k^*_{\mathcal{J}},\mathcal{J}} = \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subset \mathcal{J}} \frac{\alpha_{k^*_{\mathcal{I}},\mathcal{I}} \mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\},\mathcal{J}\setminus\mathcal{I}}}{1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}}} = \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subset \mathcal{J}} \delta_{\mathcal{K}\setminus\mathcal{J}} \frac{\delta_{k^*_{\mathcal{J}}} \prod_{t \in \mathcal{J}\setminus\mathcal{I}} (1-\delta_t)}{1 - \delta_{\mathcal{K}\setminus\mathcal{J}}\delta_{k^*_{\mathcal{J}}}} \alpha_{k^*_{\mathcal{I}},\mathcal{I}},$$

where last equality is given by spatial independence of channel statistics. We initiate the induction by setting $\mathcal{J} = \{t_1, t_2\}$ ($t_1 \leq t_2$) and verifying the $j = 2$ case,

$$\begin{aligned} \alpha_{\{t_1,t_2\}} &= \delta_{\mathcal{K}\setminus\{t_1,t_2\}} \max_{k \in \{t_1,t_2\}} \left\{ \frac{(1-\delta_{t_2})\delta_{t_1}\alpha_{\{t_1\}}}{1 - \delta_{\mathcal{K}\setminus\{t_1,t_2\}}\delta_{t_1}}, \frac{(1-\delta_{t_1})\delta_{t_2}\alpha_{\{t_2\}}}{1 - \delta_{\mathcal{K}\setminus\{t_1,t_2\}}\delta_{t_2}} \right\} \\ &= \delta_{\mathcal{K}\setminus\{t_1,t_2\}} \frac{(1-\delta_{t_2})}{1 - \delta_{\mathcal{K}\setminus\{t_1,t_2\}}\delta_{t_1}} \delta_{t_1}\alpha_{\{t_1\}} \end{aligned} \tag{73}$$

where (73) follows the one-sided fairness $\delta_{t_1}\alpha_{\{t_1\}} \geq \delta_{t_2}\alpha_{\{t_2\}}$ and the channel assumption $\delta_{t_1} \geq \delta_{t_2}$ for $t_1 \leq t_2$. For two sets $\mathcal{J}_1 = \{t_1, t_2\}$ and $\mathcal{J}_2 = \{t_3, t_2\}$ with $t_1 \leq t_3 \leq t_2$, we can verify from the expression of $\alpha_{k^*_{\mathcal{J}},\mathcal{J}}$ that

$$\delta_{t_1}\alpha_{\{t_1\}} \geq \delta_{t_3}\alpha_{\{t_3\}} \iff \delta_{k^*_{\mathcal{J}_1}}\alpha_{\mathcal{J}_1} \geq \delta_{k^*_{\mathcal{J}_2}}\alpha_{\mathcal{J}_2} \tag{74}$$

We now assume that (68) and $\delta_{k^*_{\mathcal{J}_1}}\alpha_{\mathcal{J}_1} \geq \delta_{k^*_{\mathcal{J}_2}}\alpha_{\mathcal{J}_2}$ are true for any $\mathcal{J}, \mathcal{J}_1, \mathcal{J}_2 \subset \mathcal{K}$ with $|\mathcal{J}| = |\mathcal{J}_1| = |\mathcal{J}_2| = j \leq l-1$ ($3 \leq l \leq K$), $k^*_{\mathcal{J}_1} \leq k^*_{\mathcal{J}_2}$, and $\mathcal{J}_1 \setminus \{k^*_{\mathcal{J}_1}\} = \mathcal{J}_2 \setminus \{k^*_{\mathcal{J}_2}\}$ and show that (68) and $\delta_{k^*_{\mathcal{L}_1}}\alpha_{\mathcal{L}_1} \geq \delta_{k^*_{\mathcal{L}_2}}\alpha_{\mathcal{L}_2}$ hold for any $\mathcal{L}, \mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{K}$ with $|\mathcal{L}| = |\mathcal{L}_1| = |\mathcal{L}_2| = l$, $k^*_{\mathcal{L}_1} \leq k^*_{\mathcal{L}_2}$, and $\mathcal{L}_1 \setminus \{k^*_{\mathcal{L}_1}\} = \mathcal{L}_2 \setminus \{k^*_{\mathcal{L}_2}\}$.

Therefore, $\alpha_{\mathcal{L}}$ can be written as

$$\alpha_{\mathcal{L}} = \max_{k \in \mathcal{L}} \left\{ \sum_{k \in \mathcal{J} \subset \mathcal{L}} \delta_{\mathcal{K}\setminus\mathcal{L}} \frac{\delta_{\{k\}} \prod_{t \in \mathcal{L}\setminus\mathcal{J}} (1-\delta_t)}{1 - \delta_{\mathcal{K}\setminus\mathcal{L}}\delta_{\{k\}}} \alpha_{\mathcal{J}} \right\}$$

$$\alpha_{\mathcal{L}} = \max\left\{g_1(k^*_{\mathcal{L}}, \mathcal{J}), \left\{g_1(k', \mathcal{J}')\right\}_{k' \neq k^*_{\mathcal{L}}}\right\} \tag{75a}$$

$$= \max_{\substack{k' \neq k^*_{\mathcal{L}}, \\ k' \in \mathcal{L}}} \left\{\max\left\{g_1(k^*_{\mathcal{L}}, \mathcal{J}), g_1(k', \mathcal{J}')\right\}\right\} \tag{75b}$$

$$= \sum_{k^*_{\mathcal{L}} \in \mathcal{J} \subset \mathcal{L}} \delta_{\mathcal{K}\setminus\mathcal{L}} \frac{\delta_{k^*_{\mathcal{L}}} \prod_{t\in\mathcal{L}\setminus\mathcal{J}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}}\delta_{k^*_{\mathcal{L}}}} \alpha_{\mathcal{J}}. \tag{75c}$$

where we define $g_1(i, \mathcal{T}) \triangleq \sum_{i\in\mathcal{T}\subset\mathcal{L}} \delta_{\mathcal{K}\setminus\mathcal{L}} \frac{\prod_{t\in\mathcal{L}\setminus\mathcal{T}}(1-\delta_t)}{\delta_i^{-1}-\delta_{\mathcal{K}\setminus\mathcal{L}}} \alpha_{\mathcal{T}}$.

In fact, (75c) follows a similar reasoning by dividing the subsets of $\mathcal{L}$ into four groups as in the symmetric channel. Specifically, the subsets that contain neither k' nor $k^*_{\mathcal{L}}$ will not occur inside the inner maximization of (75b). The subsets appear in both terms of the inner maximization should include both k' and $k^*_{\mathcal{L}}$ ($\mathcal{J} = \mathcal{J}'$), which yields $\alpha_{\mathcal{J}} = \alpha_{\mathcal{J}'}$ and $\prod_{t\in\mathcal{L}\setminus\mathcal{J}}(1-\delta_t) = \prod_{t\in\mathcal{L}\setminus\mathcal{J}'}(1-\delta_t)$. Since $\delta_{\{k^*_{\mathcal{L}}\}} \geq \delta_{\{k'\}}$, the subsets $\mathcal{J}$ including both k' and $k^*_{\mathcal{L}}$ satisfy $\alpha_{\mathcal{L}_1} = g_1(k^*_{\mathcal{L}}, \mathcal{J}) \geq g_1(k', \mathcal{J}) = \alpha_{\mathcal{L}_2}$, where $\mathcal{L}_1 = \mathcal{L}$ and $k' = \min_{k\in\mathcal{L}_2} k$. We notice that the rest of subsets of $\mathcal{L}$ can always be regrouped in pairs and be written as $(\mathcal{J} = \{k^*_{\mathcal{L}}, \mathcal{T}\}, \mathcal{J}' = \{k', \mathcal{T}\})$ where $k', k^*_{\mathcal{L}} \notin \mathcal{T} \subset \mathcal{L}$. The reasoning for the equality of (75c) is completed by

$$\begin{aligned}
\sum_{k^*_{\mathcal{L}} \in \mathcal{J} \subset \mathcal{L}} \frac{\prod_{t\in\mathcal{L}\setminus\mathcal{J}}(1-\delta_t)}{\delta^{-1}_{\{k^*_{\mathcal{L}}\}}-\delta_{\mathcal{K}\setminus\mathcal{L}}} \alpha_{\mathcal{J}} &= \sum_{\mathcal{T}\subset\mathcal{L}} \frac{\prod_{t\in\mathcal{L}\setminus\mathcal{T}\setminus\{k^*_{\mathcal{L}}\}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}}\delta_{\{k^*_{\mathcal{L}}\}}} \delta_{\{k^*_{\mathcal{L}}\}} \alpha_{\mathcal{T}\cup\{k^*_{\mathcal{L}}\}} \\
&\overset{(a)}{\geq} \sum_{\mathcal{T}\subset\mathcal{L}} \frac{(1-\delta_{k'})\prod_{t\in\mathcal{L}\setminus\mathcal{T}\setminus\{k^*_{\mathcal{L}},k'\}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}}\delta_{\{k^*_{\mathcal{L}}\}}} \delta_{\{k'\}} \alpha_{\mathcal{T}\cup\{k'\}} \\
&\overset{(b)}{\geq} \sum_{\mathcal{T}\subset\mathcal{L}} \frac{(1-\delta_{k^*_{\mathcal{L}}})\prod_{t\in\mathcal{L}\setminus\mathcal{T}\setminus\{k^*_{\mathcal{L}},k'\}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}}\delta_{\{k'\}}} \delta_{\{k'\}} \alpha_{\mathcal{T}\cup\{k'\}} \\
&= \sum_{k'\in\mathcal{J}'\subset\mathcal{L}} \frac{\prod_{t\in\mathcal{L}\setminus\mathcal{J}'}(1-\delta_t)}{\delta^{-1}_{\{k'\}}-\delta_{\mathcal{K}\setminus\mathcal{L}}} \alpha_{\mathcal{J}'}
\end{aligned} \tag{76}$$

where (a) follows $\delta_{k^*_{\mathcal{J}_1}} \alpha_{\mathcal{J}_1} \geq \delta_{k^*_{\mathcal{J}_2}} \alpha_{\mathcal{J}_2}$ and (b) follows $\delta_{k^*_{\mathcal{L}}} \geq \delta_{k'}$.

Given $\mathcal{L}_1 \setminus \{k^*_{\mathcal{L}_1}\} = \mathcal{L}_2 \setminus \{k^*_{\mathcal{L}_2}\}$, the relation between $\delta_{k^*_{\mathcal{L}_1}} \alpha_{\mathcal{L}_1}$ and $\delta_{k^*_{\mathcal{L}_2}} \alpha_{\mathcal{L}_2}$ is determines by

$$\begin{aligned}
\delta_{k^*_{\mathcal{L}_1}} \alpha_{\mathcal{L}_1} &= \delta_{k^*_{\mathcal{L}_1}} \sum_{k^*_{\mathcal{L}_1} \in \mathcal{J} \subset \mathcal{L}_1} \delta_{\mathcal{K}\setminus\mathcal{L}_1} \frac{\delta_{\{k^*_{\mathcal{L}_1}\}} \prod_{t\in\mathcal{L}_1\setminus\mathcal{J}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}_1}\delta_{\{k^*_{\mathcal{L}_1}\}}} \alpha_{\mathcal{J}} \\
&\overset{(c)}{=} \sum_{\mathcal{T}\subset\mathcal{L}_1\setminus\{k^*_{\mathcal{L}_1}\}} \delta_{\mathcal{K}\setminus\mathcal{L}_1\cup\{k^*_{\mathcal{L}_1}\}} \frac{\prod_{t\in\mathcal{L}_1\setminus\mathcal{T}\setminus\{k^*_{\mathcal{L}_1}\}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}_1\cup\{k^*_{\mathcal{L}_1}\}}} \delta_{k^*_{\mathcal{L}_1}} \alpha_{\mathcal{T}\cup\{k^*_{\mathcal{L}_1}\}} \\
&= \sum_{\mathcal{T}\subset\mathcal{L}_2\setminus\{k^*_{\mathcal{L}_2}\}} \delta_{\mathcal{K}\setminus\mathcal{L}_2\cup\{k^*_{\mathcal{L}_2}\}} \frac{\prod_{t\in\mathcal{L}_2\setminus\mathcal{T}\setminus\{k^*_{\mathcal{L}_2}\}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}_2\cup\{k^*_{\mathcal{L}_2}\}}} \delta_{k^*_{\mathcal{L}_1}} \alpha_{\mathcal{T}\cup\{k^*_{\mathcal{L}_1}\}} \\
&\overset{(d)}{\geq} \sum_{\mathcal{T}\subset\mathcal{L}_2\setminus\{k^*_{\mathcal{L}_2}\}} \delta_{\mathcal{K}\setminus\mathcal{L}_2\cup\{k^*_{\mathcal{L}_2}\}} \frac{\prod_{t\in\mathcal{L}_2\setminus\mathcal{T}\setminus\{k^*_{\mathcal{L}_2}\}}(1-\delta_t)}{1-\delta_{\mathcal{K}\setminus\mathcal{L}_2\cup\{k^*_{\mathcal{L}_2}\}}} \delta_{k^*_{\mathcal{L}_2}} \alpha_{\mathcal{T}\cup\{k^*_{\mathcal{L}_2}\}} = \delta_{k^*_{\mathcal{L}_2}} \alpha_{\mathcal{L}_2}
\end{aligned}$$

where (c) holds by letting $\mathcal{J} = \mathcal{T} \cup \{k^*_{\mathcal{L}_1}\}$ and (d) follows $\delta_{k^*_{\mathcal{J}_1}} \alpha_{\mathcal{J}_1} \geq \delta_{k^*_{\mathcal{J}_2}} \alpha_{\mathcal{J}_2}$ for any $j < l$, $\mathcal{J}_1 \setminus \{k^*_{\mathcal{J}_1}\} = \mathcal{J}_2 \setminus \{k^*_{\mathcal{J}_2}\}$ and $k^*_{\mathcal{J}_1} \leq k^*_{\mathcal{J}_2}$. Thus, the proof is completed. ■

In the following, we show that the α's described in Lemma C.7 characterize the capacity region. First, we recall that $\mathrm{P}_{\mathcal{F},\mathcal{T}}$ can be extended, as shown below, according to the Lemma 10 in [28]. For any disjoint and non-empty sets $\mathcal{F}, \mathcal{T} \subseteq \mathcal{K}$, the $\mathrm{P}_{\mathcal{F},\mathcal{T}}$ is expressed as

$$\mathrm{P}_{\mathcal{F},\mathcal{T}} = \sum_{\mathcal{U} \subseteq \mathcal{T}} (-1)^{|\mathcal{U}|} \delta_{\mathcal{F}+\mathcal{U}} = \sum_{\mathcal{U} \subseteq \mathcal{T}} (-1)^{|\mathcal{U}|+1} (1 - \delta_{\mathcal{F}+\mathcal{U}}) \tag{77}$$

where last equality of (77) follows an extension on $\delta_{\mathcal{F}+\mathcal{U}}$ and $1 - \delta_{\mathcal{F}+\mathcal{U}}$) justified by binomial theorem

$$0 = \sum_{\mathcal{U} \subseteq \mathcal{T}} (-1)^{|\mathcal{U}|} = \sum_{\mathcal{U} \subseteq \mathcal{T}} (-1)^{|\mathcal{U}|} \delta_{\mathcal{F}+\mathcal{U}} + \sum_{\mathcal{U} \subseteq \mathcal{T}} (-1)^{|\mathcal{U}|} (1 - \delta_{\mathcal{F}+\mathcal{U}}).$$

Then, we substitute $\mathcal{K} \setminus \mathcal{J} \cup \{k\}, \mathcal{J} \setminus \mathcal{I}$ for $\mathcal{F}, \mathcal{T}$ in (77), respectively and insert the extension of $\mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k\},\mathcal{J}\setminus\mathcal{I}}$ into (68) which yields

$$\begin{aligned}
\alpha_{\mathcal{J}} &= (1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}})^{-1} \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subset \mathcal{J}} \alpha_{\mathcal{I}} \sum_{\mathcal{U} \subseteq \mathcal{J}\setminus\mathcal{I}} (-1)^{|\mathcal{U}|+1} (1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}\cup\mathcal{U}}) \\
&= (1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}})^{-1} \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subset \mathcal{J}} \\
&\qquad \alpha_{\mathcal{I}} \Bigg(- (1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}}) + \sum_{\phi \neq \mathcal{U} \subseteq \mathcal{J}\setminus\mathcal{I}} (-1)^{|\mathcal{U}|+1} (1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}\cup\mathcal{U}}) \Bigg) \\
&= \frac{\sum_{\phi \neq \mathcal{U} \subseteq \mathcal{J}\setminus\{k^*_{\mathcal{J}}\}} \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subseteq \mathcal{J}\setminus\mathcal{U}} \alpha_{\mathcal{I}} (-1)^{|\mathcal{U}|+1} (1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}\cup\mathcal{U}})}{1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}}} \\
&\qquad - \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subset \mathcal{J}} \alpha_{\mathcal{I}}
\end{aligned} \tag{78}$$

where we change the summation order over $\mathcal{I}$ and $\mathcal{U}$ to reach (78).

We attain the recursive relation between $\sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subseteq \mathcal{J}} \alpha_{\mathcal{I}}$ and $\sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subseteq \mathcal{J}\setminus\mathcal{U}} \alpha_{\mathcal{I}}$ by moving the first term in the right hand side of (78) to the left hand side, as shown below.

$$\sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subseteq \mathcal{J}} \alpha_{\mathcal{I}} = \frac{\sum_{\phi \neq \mathcal{U} \subseteq \mathcal{J}\setminus\{k^*_{\mathcal{J}}\}} (-1)^{|\mathcal{U}|+1} (1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}\cup\mathcal{U}}) \sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subseteq \mathcal{J}\setminus\mathcal{U}} \alpha_{\mathcal{I}}}{1 - \delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k^*_{\mathcal{J}}\}}} \tag{79}$$

It is not trivial to compute directly the $\sum_{k^*_{\mathcal{J}} \in \mathcal{I} \subseteq \mathcal{J}} \alpha_{\mathcal{I}}$ from the recursion with initial condition. Thus, we show that the α's satisfying the recursive relation, i.e., (79), also verify Lemma C.8.

Lemma C.8. *For a given k ($1 \leq k \leq K$), we have*

$$\sum_{k \in \mathcal{I} \subseteq \mathcal{S}_k} \alpha_{\mathcal{I}} = \frac{(1 - \delta_{\mathcal{K}}) \alpha_{\{k\}}}{(1 - \delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}})} \tag{80}$$

for any $\mathcal{S}_k$ such that $k \in \mathcal{S}_k \subseteq \{k, k+1, \cdots, K\}$.

Proof. The proof is by the induction on the cardinality of $\mathcal{S}_k$.

In particular, for arbitrary k and $|\mathcal{S}_k| = 1$ (say $\mathcal{S}_k = \{k\}$), (80) is validated

$$\alpha_{\{k\}} = \frac{(1-\delta_{\mathcal{K}})\alpha_{\{k\}}}{(1-\delta_{\mathcal{K}-\{k\}+\{k\}})} = \alpha_{\{k\}}.$$

We now assume that (80) holds for all $\mathcal{S}_k$ with $|\mathcal{S}_k| \leq l$ and show that it also holds for all $\mathcal{S}_k$ with $|\mathcal{S}_k| = l+1$. Note that $\mathcal{J}$ is a set whose minimum element is $k^*_{\mathcal{J}}$ and $\mathcal{S}_k$ is a set with minimum element being k. Since (79) is true for any $\mathcal{J}$, we can substitute k and $\mathcal{S}_k$ for $k^*_{\mathcal{J}}$ and $\mathcal{J}$, respectively, in (79) to get

$$\begin{aligned}
\sum_{k\in\mathcal{I}\subseteq\mathcal{S}_k} \alpha_{\mathcal{I}} &= \frac{\sum_{\phi\neq\mathcal{U}\subseteq\mathcal{S}_k\setminus\{k\}}(-1)^{|\mathcal{U}|+1}(1-\delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}\cup\mathcal{U}})\sum_{k\in\mathcal{I}\subseteq\mathcal{S}_k\setminus\mathcal{U}}\alpha_{\mathcal{I}}}{1-\delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}}} \\
&= \frac{\sum_{\phi\neq\mathcal{U}\subseteq\mathcal{S}_k\setminus\{k\}}(-1)^{|\mathcal{U}|+1}(1-\delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}\cup\mathcal{U}})\frac{(1-\delta_{\mathcal{K}})\alpha_{\{k\}}}{(1-\delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}\cup\mathcal{U}})}}{1-\delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}}} \\
&= \frac{(1-\delta_{\mathcal{K}})\alpha_{\{k\}}}{1-\delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}}} \sum_{\phi\neq\mathcal{U}\subseteq\mathcal{S}_k\setminus\{k\}}(-1)^{|\mathcal{U}|+1} \\
&= \frac{(1-\delta_{\mathcal{K}})\alpha_{\{k\}}}{1-\delta_{\mathcal{K}\setminus\mathcal{S}_k\cup\{k\}\cup\mathcal{U}}} \qquad (81)
\end{aligned}$$

where the second equality holds given the induction assumption and (81) follows binomial theorem. ■

With the property in Lemma C.8, we can finally prove the optimality. Summing up all the $\alpha_{\mathcal{J}}$ and using (66c), we can get the following constraint,

$$\begin{aligned}
1 = \sum_{\mathcal{J}\subseteq\mathcal{K}} \alpha_{\mathcal{J}} &= \sum_{k=1}^{K} \sum_{k\in\mathcal{I}\subseteq\{k,k+1,\cdots,K\}} \alpha_{\mathcal{I}} && (82a) \\
&= \sum_{k=1}^{K} \frac{(1-\delta_{\mathcal{K}})\alpha_{\{k\}}}{1-\delta_{\mathcal{K}\setminus\{k,k+1,\cdots,K\}\cup\{k\}}} && (82b) \\
&= \sum_{k=1}^{K} \frac{(1-\delta_{\mathcal{K}})\alpha_{\{k\}}}{1-\delta_{\{1,2,\cdots,k\}}} && (82c)
\end{aligned}$$

where (82b) comes from Lemma C.8. Additionally, we rewrite (66a) as

$$\frac{R_k}{(1-\delta_{\mathcal{K}})\log|\mathcal{X}|} \leq \alpha_{\{k\}}, \qquad (83)$$

and apply this inequality to (82c), which yields

$$\log|\mathcal{X}| \geq \sum_{k=1}^{K} \frac{R_k}{1-\delta_{\{1,2,\cdots,k\}}}. \qquad (84)$$

We substitute $\pi(1), \pi(2), \cdots, \pi(K)$ for $1, 2, \cdots, K$, respectively, to complete the proof of optimality.

$$\log|\mathcal{X}| \geq \sum_{\pi(k)=1}^{K} \frac{R_{\pi(k)}}{1-\delta_{\{\pi(1),\pi(2),\cdots,\pi(k)\}}}. \tag{85}$$

3.C SKETCH OF PROOF OF COROLLARY 3.4

In the symmetric Gaussian channel, since $R_{sym} = R_k$ holds for all k, we can focus on the R_1 without loss of generality. In particular, we apply the RVs choice given in Section 3.2.3, and compute the quantities in (41a) and (41b) specifically for user 1. Then, we have,

$$\begin{aligned}
&I(V_1; Y_1^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \,|\, S^{(1)}, Q^{(1)}) \\
&\quad = K^{-1} I(\boldsymbol{v}_k; \boldsymbol{H}_{\mathcal{K}}\boldsymbol{v}_k + \mathrm{diag}\{\boldsymbol{z}_1, \hat{\boldsymbol{z}}_{\{1,2\}\setminus\{1\}}, \cdots, \hat{\boldsymbol{z}}_{\{1,K\}\setminus\{1\}}\} \,|\, \boldsymbol{H}_{\mathcal{K}}) \\
&\quad = K^{-1}\mathbb{E}\log\det\left(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{H}_{\mathcal{K}}^H \boldsymbol{\Lambda}_1 \boldsymbol{H}_{\mathcal{K}}\right) = K^{-1} a_1 \qquad (86a) \\
&I(V_{i\to\mathcal{J}}; Y_1^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \,|\, S^{(j)}, Q^{(j)}) \\
&\quad = \binom{K}{j}^{-1} I(\boldsymbol{v}_{i\to\mathcal{J}}; \boldsymbol{H}_{\tilde{\mathcal{J}}}\boldsymbol{v}_{i\to\mathcal{J}} + \mathrm{diag}\{\boldsymbol{z}_1, \hat{\boldsymbol{z}}_{\mathcal{U}_{j+1}\setminus\mathcal{J}}, \cdots, \hat{\boldsymbol{z}}_{\mathcal{U}_K\setminus\mathcal{J}}\} \,|\, \boldsymbol{H}_{\tilde{\mathcal{J}}}) \\
&\quad = \binom{K}{j}^{-1} \mathbb{E}\log\det\left(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{H}_{\tilde{\mathcal{J}}}^H \boldsymbol{\Lambda}_j \boldsymbol{H}_{\tilde{\mathcal{J}}}\right) = \binom{K}{j}^{-1} a_j \qquad (86b) \\
&I(\{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \,|\, Y_1^{(i)}, S^{(i)}, Q^{(i)}) \\
&\quad = \sum_{l=1}^{i} \binom{K}{i}^{-1} I(\boldsymbol{v}_{\mathcal{I}}; \boldsymbol{H}_l\boldsymbol{v}_{\mathcal{I}} + \hat{\boldsymbol{z}}_{\mathcal{J}\setminus\mathcal{I}} \,|\, \boldsymbol{H}_1\boldsymbol{v}_{\mathcal{I}} + \boldsymbol{z}_1, \boldsymbol{H}_1, \boldsymbol{H}_l) \\
&\quad = \sum_{l=1}^{i} \binom{K}{i}^{-1} \mathbb{E}\log\det\left(\boldsymbol{I} + \frac{\mathsf{SNR}}{\beta_i}\boldsymbol{H}_l(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{H}_1^H\boldsymbol{H}_1)^{-1}\boldsymbol{H}_l^H\right) \\
&\quad = \sum_{l=1}^{i} \binom{K}{i}^{-1} b_{l,i} \qquad (86c)
\end{aligned}$$

where we define $\tilde{\mathcal{J}} \triangleq \{1\} \cup \{j+1, \cdots, K\}$, $|\mathcal{U}_{j+1}| = \cdots = |\mathcal{U}_K| = j+1$ and $(\mathcal{U}_l \setminus \mathcal{J}) \cap (\mathcal{U}_m \setminus \mathcal{J}) = \phi$ for $l, m \in \{j+1, \cdots, K\}$ and $l \neq m$. Note that the last two equalities hold only when $j = i+1$; for other i, j, both quantities are zero according to the values of $Q(i)$ and $Q(j)$. Thus, we could rewrite the Gaussian rate region as below.

$$R_{sym} \leq \alpha_1 K^{-1} a_1 \tag{87a}$$

$$0 \leq \alpha_j \binom{K}{j}^{-1} a_j - \alpha_{j-1} \binom{K}{j-1}^{-1} \sum_{l}^{j-1} b_{l,j-1} \tag{87b}$$

Further, we have

$$\alpha_t \geq \frac{\binom{K}{t}\sum_{l\leq t} b_{l,t}}{\binom{K}{t-1}a_t}\alpha_{t-1}. \tag{88}$$

For a given set of β and for a given SNR, a_j and $b_{l,j-1}$ are fixed. Hence, the maximum achievable rate of R_{sym} can be reached when α_1 is maximized. However, the selection of α_1 is as well affected by all the other α via $\sum_{j=1}^{K}\alpha_j = 1$. We apply $j-1$ times (88) for $t = j, j-1, \cdots, 1$. Then, we obtain

$$\alpha_j \geq \frac{\binom{K}{j}\sum_{l\leq j} b_{l,j}}{\binom{K}{j-1}a_j}\alpha_{j-1} \geq \cdots \geq \frac{\binom{K}{j}}{K}\prod_{t=2}^{j}\frac{\sum_{l\leq t} b_{l,t}}{a_t}\alpha_1 \tag{89}$$

With a similar reasoning in Lemma C.6, we note that each α_j should meet its lower bound to maximize α_1, which means that the above inequality provides the exact optimal ratio $\frac{\alpha_j}{\alpha_1}$ by letting all the equalities hold: suppose that we select an $A = \{\alpha_j, j \in \mathcal{K}\}$ such that $\sum_{j=1}^{K}\alpha_j = 1$ and only one of the above inequalities (89) holds strictly. Then, we can always reduce some of the α's values to make sure that all the equalities on α_j for all j, which would in turn lower the value of summation of α, i.e., $\sum_{j=1}^{K}\alpha_j = c < 1$. In this case, the optimal $\alpha_j^* = \frac{\alpha_j}{c}$ is chosen and no larger α_1 can be selected to satisfy the above inequalities. Thus, we have

$$\alpha_1^* = \left(1 + \sum_{j=2}^{K}\frac{\binom{K}{j}}{K}\prod_{t=2}^{j}\frac{\sum_{l\leq t} b_{l,t}}{a_t}\right)^{-1} \tag{90}$$

Inserting α_1^* to (87a), we obtain the optimal symmetric rate (49) in Gaussian case.

3.D PROOF OF EXPRESSION (55)

We let the compression noise to be $\beta_i = r_i \mathsf{SNR}^{d_i}$ where $r_i = O(\mathsf{SNR})$ is arbitrary but finite number when SNR tends to infinite. When $\mathsf{SNR} \to \infty$, if we assume $d_i > 0$, the side information are polluted by noise at comparable level as signal and $\mathsf{DoF}_{\text{sym}}$ tends to one; if $d_i < 0$, the cost of forwarding side information scales with SNR and again ruins the $\mathsf{DoF}_{\text{sym}}$. Therefore, we let $\beta_i = 1, \forall\ i$ to evaluate DoF.

One can verify, from (50a) and (50b), that, at high SNR,

$$a_t = |\mathcal{T}|\log\mathsf{SNR} + O(\log\mathsf{SNR}), \tag{91a}$$

$$b_{l,t} = \begin{cases} O(\log\mathsf{SNR}), & l = 1 \\ \log\mathsf{SNR} + O(\log\mathsf{SNR}), & l \neq 1 \end{cases} \tag{91b}$$

provided that $n_{\mathrm{t}} \geq K$. Note that $|\mathcal{T}| = K - t + 1$, we have

$$\mathsf{DoF}_{\mathrm{sym}} = \frac{R_{\mathrm{sym}}}{\log \mathsf{SNR}} \overset{\mathsf{SNR}\to\infty}{\longrightarrow} \left(K + \sum_{j=2}^{K} \binom{K}{j} \prod_{t=2}^{j} \frac{t-1}{K-t+1} \right)^{-1} K \tag{92a}$$

$$= \left(K + \sum_{j=2}^{K} \binom{K}{j} \binom{K-1}{j-1}^{-1} \right)^{-1} K \tag{92b}$$

$$= \left(\sum_{j=1}^{K} \frac{1}{j} \right)^{-1}. \tag{92c}$$

Part II

Broadcasting with Heterogeneous State Feedback

4

BC with State Feedback from partial Receivers

4.1 OVERVIEW

Remarkable gain by linear coding has been reported [43,44] in two user EBC with state feedback from single user. Inspired by that, we concentrate in the chapter on the memoryless state-dependent BC with *heterogeneous* state feedback. When both delayed CSI are accessible to the source, the capacity of two user EBC is acquirable with either non-linear schemes leveraging random coding, e.g., the aforementioned block Markov scheme and Distributed Compression scheme, or linear algorithms using bits *XOR* operations [27,28,55,80]. Nevertheless, this may not be true for the heterogeneous state feedback case. As such, we focus on non-linear coding and attempt to answer the question: whether non-linear scheme can have a superior performance to the existing linear methods?

Given that the DoF in this channel collapses to one as shown in [41], there seems to be little chance that the heterogeneous state feedback is useful for GBC. As such, we only specify the regions in EBC. The model of general BC and EBC are defined in Section 1.3. Since the complexities of the schemes scale with K, we consider mainly the two-user and three-user cases in the chapter. The main contributions of this work are listed below.

- First, we modify the block Markov scheme described in Chapter 2 and the Distributed Compression scheme proposed in Chapter 3 in order to make them suited to the case with insufficient state feedback. To be specific, the pmf, the codebook construction, and the encoding process of these schemes are forced to rest upon the knowledge of the available CSI. In addition, the corresponding inner bounds are derived for two user BC and K-user BC, respectively.

- We then specify both inner bounds in a two user EBC that has single user state feedback, where performance gain over no CSIT case is shown to be possible thanks to the intermittent coding opportunity to exchange overheard observations. Different from [43], we demonstrate that the claimed improvement verifies for a more general pmf, though not for all pmf in EBC. Further, rate region for certain specific three user EBC are established, where we reveal that adding a user–who does not feedback its outdated state to the source–into a two user system with delayed and/or perfect state could be beneficial.

- While comparing the proposed inner bound and that in [43,44] in the two

user symmetric spatially independent EBC, we prove that the non-linear scheme based region is strictly larger than the linear ones.[1] It is further illustrated that the enhancement results from that the non-linear scheme generates side information in a more efficient manner. Remarkably, one constraint of the proposed inner bound even coincides with one of the outer bound restrictions.

Despite that global state are assumed at the receivers' side for simplicity, the superior rate performance to be shown in the coming sections can also be achieved when limited users have access to the global state, e.g., in a two-user EBC, the user who does not feedback its state obtains full CSI while the other only has its proper CSI. This is due to the fact that the transmitter designs the codewords based on an algorithm relying on the state feedback from users in $\mathcal{J}$ and these users in $\mathcal{J}$ are also aware of the protocol. As long as the states $S_{\mathcal{J}}$ are spread locally in $\mathcal{J}$, the communications between source and users in $\mathcal{J}$ are coherent. Furthermore, we can even refine the state availability with this reasoning. Apart from the self state $S_{\{k\}}$, knowing $S_{\mathcal{J}}$ at the user k is sufficient to achieve the same rate performance for EBC in this chapter. Therefore, unlike the homogeneous feedback case, a scheme with heterogeneous state feedback can even reduce the cost and complexity at the destinations as lower demands on exchanging CSI are required.

However, in some other channels, this may not hold. Let us take Gaussian erasure BC [81] (also known as Bursty BC in [44]) as an example. The channel state includes the erasure events and the channel coefficients that follow discrete, continuous distributions, respectively. To avoid the degenerate rate or DoF performance, the channel coefficients are required by *all* nodes with/without delay.[2] Under such circumstance, the benefits become less attractive provided that the feedback cost is normally dominated by sending channel coefficients that are extracted from infinite alphabet.

The remainder of this chapter is organized as: in Section 4.2, we provide a region in two user BC which is achieved by tailored block Markov scheme, specify it in EBC and then compare it with some other inner bounds and an outer bound. In Section 4.3, we propose an adapted DC scheme and derive the corresponding region in subclass channels of multiple user BC/EBC. Details of computations are collected in appendices. Finally, the remarks in Section 4.4 close the chapter.

4.2 BLOCK MARKOV SCHEME IN D_1N_1 CASE

4.2.1 *Inner bounds and an outer bound*

In this part, we make a slight modification to the non-linear scheme proposed in the Chapter 2, which involves the block Markov coding and Wyner-Ziv compression,

[1] Extending the comparison to a more general pmf is possible. However, for brevity, we evaluate the inner bounds when $\delta_1 = \delta_2 = \delta$ and $\delta_{12} = \delta^2$.

[2] To avoid confusion, the heterogeneous CSIT in this chapter means that the states associated with *discrete* distribution are partially known to the source.

and derive the corresponding inner bound $\mathcal{R}_{BM}$ based on the adapted scheme.[3]

Theorem 4.1 (Inner bound $\mathcal{R}_{BM}$)**.** *The rate pair (R_1, R_2) is achievable for two-user BC in the DN (D_1N_1) case with a block Markov scheme if (R_1, R_2) are in the region $\mathcal{R}_{BM}$ defined by following constraints,*

$$\begin{aligned} R_1 \leq \min\{&I(V_1; Y_1\hat{Y}|V_{12}S_1S_2Q), I(V_1; Y_1\hat{Y}|V_{12}S_1S_2Q) + I(V_{12}; Y_1|S_1S_2Q) \\ &+ I(\hat{Y}; Y_1|V_{12}S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q)\}, \end{aligned} \tag{93a}$$

$$\begin{aligned} R_2 \leq \min\{&I(V_2; Y_2\hat{Y}|V_{12}S_1S_2Q), I(V_2; Y_2\hat{Y}|V_{12}S_1S_2Q) + I(V_{12}; Y_2|S_1S_2Q) \\ &+ I(\hat{Y}; Y_2|V_{12}S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q)\}, \end{aligned} \tag{93b}$$

for some pmf's verifying,

$$\begin{aligned} p(y_1, y_2, \hat{y}, x, v_1, v_2, v_{12}, s_1, s_2, q) =& p(y_1, y_2|x, s_1, s_2)p(\hat{y}|x, v_{12}, v_1, v_2, s_1) \\ &\times p(x|v_{12}, v_1, v_2, q)p(v_1|q)p(v_2|q)p(v_{12}|q)p(s_1)p(s_2)p(q). \end{aligned}$$

Sketch of proof. Since the scheme that attains the region above is, to some degree, similar to the scheme in Appendix of Chapter 2, we omit the detailed proof and highlight the distinctions. In particular, the pmf's component $p(\hat{y} \mid x, v_{12}, v_1, v_2, s, q)$, which represents how the side information is constructed from statistic point of view, is adjusted to $p(\hat{y} \mid x, v_{12}, v_1, v_2, s_1, q)$ such that the side information $\hat{y}$ depends on s_1 rather than (s_1, s_2). In addition, the auxiliary common message rate is set to $\hat{R} \triangleq I(\hat{Y}; V_1V_2X|V_{12}S_1)$. In the encoding part, the source searches for an index l_b which ensures that the corresponding sequence belongs to the joint typical set $\mathcal{T}_\delta^n(V_{12}V_1V_2S_1X\hat{Y})$. ■

We apply the Theorem 4.1 in EBC and obtain the corollary below.

Corollary 4.2 (Inner bound $\mathcal{R}'_{BM}$)**.** *For the two-user EBC with delayed CSI of only one receiver at the transmitter, the rate pair (R_1, R_2) is achievable as long as it is in the region $\mathcal{R}'_{BM}$ that verifies,*

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1, \tag{94a}$$

$$\frac{R_1}{\frac{1-\delta_2}{1-\delta_2+\delta_{12}}} + \frac{R_2}{1-\delta_2} \leq 1. \tag{94b}$$

Sketch of proof. We apply the RVs choices listed below to the Theorem 4.1.

- Input X is picked as a deterministic function of the encoded signals and auxiliary time-sharing RV (V_1, V_2, V_{12}, Q),

$$X = \begin{cases} X_1 = V_1 + V_{12}, & Q = 1, w.p.\ \alpha_1 \\ X_2 = V_2 + V_{12}, & Q = 2, w.p.\ \alpha_2. \end{cases} \tag{95}$$

[3]Since both block Markov strategy and Distributed Compression method will be investigated in the chapter, we denote the corresponding region with $\mathcal{R}$ and subscripts BM/DC, respectively.

where $\alpha_1 + \alpha_2 = 1$ with each α_k denoting the normalized duration when the signal V_k dedicated for user k is activated. Different from the linear scheme in [43] which transmits the independent messages and the side information sequentially, the transmitter in adapted block Markov scheme conveys the auxiliary common information and private message *in parallel* via X. We defer the analysis of the optimal V_{12} to the Appendix while the choice on V_k's pmf will be discussed after the proof.

- Side information $\hat{Y}$ is chosen as a deterministic function of (X, S, Q),

$$\hat{Y} = \begin{cases} X = X_1 = V_1 + V_{12}, & S_1 = 0, Q = 1, \\ X = X_2 = V_2 + V_{12}, & S_1 = 1, Q = 2, \\ 0, & \textit{otherwise}. \end{cases} \tag{96}$$

The opportunistic coding gain by such choices are explained as follows. When $Q = 1$, the encoder assigns the bits erased at user 1 to the side information, among which $1 - \frac{\delta_{12}}{\delta_1}$ proportion is known to the user 2 and thus can be decoded at Y_2 without cost. When $Q = 2$, the bits overheard at destination 1 are allocated to the side information, which is helpful to destination 2 whenever these bits are missed by it. The useful part accounts for $\frac{\delta_2 - \delta_{12}}{1-\delta_1}$ proportion of the overheard bits. The side information is encoded as if the desired but erased bits for user 1 are seen at user 2 while the unwanted but received bits at Y_1 are not observed by Y_2. Admittedly, this may not always be the case. Nonetheless, for those channel realizations correspond to such assumptions, side information is beneficial and offers an *opportunistic* gain.

The circumstantial proof is provided in Appendix 4.A. ■

Unlike the conventional schemes designed for homogeneous state feedback (D_2 case), the coding gain may not always be accessible with the heterogeneous state feedback. In particular, the rate R_2 is reduced when $(Q,\ S_1,\ S_2) = (1, 0, 0)$ as compared to that by conventional schemes since the private rate shall be limited when X_1 is absent at Y_2 to guarantee a successful delivery of auxiliary common message to receiver 2. When $(Q,\ S_1,\ S_2) = (2, 1, 1)$, though no additional cost is imposed on decoding side information, conveying the auxiliary common information is yet considered a waste given that user 2 has already obtained the desired signal. On the contrary, appreciable enhancement is indeed available when $(Q,\ S_1,\ S_2) = (1, 0, 1)$ and $(Q,\ S_1,\ S_2) = (2, 1, 0)$.

A corner point is named on-axis (*resp.* off-axis) corner point whenever its rates satisfy one of R_1, R_2 equals to zero (*resp.* $R_1 \neq 0$ and $R_2 \neq 0$). The on-axis corner points for (94a) are $(1 - \delta_1, 0)$ and $(0, 1 - \delta_{12})$ while for (94b) are $(\frac{1-\delta_2}{1-\delta_2+\delta_{12}}, 0)$ and $(0, 1 - \delta_2)$. By observation, the achievable rate of R_2 is restricted by $1 - \delta_2$ whereas the one of R_1 depends on the relation of $\delta_1, \delta_2, \delta_{12}$. In the sequel, we discuss the region for different δ's.

1. $\frac{1-\delta_2}{1-\delta_2+\delta_{12}} < 1 - \delta_1$:
   Under such a condition, (94a) is redundant and the obtained region is determined solely by (94b). In fact, this is a degenerate case where even

the single user rate $(1-\delta_1, 0)$ cannot be achieved. In this case, the source still forces the destination 2 to decode the side information while its proper channel is suffering from severe elimination. To make the auxiliary message decodable at user 2, the private message for user 1 is encoded in a rate lower than single user rate. As a matter of fact, the degenerate case corresponds to a scenario discussed in the scheme description in Appendix of Chapter 2 when (29) does not hold. To avoid the degraded performance, the adapted scheme here should be adjusted similarly, i.e., the source only encodes and sends V_1 and V_2 while each destination decodes simply with its proper channel output when $\frac{1-\delta_2}{1-\delta_2+\delta_{12}} < 1-\delta_1$.

2. $\frac{1-\delta_2}{1-\delta_2+\delta_{12}} \geq 1-\delta_1$:

In this case, each of the constraints (94a), (94b) serves as one boundary of the achievable region, where the two constraints together give rise to another corner point CP_1: (R_{1,CP_1}, R_{2,CP_1}), as defined below,

$$\begin{aligned} R_{1,CP_1} &\triangleq \frac{(\delta_2-\delta_{12})(1-\delta_1)}{(1-\delta_{12})-(1-\delta_1)(1-\delta_2+\delta_{12})}, \\ R_{2,CP_1} &\triangleq (1-\delta_{12})\frac{(1-\delta_2)-(1-\delta_1)(1-\delta_2+\delta_{12})}{(1-\delta_{12})-(1-\delta_1)(1-\delta_2+\delta_{12})}. \end{aligned}$$

The sum rate outperforms the single user rate as shown in the following,

$$\begin{aligned} R_{sum,CP_1} &= 1+\frac{\delta_{12}^2-\delta_1\delta_2+\delta_1\delta_{12}(\delta_2-\delta_{12})}{\delta_1+\delta_2-2\delta_{12}-\delta_1(\delta_2-\delta_{12})} \qquad (97) \\ &= 1-\delta_1+\frac{(\delta_1-\delta_{12})(1-\delta_2+\delta_{12})[\frac{1-\delta_2}{1-\delta_2+\delta_{12}}-(1-\delta_1)]}{(\delta_2-\delta_{12})+(1-\delta_2+\delta_{12})[\frac{1-\delta_2}{1-\delta_2+\delta_{12}}-(1-\delta_1)]} \geq 1-\delta_1 \\ &= 1-\delta_2+\frac{(\delta_2-\delta_{12})^2(1-\delta_1)}{(\delta_2-\delta_{12})+(1-\delta_2+\delta_{12})[\frac{1-\delta_2}{1-\delta_2+\delta_{12}}-(1-\delta_1)]} \geq 1-\delta_2. \end{aligned}$$

Moreover, one can easily verify that the corner point CP_1 is achievable by letting,

$$\begin{aligned} H(V_1) &= 1, \\ H(V_2) &= \frac{(1-\delta_2)-(1-\delta_1)(1-\delta_2+\delta_{12})}{\delta_1-\delta_{12}}, \\ \alpha_1 &= \frac{(\delta_2-\delta_{12})(1-\delta_1)}{(1-\delta_{12})-(1-\delta_1)(1-\delta_2+\delta_{12})}, \\ \alpha_2 &= \frac{\delta_1-\delta_{12}}{(1-\delta_{12})-(1-\delta_1)(1-\delta_2+\delta_{12})}. \end{aligned}$$

where $0 \leq \alpha_1 \leq 1,\ 0 \leq \alpha_2 \leq 1,\ 0 \leq h(V_2) \leq 1$ are satisfied. Although the entropy $h(V_k)$ and the normalized duration α_k that attain CP_1 are not unique, the value of $h(V_1)$ and $h(V_2)$, differing from the linear schemes, shall not be set to one at the same time. Another simple observation is that, in contrast to the D_2 case, the normalized transmission lengths α's are not equal.

In order to better understand the EBC capacity in D_1N_1 case, we include an outer bound as a reference as follows.

Theorem 4.3 (Outer bound). *Any achievable rate pair (R_1, R_2) for the EBC with single user state feedback (D_1N_1) must satisfy*

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1, \tag{98a}$$

$$\frac{R_1}{1-\delta_{12}} + \frac{R_2}{1-\delta_2} \leq 1. \tag{98b}$$

Sketch of proof. In fact, the above region describes the capacity of EBC with state feedback from both receivers (D_2 case). The capacity of D_1N_1 case is outer bounded by the capacity when more information is available at the transmitter. The related proof on this capacity is quite standard and thus omitted in the thesis. Both the achievability and the converse are provided in [27, 28] with details. ■

With respect to the two user EBC, the capacity regions of D_2 case and of N_1P_1 case could be considered as outer bounds on D_1N_1's capacity. The N_1P_1's capacity for spatially independent symmetric EBC collapses with the D_1P_1 capacity as demonstrated in [43], where only the erasure probability for user 2, namely δ_2, is involved in both converse and achievability ($R_1 \leq 1-\delta_1$ is single user rate obtained by TDMA strategy). Hence, we can extend the capacity regions of N_1P_1 and D_1P_1 cases to the general two-user EBC, as shown below.

$$\mathcal{C}_{N_1P_1} = \mathcal{C}_{D_1P_1} = \left\{ R_1, R_2 \in \mathbb{R}_+ \middle| \begin{array}{l} R_1 + \frac{R_2}{1-\delta_2} \leq 1, \\ \frac{R_1}{1-\delta_1} + 0 \cdot R_2 \leq 1 \end{array} \right\} \tag{99}$$

It can be easily shown that $\mathcal{C}_{N_1P_1}$ encompasses the region in Theorem 4.3 while the rate pair $(1-\delta_1, \delta_1(1-\delta_2))$ inside $\mathcal{C}_{N_1P_1}$ is not attainable by the schemes in D_2 case. Accordingly, (98a), (98b) are preferred as a tighter outer bound.

Remark 4.1. *Finding an asymmetric outer bound for D_1N_1 case is not in the scope of this thesis in that single-letterizing a region when the channel input is dependent on feedback (and thus on previous channel outputs) is quite involved.*

We compare the inner bound (94a) and (94b) with the outer bound in Theorem 4.3 in the non-degenerate case. A simple inspection reveals that (94a) coincides with (98a) whereas (94b) provides strictly smaller rate than (98b). A possible reason for the collapse of one boundary in the inner and outer bound will be that under certain circumstances, the EBC formulates a virtual "pipe" from user 1 to 2 with the help of the single user state so that the EBC is degraded. The scheme is designed such that the side information is always useful to receiver 1 when $Q = 1$ and is always cost free at Y_1 when $Q = 2$.

4.2.2 *Inner bounds analysis: non-linear versus linear coding*

To evaluate the inner bound $\mathcal{R}_{BM}^{'}$, we introduce two achievable regions by *linear* schemes based on *XOR* operations on bits, which are described in [44] and [43] and denoted with $\mathcal{R}_{LNC_1}$ and $\mathcal{R}_{LNC_2}$, respectively. To make a fair comparison, we consider a symmetric spatially independent EBC which is used by [43, 44], determined succintly by δ.

$$\mathcal{R}_{LNC_1} \triangleq \left\{ (R_1, R_2) \in \mathbb{R}_+ \left| \begin{array}{c} \frac{R_1}{\frac{1-\delta}{1-\delta^2+\delta^3}} + \frac{R_2}{1-\delta} \leq 1, \\ \frac{R_1}{1-\delta} + \frac{R_2}{(1+\delta-\delta^2)(1-\delta)} \leq 1 \end{array} \right. \right\}, \tag{100a}$$

$$\mathcal{R}_{LNC_2} \triangleq \left\{ (R_1, R_2) \in \mathbb{R}_+ \left| \begin{array}{c} \frac{R_1}{\frac{1-\delta}{1-\delta^2+\delta^3}} + \frac{R_2}{1-\delta} \leq 1, \\ \frac{R_1}{1-\delta} + \frac{R_2}{\frac{1+\delta-2\delta^2+\delta^3}{1-\delta}} \leq 1 \end{array} \right. \right\}. \tag{100b}$$

$\mathcal{R}_{LNC_1}$ is actually defined and proved explicitly in Section 1.3. Therefore, we recall briefly the principle of the re-transmission schemes that obtain the off-axis corner points of both regions: the source transmits over T time slots which are divided into three phases. In the first two phases, uncoded bits of the private information that is intended for user k are transmitted in phase k with duration $\alpha_k T$. In the third phase, side information that both receivers might be interested in is sent according to S_1. To be specific, the source sends the *XOR*ed bits of the $\alpha_1 T\delta$ bits of V_1 missed by Y_1 in phase 1 and $\alpha_2 T(1-\delta)$ bits of V_2 collected by Y_1 in phase 2. The two regions $\mathcal{R}_{LNC_1}$, $\mathcal{R}_{LNC_2}$ follow different choice of α's. In particular, $\delta\alpha_1 = (1-\delta)^2\alpha_2$ and $\alpha_3 = (1-\delta)\alpha_2$ are selected for $\mathcal{R}_{LNC_2}$ whereas the α's for $\mathcal{R}_{LNC_1}$ are chosen such that the number of erased bits at Y_1 in phase 1 matches that of collected bits at Y_1 in phase 2 and phase 2, 3 are of equal length, i.e., $\delta\alpha_1 = (1-\delta)\alpha_2$ and $\alpha_2 = \alpha_3$.

Note that $(\delta_1 = \delta_2 = \delta,\ \delta_{12} = \delta^2)$ falls into non-degenerate case for $\mathcal{R}_{BM}^{'}$. Analogously, it is also a non-degenerate case for $\mathcal{R}_{LNC_1}$ and $\mathcal{R}_{LNC_2}$ since one can easily verify that the on-axis corner points $(0, 1-\delta)$ and $(1-\delta, 0)$ are attainable by the linear *XOR* schemes. It is worth mentioning that for $\delta_1 \leq \delta_2$, a region with maximum sum rate being $(1-\delta_1) + \alpha_2[(1-\delta_2)(1+\delta_2-\delta_2^2) - (1-\delta_1)]$ is achievable if we extend the linear schemes to the asymmetric spatially independent EBC according to a similar principle, that is, *XOR* the erased and collected bits in different phases. Specifically, the corresponding maximum sum rate is not larger than the individual rate $1-\delta_1$ when $(1-\delta_2)(1+\delta_2-\delta_2^2) \leq (1-\delta_1)$. Therefore, even though it is not stated in references [43, 44], the sum rate by linear schemes may collapse to single user rate as well when extended to asymmetric channels.

In the symmetric spatially independent scenario, the relation between $\mathcal{R}_{LNC_1}$ and $\mathcal{R}_{LNC_2}$ is decided by the off-axis corner points as both regions have common on-axis corner points. Provided that the first condition in $\mathcal{R}_{LNC_1}$ is the same to $\mathcal{R}_{LNC_2}$, what remains is to verify that the corner point $((1-\delta)^2,\ (1-\delta)\delta + (1-\delta)^3\delta^2)$ of $\mathcal{R}_{LNC_2}$ satisfies the second condition of $\mathcal{R}_{LNC_1}$, that is, $\frac{(1-\delta)^2}{1-\delta} + \frac{(1-\delta)\delta+(1-\delta)^3\delta^2}{(1+\delta-\delta^2)(1-\delta)} = 1 - \frac{\delta^3(1-\delta)}{1+\delta-\delta^2} \leq 1$. Hence, we have $\mathcal{R}_{LNC_2} \subseteq \mathcal{R}_{LNC_1}$.

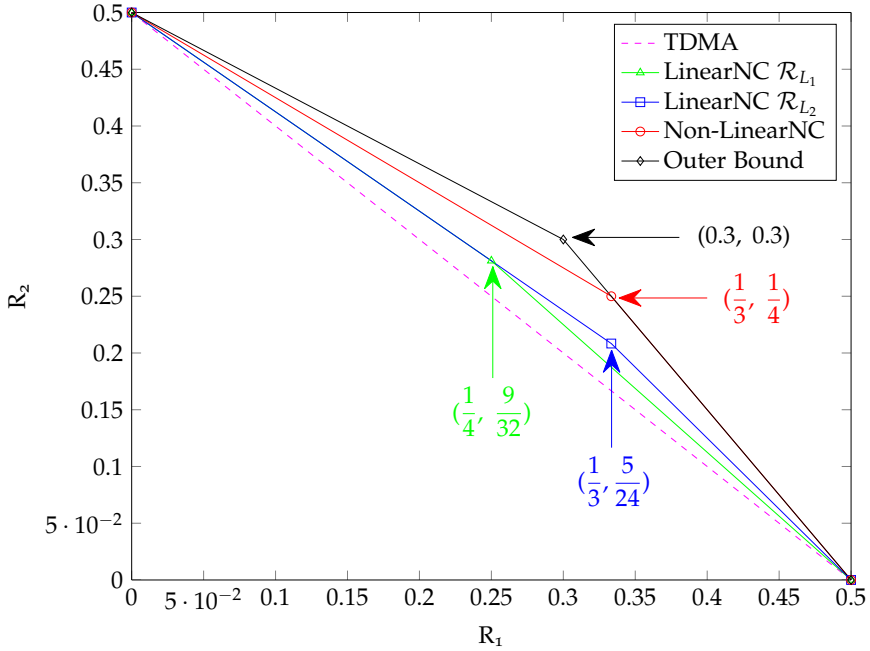

Figure 15: Two-user EBC in D_1N_1 case: $\delta_1 = \delta_2 = \frac{1}{2}$ and $\delta_{12} = \frac{1}{4}$.

Proposition 4.4. *Any rate pairs (R_1, R_2) in $\mathcal{R}_{LNC_1}$ and $\mathcal{R}_{LNC_2}$ are achievable by the scheme associated with $\mathcal{R}'_{BM}$.*

Proof. Since $\mathcal{R}_{LNC_2} \subseteq \mathcal{R}_{LNC_1}$ is justified, we only need to show $\mathcal{R}_{LNC_1} \subseteq \mathcal{R}'_{BM}$. The single user rate $1-\delta$ can be get by omitting the side information in the block Markov scheme. Furthermore, we check whether the off-axis corner point $(\frac{1-\delta}{1+\delta}, \frac{1-\delta}{1+\delta}\delta(1+\delta-\delta^2))$ of $\mathcal{R}_{LNC_1}$ verifies the condition in $\mathcal{R}_{BM}$ as

$$\frac{\frac{1-\delta}{1+\delta}}{1-\delta} + \frac{\frac{1-\delta}{1+\delta}\delta(1+\delta-\delta^2)}{1-\delta^2} = 1 - \frac{\delta^3}{(1+\delta)^2} \leq 1,$$
$$\frac{\frac{1-\delta}{1+\delta}}{\frac{1-\delta}{1-\delta+\delta^2}} + \frac{\frac{1-\delta}{1+\delta}\delta(1+\delta-\delta^2)}{1-\delta} = 1 - \frac{\delta(1-\delta)^2}{1+\delta} \leq 1.$$

Thus, we have $\mathcal{R}_{LNC_1} \subseteq \mathcal{R}'_{BM}$. ■

Remark 4.2. *It is readily seen that for any $0 < \delta < 1$, $\mathcal{R}'_{BM}$ by non-linear scheme is strictly larger than $\mathcal{R}_{LNC_2}$ and $\mathcal{R}_{LNC_1}$ obtained with the linear coding algorithms relying on bit-wise XOR.*

Let us assume $\delta = 0.5$ and illustrate the improvement in Fig. 15. As shown in the plot, thanks to the opportunistic coding gain, all three inner bounds outperform the TDMA region. As predicted, the adapted block Markov scheme provides a remarkable gain over the linear ones. The rate enhancement might result from the powerful non-linear coding tools and/or a properly designed structure, i.e., superimpose the side information over the private information. Though it has

been claimed previously, it is yet worth emphasizing that the vertex of $\mathcal{R}_{BM}^{'}$ is on the boundary of outer bound as one of the boundary in $\mathcal{R}_{BM}^{'}$ coincides with it.

In the block Markov scheme, the side information is useful when the elements in state set are with non-zero probability and erasure is possible. Given that, it seems feasible to extend the enhancement to certain specific finite state channel, e.g., zero is in the state set with non-zero probability. To that end, we provide an example as follows.

Example 4.2.1 (Finite Field Deterministic BC in D_1N_1 case [2, 54])**.** *We consider a symmetric spatially independent finite field deterministic BC, which can be described briefly by $Y_k = H_kX$. H_k associated with S_k is chosen uniformly from the set $\mathcal{H}$ over finite field that contains element 0, $k = 1, 2$. We set the probability $\Pr(H_k = 0) = p_0 = \frac{1}{|\mathcal{H}|}$. In addition, we assume $|\mathcal{Y}_k| = |\mathcal{Y}|$.*

We apply the same input choice as (95) *to the region by* (112a)-(112d) *and adapt the $\hat{Y}$'s choice in* (96) *by letting the condition $S_1 = 1, Q = 2$ be $S_1 \neq 0, Q = 2$, see Appendix 4.B. The corresponding region is given by*

$$\mathcal{R}_{FS} \triangleq \left\{ R_1, R_2 \in \mathbb{R}_+ \;\middle|\; \begin{array}{l} \frac{R_1}{1-p_0} + \frac{R_2}{1-p_0^2} \leq \log|\mathcal{Y}|, \\ \frac{R_1}{\frac{1-p_0}{1-p_0+p_0^2}} + \frac{R_2}{1-p_0} \leq \log|\mathcal{Y}|. \end{array} \right\} \tag{101}$$

where the sum rate by off-axis corner point is $\frac{(1-p_0)(2-p_0^2)}{2-p_0}\log|\mathcal{Y}| = (1-p_0)\log|\mathcal{Y}| + \frac{(1-p_0)^2p_0}{2-p_0}\log|\mathcal{Y}| \geq (1-p_0)\log|\mathcal{Y}|$.

This improvement tends to zero as $|\mathcal{H}|$ goes to infinity. Therefore, a state set with *bounded* cardinality is crucial for exploiting such opportunistic gain from the heterogeneous state feedback.

4.3 DISTRIBUTED LOSSY CODING SCHEME

Following the analysis of last section, the gain over the schemes in [43, 44] may result from the superimposed codewords structure, the non-linear coding techniques, and the carefully designed parameters which render a subtle balance over information. In order to locate precisely the origin of advantage, we prefer a scheme without superposition in the codewords. Since the DC scheme in Chapter 3 does not include superposition and is simple as compared to general schemes, we present in the sequel a uniform scheme for K-user D_lN_{K-l} case by adjusting the DC scheme. The adapted scheme also hinges on the aforementioned distributed lossy coding technique. In addition, we will specify it in multiple user EBC case.

Theorem 4.5 (Inner bound $\mathcal{R}_{DC}$)**.** *A rate tuple $(R_1, \ldots, R_K)$ is achievable for the K-user BC in D_lN_{K-l} case ($S_{\mathcal{L}}$ are known to the source) provided that*

$(R_1, \ldots, R_K) \in \mathcal{R}_{DC}$ *where* $\mathcal{R}_{DC}$ *is determined by the constraints below,*

$$R_k \leq \alpha_1 I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \mid S^{(1)}, Q^{(1)}), \tag{102a}$$

$$0 \leq \min_{\substack{i,j,k,\mathcal{J}:\\ i<j,k\in\mathcal{J}}} \Big\{ \alpha_j I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \mid S^{(j)}, Q^{(j)}) + \alpha_i I(Y_k^{(i)}; \hat{Y}_{i\to\mathcal{J}} \mid S^{(i)} Q^{(i)}) - \alpha_i I(\{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \mid S_{\mathcal{L}}^{(i)} Q^{(i)}) \Big\}, \tag{102b}$$

for some K-tuple $(\alpha_1, \ldots, \alpha_K) \in \mathbb{R}_+^K$ *with* $\sum_k \alpha_k = 1$, $\mathcal{L} = \{1, 2, \cdots, l\}$, *and some pmfs verify*

$$\prod_{j=1}^{K} p(y_1^{(j)}, \ldots, y_K^{(j)} \mid x^{(j)}, s^{(j)}) p(x^{(j)} \mid \{v_{\mathcal{J}}\}_{\mathcal{J}}, q^{(j)}) p(s^{(j)}) p(q^{(j)}) \times \prod_{j=1}^{K} \prod_{i=0}^{j-1} \prod_{\mathcal{J}} p(v_{i\to\mathcal{J}}) p(\hat{y}_{i\to\mathcal{J}} \mid \{v_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, s_{\mathcal{L}}^{(i)}, q^{(i)})$$

Sketch of proof. The proof follows an adaptation of the DC scheme described in Chapter 3 tailored for scenario with heterogeneous CSIT. In particular, $\hat{Y}_{i\to\mathcal{J}}$ is forced to build only on the state available at the source $S_{\mathcal{L}}^{(i)}$. This leads to several revisions as listed: using $p(\hat{y}_{i\to\mathcal{J}} \mid \{v_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, s_{\mathcal{L}}^{(i)}, q^{(i)})$ for the pmf, depending on $p(\{\hat{y}_{i\to\mathcal{J}}\}^n \mid \{v_{\mathcal{I}}^n\}_{\mathcal{I}\subset\mathcal{J}}, \{s_{\mathcal{L}}^{(i)}\}^n, \{q^{(i)}\}^n)$ in the codebook construction, relying on the joint typical set $\mathcal{T}_\delta^{n_i}(\hat{Y}_{i\to\mathcal{J}}, \{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}, S_{\mathcal{L}}, Q)$ in the encoding. ■

4.3.1 *EBC in D_1N_1 case*

In the two user EBC with D_1N_1 state, we have the following corollary.

Corollary 4.6. *The inner bound in Corollary 4.2 is also achievable with the tailored distributed lossy coding scheme.*

Proof. We let $H(V_1) = H(V_2) = H(V_{12}) = 1$, apply the choice of X and $\hat{Y}$ as shown below to the region in Theorem 4.5 when $K = 2$, $\mathcal{L} = \{1\}$.

$$X = \begin{cases} X_1 = V_1, & Q = 1, w.p.\ \alpha_1, \\ X_2 = V_2, & Q = 2, w.p.\ \alpha_2, \\ X_3 = V_{12}, & Q = 3, w.p.\ \alpha_3. \end{cases} \tag{103a}$$

$$\hat{Y} = \begin{cases} V_1, & S_1 = 0, Q = 1, \\ V_2, & S_1 = 1, Q = 2, \\ 0, & otherwise. \end{cases} \tag{103b}$$

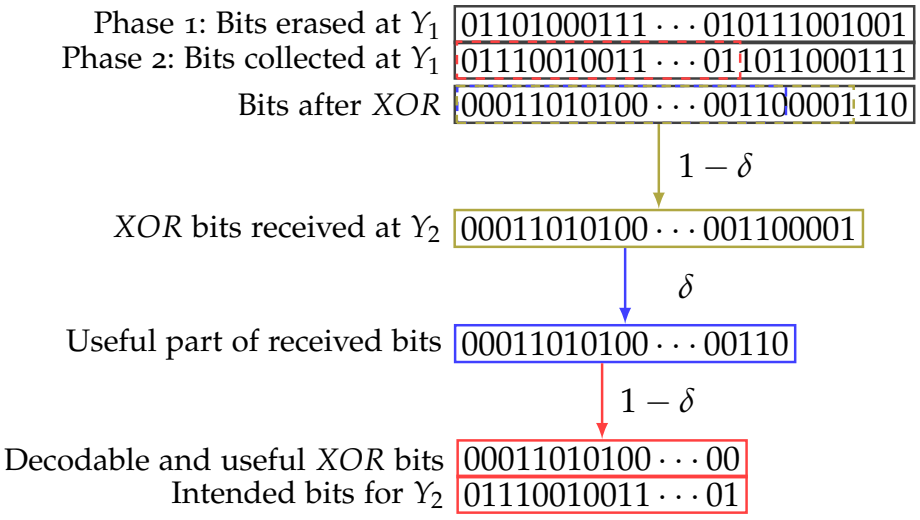

Figure 16: The side information exploitation in the linear scheme.

Then, we have

$$R_1 \leq \alpha_1, \tag{104a}$$
$$R_2 \leq \alpha_2(1-\delta_{12}), \tag{104b}$$
$$0 \leq \alpha_3(1-\delta_1) - \alpha_1\delta_1, \tag{104c}$$
$$0 \leq \alpha_3(1-\delta_2) - \alpha_1\delta_{12} - \alpha_2(\delta_2-\delta_{12}). \tag{104d}$$

To facilitate the proof, we ignore the standard procedure of running the FME on the constraints, instead, we verify that all the corner points of region in Corollary 4.2 are within the range determined by the above conditions because of the convexity of rate region. For $(1-\delta_1, 0)$, we set $\alpha_1 = 1-\delta_1$, $\alpha_2 = 0$, and $\alpha_3 = \delta_1$. To obtain $(0, 1-\delta_2)$, we let $\alpha_1 = 0$, $\alpha_2 = \frac{1-\delta_2}{1-\delta_{12}}$, and $\alpha_3 = \frac{\delta_2-\delta_{12}}{1-\delta_{12}}$. Regarding CP_1, we choose $\alpha_1 = R_{1,CP_1}$, $\alpha_2 = \frac{R_{2,CP_1}}{1-\delta_{12}}$, $\alpha_3 = \frac{\delta_1(\delta_2-\delta_{12})}{(1-\delta_{12})-(1-\delta_1)(1-\delta_2+\delta_{12})}$. ■

Remark 4.3. *In the symmetric spatially independent EBC, the maximum sum rate* $\frac{1-\delta}{2-\delta}(2-\delta^2)$ *in Corollary 4.2 is obtained at the corner point* $(\frac{1-\delta}{2-\delta}, \frac{(1-\delta)(1-\delta^2)}{2-\delta})$. *Surprisingly, such an asymmetric corner point can be achieved by assigning equal length to first two phases in the adapted DC scheme, that is,* $\alpha_1 = \alpha_2 = \frac{1-\delta}{2-\delta}$. *Moreover, if the same* α*'s are applied to the linear schemes in [44], strictly smaller sum rate can be achieved, i.e.,* $R_{sum} = \frac{1-\delta}{2-\delta}(2-\delta^2) - \frac{\delta(1-\delta)^2}{2-\delta}\max(\delta, 1-\delta)$.

We notice that non-linear scheme without superposition can still achieves the same superior performance over linear schemes. To understand where such advantage comes from, we first illustrate in Fig. 16 how receiver 2 exploits the side information in the linear schemes [44]. The first sequence stands for the bits intended for user 1 and erased at Y_1 in phase 1 while the second sequence represents the bits intended for user 2 and collected at Y_1 in phase 2. The third sequence, which is transmitted in phase 3 of length $\alpha_3 T$, is obtained by performing a XOR operation over first and second sequences. Then, at destination 2, $(1-\delta)\alpha_3 T$

bits are received.[4] Among the $(1-\delta)\alpha_3 T$ observed bits at Y_2, only δ proportion, who were erased at Y_2 in phase 2, is useful for destination 2. In order to decode these $\delta(1-\delta)\alpha_3 T$ useful bits, the decoder 2 conducts a *XOR* operation on the observed useful bits and the corresponding bits in first sequence. The *XOR* is feasible merely when the receiver 2 is aware of the *respective* bits erased at Y_1 in phase 1. The fact that $(1-\delta)$ proportion of the interested bits was witnessed at Y_2 implies that decoding the *XOR* bits is performed with the cost of losing $\delta^2(1-\delta)\alpha_3 T$ useful bits. Therefore, the decoder 2 possesses finally $\delta(1-\delta)^2\alpha_3 T$ bits that are received at, beneficial for, and decodable at the destination 2. In contrast, user 1 can avoid such loss when decoding the bits in first sequence given the side information bits. Since the bits in second sequence are all known to Y_1, *XOR* can be executed at destination 1 with no cost, which renders the useful bits, analogously defined for receiver 1, all decodable at Y_1.

A simple observation indicates that *XOR* the overheard bits together with wanted bits in the linear scheme is crucial to profit from the undesired observations. However, such advantage becomes costly to be taken of when part of the delayed state is missing at the source. The proposed non-linear scheme, on the contrary, does not suffer the $1-\delta$ loss at destination 2 in that it is not contingent upon the bit-wise operations, e.g., *XOR*. In particular, *all* the bits observed by Y_2 and simultaneously erased at Y_1 in phase 1 account for the (decodable) side information in the non-linear scheme whereas among the bits that are received by Y_2 and simultaneously removed at Y_1 in phase 1, only those *correspond to* the *XOR*ed useful bits are treated as (decodable) side information in the linear algorithm. Consequently, the *XOR*-based linear schemes do not seem to achieve the same rate region as the proposed non-linear schemes. When we return to the D_2 case, with both users' state at the source, the linear algorithm is able to design the *XOR*ed bits in such a way that all the received and useful bits at destination 2 are decodable, which yields a capacity-achieving rate region. Therefore, only when partial state is absent at the source, the shortcoming of a bit-wise *XOR* operation dedicated to interference mitigation is revealed.

Remark 4.4. *Provided that the distinct private and common messages in the distributed lossy coding scheme are sent sequentially likewise in [43, 44], it is readily seen that the performance gain comes from the non-linear coding techniques.*

Remark 4.5. *Owing to our scheme's close relation with the SW scheme [20] without binning, we obtain the respective region $\mathcal{R}'_{SW}$ by applying RVs choices* (103a), (103b) *for EBC to* (15a), (15b), *and* (15c) *and compare it with $\mathcal{R}'_{BM}$. For simplicity's sake, we concentrate on $\delta_1 > \delta_2$ case which ensures non-degenerate performance of our region as $\frac{1-\delta_2}{1-\delta_2+\delta_{12}} > 1-\delta_1$ is verified. The rate region is defined as follows (see Appendix 4.C for details),*

$$\mathcal{R}'_{SW} \triangleq \left\{ R_1, R_2 \in \mathbb{R}_+ \;\middle|\; \begin{array}{c} \frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1, \\ \frac{R_1}{\frac{(1-\delta_1)(1-\delta_2+\delta_1)}{1-\delta_2+\delta_{12}}} + \frac{R_2}{1-\delta_2} \leq 1. \end{array} \right\}. \tag{105}$$

[4] For an easy comprehension, the sequence when certain erasure or perfect recovery events are imposed is presented as a truncated version of the original sequence in the plot. However, the readers should bear in mind that the erasure or perfect recovery occurs *randomly* rather than consecutively.

Note that the first condition of $\mathcal{R}'_{SW}$ provides two on-axis corner points $(1-\delta_1, 0)$ and $(0, 1-\delta_{12})$ which coincide with those of $\mathcal{R}'_{BM}$, respectively. However, the second condition is a more demanding constraint than the one in $\mathcal{R}'_{BM}$, as is justified below.

$$\frac{(1-\delta_1)(1-\delta_2+\delta_1)}{1-\delta_2+\delta_{12}} = \frac{1-\delta_2+\delta_1(\delta_2-\delta_1)}{1-\delta_2+\delta_{12}} \leq \frac{1-\delta_2}{1-\delta_2+\delta_{12}} \tag{106}$$

The looseness of that condition results from the mismatched constraint (15c)*, which originates from the separation of source and channel coding. Consequently, the SW scheme without binning cannot achieve $\mathcal{R}'_{BM}$ by selecting the RVs in the same way as* (103a), (103b) *when $\delta_1 > \delta_2$. Nevertheless, one can verify that the SW scheme [20] could obtain a larger sum-rate than our adapted schemes (block Markov scheme and DC scheme) when $\delta_1 < \delta_2$. Finding a general scheme that outperforms SW scheme and our schemes under all kind of δ condition is an interesting direction for the future work.*

4.3.2 *Three user EBC*

In the following, we present two inner bounds for three user EBC in different state scenarios, D_2N_1 and $D_1N_1P_1$. These two cases can be regarded as adding a no-feedback user to two user EBC with D_2 and D_1P_1 states, respectively. Our goal here is to explore the potential of opportunistic coding gain in beyond two user channel and examine whether an extra isolated user (no feedback) can enlarge the capacity. Note that the D_1N_1 capacity is not thoroughly characterized, we omit the corresponding D_1N_2 case. For brevity, we investigate hereafter the symmetric spatially independent subclass of EBC.

D_2N_1 **case**

For the sake of simplicity, we consider only the partial symmetric rate for D_2N_1 case such that $R_1 = R_2 = R_{sym}$ in (R_1, R_2, R_3).

Corollary 4.7 (DC inner bound in D_2N_1 case)**.** *In a symmetric spatially independent EBC with D_2N_1 state, a rate triple (R_{sym}, R_{sym}, R_3) in region $\mathcal{R}_{D_2N_1}$ is achievable if*

$$\frac{2+\delta}{1-\delta^2}R_{sym} + \frac{1+\delta-\delta^2}{1-\delta^3}R_3 \leq 1, \tag{107a}$$

$$\frac{(1-\delta+\delta^3)(2-\delta+\delta^2)}{(1-\delta)(1-\delta+\delta^2)}R_{sym} + \frac{1}{1-\delta}R_3 \leq 1, \tag{107b}$$

$$\frac{(2+\delta)(1-\delta+\delta^2)}{(1-\delta)(1+\delta^2)}R_{sym} + \frac{1+\delta+2\delta^2+2\delta^3-\delta^4}{(1-\delta)(1+\delta+2\delta^2+\delta^3+\delta^4)}R_3 \leq 1. \tag{107c}$$

Sketch of proof. The rate region is justified by applying the following RVs to the one in Theorem 4.5 when $K = 3$ and $\mathcal{L} = \{1, 2\}$.

- X as a deterministic function of $(Q, V\text{'s})$, for $j = 1, 2, 3$,

$$X^{(j)} = V_{Q_1^{(j)} \to Q_2^{(j)}}. \tag{108}$$

- $\hat{Y}_{1\to\{1,2,3\}\setminus\{k\}}$ as a function of $(Q, V\text{'s}, S_1, S_2)$, for $k = 1, 2, 3$,

$$\hat{Y}_{1\to 12} = \begin{cases} V_1, & (S_1, S_2) = (0,1),\ Q = (0,1), \\ V_2, & (S_1, S_2) = (1,0),\ Q = (0,2), \\ 0, & \textit{otherwise}. \end{cases} \tag{109a}$$

$$\hat{Y}_{1\to 13} = \begin{cases} V_1, & (S_1, S_2) = (0,0),\ Q = (0,1), \\ V_3, & (S_1, S_2) = (1,0),\ Q = (0,3), \\ 0, & \textit{otherwise}. \end{cases} \tag{109b}$$

$$\hat{Y}_{1\to 23} = \begin{cases} V_2, & (S_1, S_2) = (0,0),\ Q = (0,2), \\ V_3, & (S_1, S_2) = (0,1),\ Q = (0,3), \\ 0, & \textit{otherwise}. \end{cases} \tag{109c}$$

  As both delayed state $(S_1,\ S_2)$ are known, the $\hat{Y}_{1\to 12}$ is designed in a similar manner to D_2 case, i.e., $\hat{Y}$ is for V_1 when $(S_1, S_2) = (0,1)$ and for V_2 when $(S_1, S_2) = (1,0)$. For $\hat{Y}_{1\to 13}$, the side information is regulated based on the same principle of the distributed lossy coding scheme in D_1N_1 case. Whenever neither of the terminal 1 and 2 acquires V_1, we set $\hat{Y}_{1\to 13}$ equal to V_1 wishing that it is obtained by user 3 and thus no extra penalty will be given (to destination 3) when decoding the corresponding portion of $\hat{Y}_{1\to 13}$. When V_3 is seen by user 1 and removed at user 2, $\hat{Y}_{1\to 13}$ is associated with V_3 so that the auxiliary information could, by chance, assist user 3. By replacing the role of user 1 with user 2, the design of $\hat{Y}_{1\to 23}$ follows.

- $\hat{Y}_{i\to\{1,2,3\}}$ determined by $(Q, V\text{'s}, S_1, S_2)$, for $i = 1, 2$,

$$\hat{Y}_{1\to 123} = \begin{cases} V_3, & (S_1, S_2) = (1,1),\ Q = (0,3), \\ 0, & \textit{otherwise} \end{cases} \tag{110a}$$

$$\hat{Y}_{2\to 123} = \begin{cases} V_{12}, & (S_1, S_2) = (0,0),\ Q = (1,\{12\}), \\ V_{13}, & (S_1, S_2) = (0,1),\ Q = (1,\{13\}), \\ V_{23}, & (S_1, S_2) = (1,0),\ Q = (1,\{23\}), \\ 0, & \textit{otherwise} \end{cases} \tag{110b}$$

  In general, when conceiving the side information, the lower order auxiliary messages should be considered in priority since multicasting to a larger group of users may lead to a rate loss caused by the weakest user. As such, we let V_1 (resp. V_2) correspond to $\hat{Y}_{1\to 12}$ rather than $\hat{Y}_{1\to 123}$ if $(S_1, S_2) = (0,1)$ (resp. $(S_1, S_2) = (1,0)$), $Q = (0,1)$ (resp. $Q = (0,2)$), as otherwise the rate is limited by more links. $\hat{Y}_{1\to 123} = V_3$ is selected when $(S_1, S_2) = (1,1),\ Q = (0,3)$ as it can be retrieved at destination 1, 2 with no supplemental cost. In this case, opportunistic coding gain appears when $S_3 = 0$.

As long as V_{12} is missed by both desired terminals, we assign it to $\hat{Y}_{2\to 123}$ in the hope that receiver 3 learns V_{12} which renders this segment of the information in $\hat{Y}_{2\to 123}$ costless at Y_3. If the receiver 2 observes V_{13} while at least one of user 1, 3 is not aware of the signal, we should associate it with $\hat{Y}_{2\to 123}$. However, we assume, for brevity, that this occurs merely when $(S_1, S_2) = (0, 1)$ (rather than $S_2 = 1$) due to the lack of S_3. Analogously, V_{23} is associated with $\hat{Y}_{2\to 123}$ when $(S_1, S_2) = (1, 0),\ Q = (1, \{23\})$.

The rest of the calculation is deferred to Appendix 4.D. ■

Before analyzing the sum rate performance, we define the maximum single user rates related to each constraint of $\mathcal{R}_{D_2N_1}$ as $R_{A_{sym}} \triangleq \frac{1-\delta^2}{2+\delta}$, $R_{A_3} \triangleq \frac{1-\delta^3}{1+\delta-\delta^2}$ determined by (107a), $R_{B_{sym}} \triangleq \frac{(1-\delta)(1-\delta+\delta^2)}{(1-\delta+\delta^3)(2-\delta+\delta^2)}$, $R_{B_3} \triangleq 1-\delta$ restricted by (107b), and $R_{C_{sym}} \triangleq \frac{(1-\delta)(1+\delta^2)}{(2+\delta)(1-\delta+\delta^2)}$, $R_{C_3} \triangleq \frac{(1-\delta)(1+\delta+2\delta^2+\delta^3+\delta^4)}{1+\delta+2\delta^2+2\delta^3-\delta^4}$ ruled by (107c). Moreover, we define $\delta' \approx 0.26$, $\delta'' \approx 0.66$, and $\delta''' \approx 0.80$ as the real positive root of $3\delta^2+3\delta-1=0$, $\delta^4-\delta^3+\delta^2+\delta-1=0$, and $\delta^4+2\delta-2=0$, respectively. Next, we discuss the maximum achievable sum rate for different δ values. We remind the reader that $0<\delta'<\frac{1}{2}<\delta''<\delta'''<1$.

- $0 \le \delta \le \frac{1}{2}$: $R_{A_{sym}} \le R_{C_{sym}} \le R_{B_{sym}}$, $R_{C_3} \le R_{B_3} \le R_{A_3}$.

  The rate region is determined by (107a) and (107c) which leads to an off-axis corner point $(R_{sym}, R_3) = (\frac{1-\delta^2}{3+\delta-2\delta^2+\delta^3}, \frac{(1-\delta)^2(1+\delta+\delta^2)}{3+\delta-2\delta^2+\delta^3})$. The sum rate equals to $\frac{3-\delta-2\delta^2-\delta^3+\delta^4}{3+\delta-2\delta^2+\delta^3} = 1-\delta(1-\frac{4(\frac{1}{2}-\delta)(\frac{1}{2}+\delta)+\delta+\delta^3}{3+\delta-2\delta^2+\delta^3}) \ge 1-\delta$. When $\delta > \delta'$, the sum-rate of this corner point outperforms the D_2 rate $\frac{2(1-\delta^2)}{(2+\delta)}$. Unfortunately, this does not hold for $\delta \le \delta'$ since the opportunistic coding gain vanishes faster than the stable improvement by coding between receivers who feedback states as δ decreases.

- $\delta'' \le \delta \le \delta'''$: $R_{A_{sym}} \le R_{B_{sym}} \le R_{C_{sym}}$, $R_{B_3} \le R_{C_3} \le R_{A_3}$.

  In this case, the activated constraints are (107a) and (107b). The corresponding corner point is $(R_{sym}, R_3) = (\frac{2(1-\delta^2)(1-\delta+\delta^2)}{6-2\delta-2\delta^2+4\delta^3-\delta^4-\delta^5+\delta^6}, \frac{(1-\delta)(1+\delta+\delta^2)(2-\delta-\delta^4)}{6-2\delta-2\delta^2+4\delta^3-\delta^4-\delta^5+\delta^6})$, which gives the sum rate $\frac{2(1-\delta^2)}{2+\delta} + \frac{\delta(1-\delta)(-2+8\delta-6\delta^3+\delta^4-3\delta^5-3\delta^6)}{(6-2\delta-2\delta^2+4\delta^3-\delta^4-\delta^5+\delta^6)(2+\delta)} \ge \frac{2(1-\delta^2)}{2+\delta}$ when $\delta' \le \delta \le \delta'''$.

- $\frac{1}{2} \le \delta \le \delta''$: $R_{A_{sym}} \le R_{C_{sym}} \le R_{B_{sym}}$, $R_{B_3} \le R_{C_3} \le R_{A_3}$.

  Upon inspection, we see that the off-axis corner point either by (107a), (107c) or by (107a), (107b) is achievable. Note that both corner points allows a better performance when $\delta' \le \delta \le \delta'''$, in terms of sum rate, as compared to D_2. Hence, for $\frac{1}{2} \le \delta \le \delta''$, a superior total rate is attainable.

- $\delta''' \le \delta \le 1$: $R_{B_{sym}} \le R_{A_{sym}} \le R_{C_{sym}}$, $R_{B_3} \le R_{C_3} \le R_{A_3}$.

  The system falls into the degenerate case as the sum rate is dominated by D_2 rate, i.e., $R_{sum} \le \max\{R_{B_{sym}}, R_{B_3}\} \le R_{A_{sym}} \le \frac{2(1-\delta^2)}{2+\delta}$. If δ is large enough, quite some signals desired at user 1, 2 are removed and the source relies strongly on receiver 3's observation. Nonetheless, destination 3 itself is also

experiencing intense erasure. As a consequence, involving terminal 3 in the information interchange may be too costly.

Under certain circumstances ($\delta' \leq \delta \leq \delta'''$), inserting a user who does not feedback its state could still enlarge the sum rate of D_2. However, when erasure happens too frequently or too rarely, the improvement shrinks. Since for an extreme δ (really low or high) either user 3 does not have adequate observations or the others do not need its observation, it is preferable to execute the information swapping within the group of the users that provide the delayed state information.

$D_1N_1P_1$ case

Corollary 4.8 (Inner bound in $D_1N_1P_1$ case)**.** *For a symmetric spatially independent EBC with $D_1N_1P_1$ state, a rate triple (R_1, R_2, R_3) in the region $\mathcal{R}_{D_1N_1P_1}$ is achievable provided that $R_3 \leq 1-\delta$ and*

$$\frac{R_1}{1-\delta} + \frac{R_2}{1-\delta^2} \leq \delta, \tag{111a}$$

$$\frac{R_1}{\frac{1-\delta}{1-\delta+\delta^2}} + \frac{R_2}{1-\delta} \leq \delta. \tag{111b}$$

Sketch of proof. Though the above region is not a direct application of Theorem 4.5, we put it in a corollary for simplicity. We describe the achievable scheme as: the encoder carries out a strategy that switches between the user 3 with P state and the rest. Specifically, the source sends out the encoded signal that assigns to the message for user 3 when $S_3 = 1$, which ensures $R_3 \leq 1-\delta$, and allocates the slots to user 1 and 2 if $S_3 = 0$. When $S_3 = 0$, user 3 is deactivated and the source and the rest two users formulate a two user BC with D_1N_1 state. During the interval, the scheme attains Corollary 4.2 is conducted. Since it only works when $S_3 = 0$, the proportion δ is considered in the constraints. ■

Note that by either adding a no-feedback user to D_1P_1 case or inserting a user providing delayed CSI into N_1P_1 case, we arrive at a three-user BC in the $D_1N_1P_1$ condition. In the capacity-achieving strategy for N_1P_1, the secondary user (no feedback) is allowed to take the channel only when the primary user is experiencing an erasure. During such period, involving a node that provides delayed state to cooperate with secondary user has been shown to be preferable. Thus, $\mathcal{R}_{D_1N_1P_1} \supset \mathcal{C}_{N_1P_1}$. Along with $\mathcal{C}_{N_1P_1} = \mathcal{C}_{D_1P_1}$ demonstrated in [43], we have $\mathcal{R}_{D_1N_1P_1} \supset \mathcal{C}_{D_1P_1}$ which validates the usefulness of the user without feedback added into D_1P_1 scenario.

4.4 SUMMARY AND CLOSING REMARKS

In this chapter, the multiple user BC with *heterogeneous* state feedback is studied. We derived two inner bounds based on the strategies adapted from the homogeneous state feedback schemes proposed in Chapter 2 and 3, namely, block Markov

scheme and Distributed Compression method. In a two user EBC with delayed CSI of a single user at the transmitter's side, a rate region that can be obtained by both tailored schemes was evaluated. For non-degenerate cases, we demonstrated that employing non-linear coding techniques enlarges the inner bounds by linear methods [43, 44], which differs from D_2 case. Surprisingly, one constraint of our inner bound even collapses to the outer bound.

Note that the performance gain does not come from the scheme's superimposed structure, instead, it relies on constructing the side information in a more efficient way. On the contrary to the linear scheme where overheard information at distinct users are associated bit-wisely via XOR, the proposed method generates the codewords through random coding and assigns the side information for each terminal individually such that no correspondence on the bits across the phases is needed.

Additionally, with Distributed Compression method, notable enhancements were revealed to be possible when including a receiver, who does not feedback its state, into two user EBC subjected to D_2 and D_1P_1 state restrictions. While the analysis on the proposed inner bounds suggests improvement over single user rate is made possible with heterogeneous state for some specific pmfs in EBC, whether there exists another coding scheme that outperforms the no feedback rate for all $\delta_{\mathcal{J}}$'s is still an open question, which remains as future work.

4.A PROOF OF COROLLARY 4.2

Before we start applying the RVs indicated for the corollary, some basic manipulations on the rate region in Theorem 4.1 can be performed. We rewrite the rate region as,

$$R_1 \leq I(V_1; Y_1\hat{Y}|V_{12}S_1S_2Q) + \{I(V_{12}\hat{Y}; Y_1|S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q)\}^-,$$
$$R_2 \leq I(V_2; Y_2\hat{Y}|V_{12}S_1S_2Q) + \{I(V_{12}\hat{Y}; Y_2|S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q)\}^-.$$

where $\{f\}^-$ is equivalent to $\min\{0, f\}$. Then, we chose certain probability functions from all the possible set of pmf's such that the part inside $\{\cdot\}^{-1}$ is non-negative, i.e., $0 \leq I(V_{12}\hat{Y}; Y_k|S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q)$. Therefore, we have a novel rate region described by two separate groups of restrictions: two conditions on the individual rates and two constraints on the side information delivery, which reads as,

$$R_1 \leq I(V_1; Y_1\hat{Y}|V_{12}S_1S_2Q), \tag{112a}$$
$$R_2 \leq I(V_2; Y_2\hat{Y}|V_{12}S_1S_2Q), \tag{112b}$$
$$0 \leq I(V_{12}\hat{Y}; Y_1|S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q), \tag{112c}$$
$$0 \leq I(V_{12}\hat{Y}; Y_2|S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q). \tag{112d}$$

With the aforementioned RVs selections, each distinct mutual information is provided by

$$\begin{aligned}
&I(V_1; Y_1\hat{Y}|V_{12}S_1S_2Q) \\
&\quad = \alpha_1[(1-\delta_1)H(Y_1 = X_1|V_{12}, S_1 = 1) + \delta_1 H(\hat{Y} = X_1|V_{12}, S_1 = 0)] \\
&\quad = \alpha_1 H(V_1), && \text{(113a)} \\
&I(V_2; Y_2\hat{Y}|V_{12}S_1S_2Q) \\
&\quad = \alpha_2[(1-\delta_2)H(X_2|V_{12}, S_2 = 1) + (\delta_2 - \delta_{12})H(X_2|V_{12}, (S_1, S_2) = (1, 0))] \\
&\quad = \alpha_2(1-\delta_{12})H(V_2), && \text{(113b)} \\
&I(V_{12}; Y_k|S_1S_2Q) \\
&\quad = (1-\delta_k)I(V_{12}; X|S_k = 1, S_{\bar{k}}, Q) \\
&\quad = \alpha_1(1-\delta_k)[H(X_1) - H(V_1)] + \alpha_2(1-\delta_k)[H(X_2) - H(V_2)], && \text{(113c)} \\
&I(\hat{Y}; Y_1|V_{12}S_1S_2Q) \\
&\quad = \alpha_2(1-\delta_1)I(X; X|V_{12}, S_1 = 1, S_2, Q = 2) \\
&\quad = \alpha_2(1-\delta_1)H(V_2), && \text{(113d)} \\
&I(\hat{Y}; Y_2|V_{12}S_1S_2Q) \\
&\quad = \alpha_1(\delta_1 - \delta_{12})H(X_1|V_{12}, (S_1, S_2) = (0, 1)) \\
&\qquad + \alpha_2(1-\delta_1-\delta_2+\delta_{12})H(X_2|V_{12}, (S_1, S_2) = (1, 1)) \\
&\quad = \alpha_1(\delta_1 - \delta_{12})H(V_1) + \alpha_2(1-\delta_1-\delta_2+\delta_{12})H(V_2), && \text{(113e)} \\
&I(\hat{Y}; V_1V_2|V_{12}S_1Q) \\
&\quad = \alpha_1\delta_1 H(X_1|V_{12}, S_1 = 0) + \alpha_2(1-\delta_1)H(X_2|V_{12}, S_1 = 1) \\
&\quad = \alpha_1\delta_1 H(V_1) + \alpha_2(1-\delta_1)H(V_2). && \text{(113f)}
\end{aligned}$$

Hence, the corresponding rate region is specified with a few algebraic operations, i.e.,

$$R_1 \leq \alpha_1 H(V_1), \tag{114a}$$
$$R_2 \leq \alpha_2(1-\delta_{12})H(V_2), \tag{114b}$$
$$0 \leq (1-\delta_1)(\alpha_1 H(X_1)+\alpha_2 H(X_2)) - \alpha_1 H(V_1) - \alpha_2(1-\delta_1)H(V_2), \tag{114c}$$
$$0 \leq (1-\delta_2)(\alpha_1 H(X_1)+\alpha_2 H(X_2)) - \alpha_1(1-\delta_2+\delta_{12})H(V_1) - \alpha_2(1-\delta_{12})H(V_2). \tag{114d}$$

We notice that the entropy $(H(X_1),\ H(X_2))$ and $(H(V_1),\ H(V_2))$ are separable in these conditions and thus there is no loss by imposing $H(X_1) = H(X_2) = H(X) = 1$. This can be justified by the fact that no extra restrictions on $H(V_1),\ H(V_2),\ \alpha_1,\ \alpha_2$ will be provoked when V_{12} is chosen uniformly from $\{0,1\}$. Indeed, (113c) is the only mutual information that concern $H(X)$ and it can be maximized with this choice, that is, $H(V_{12}) = H(X_1) = H(X_2) = 1$. We obtain the following rate region by including $\alpha_1 + \alpha_2 = 1$.

$$R_1 \leq \alpha_1 H(V_1) \tag{115a}$$
$$R_2 \leq \alpha_2(1-\delta_{12})H(V_2) \tag{115b}$$
$$0 \leq 1-\delta_1 - \alpha_1 H(V_1) - \alpha_2(1-\delta_1)H(V_2) \tag{115c}$$
$$0 \leq 1-\delta_2 - \alpha_1(1-\delta_2+\delta_{12})H(V_1) - \alpha_2(1-\delta_{12})H(V_2) \tag{115d}$$

We observe that α_k and $H(V_k)$ always appear at the same time in the constraints. Thus, we define $r_1 = \alpha_1 H(V_1)$ and $r_2 = \alpha_2(1-\delta_{12})H(V_2)$, which reduces the round of FME needed, and rewrite the rate constraints as $R_1 \leq r_1,\ R_2 \leq r_2$,

$$0 \leq (1-\delta_1) - r_1 - \frac{(1-\delta_1)r_2}{1-\delta_{12}}, \tag{116a}$$
$$0 \leq (1-\delta_2) - (1-\delta_2+\delta_{12})r_1 - r_2, \tag{116b}$$

where verifying the feasibility of such $(r_1,\ r_2)$ is relegated to Section 4.2 in which we discuss the achievability of corner point. After running FME to the rate constraints, we obtain the region in Corollary 4.2.

4.B PROOF OF EXPRESSION (101)

We apply the RVs as described to have,

$$I(V_1; Y_1\hat{Y}|V_{12}SQ) = \alpha_1(1-p_0)H(V_1) + \alpha_1 p_0 H(V_1) = \alpha_1 H(V_1), \tag{117a}$$
$$I(V_2; Y_2\hat{Y}|V_{12}SQ) = \alpha_2(1-p_0^2)H(V_2), \tag{117b}$$
$$I(V_{12}; Y_k|SQ) = (1-p_0)[\alpha_1(H(X_1)-H(V_1)) + \alpha_2(H(X_2)-H(V_2))], \tag{117c}$$
$$I(\hat{Y}; Y_1|V_{12}SQ) = \alpha_2(1-p_0)H(V_2), \tag{117d}$$
$$I(\hat{Y}; Y_2|V_{12}SQ) = \alpha_1 p_0(1-p_0)H(V_1) + \alpha_2(1-p_0)^2 H(V_2), \tag{117e}$$
$$I(\hat{Y}; V_1V_2|V_{12}S_1Q) = \alpha_1 p_0 H(V_1) + \alpha_2(1-p_0)H(V_2). \tag{117f}$$

We let $H(X_1) = H(X_2) = \log|\mathcal{Y}|$ and run FME on the $\alpha_1 H(V_1)$ and $\alpha_2 H(V_2)$ with a similar reasoning as in Appendix 4.A. Finally, we get the region $\mathcal{R}_{FS}$.

4.C PROOF OF EXPRESSION (105)

In the following, we apply (103a), (103b) to (15a), (15b), and (15c), which starts with a similar calculation of the mutual information as expressed below.

$$I(V_1; Y_1\hat{Y}|V_{12}SQ) = \alpha_1 H(V_1), \tag{118a}$$
$$I(V_2; Y_2\hat{Y}|V_{12}SQ) = \alpha_2(1-\delta_{12})H(V_2), \tag{118b}$$
$$I(V_{12}; Y_k\hat{Y}|SQ) = \alpha_3(1-\delta_k)H(V_{12}), \tag{118c}$$
$$I(\hat{Y}; Y_1|SQ) = \alpha_2(1-\delta_1)H(V_2), \tag{118d}$$
$$I(\hat{Y}; Y_2|SQ) = \alpha_1(\delta_1-\delta_{12})H(V_1) + \alpha_2(1-\delta_1-\delta_2+\delta_{12})H(V_2), \tag{118e}$$
$$I(\hat{Y}; V_1V_2V_{12}|S_1Q) = \alpha_1\delta_1 H(V_1) + \alpha_2(1-\delta_1)H(V_2). \tag{118f}$$

Then, we substitute the above expressions in (15a), (15b), and (15c), which provides,

$$R_1 \le \alpha_1(1-\delta_1)H(V_1) + \alpha_3(1-\delta_1)H(V_{12}), \tag{119a}$$
$$R_2 \le \alpha_2(1-\delta_2)H(V_2) + \alpha_3(1-\delta_2)H(V_{12}) - \alpha_1\delta_{12}H(V_1), \tag{119b}$$
$$R_1+R_2 \le \alpha_1(1-\delta_1)H(V_1) + \alpha_2(1-\delta_{12})H(V_2) + \alpha_3(1-\delta_1)H(V_{12}), \tag{119c}$$
$$R_1+R_2 \le \alpha_1(1-\delta_{12})H(V_1) + \alpha_2(1-\delta_2)H(V_2) + \alpha_3(1-\delta_1)H(V_{12}), \tag{119d}$$
$$R_1+R_2 \le \alpha_1(1-\delta_1)H(V_1) + \alpha_2(1-\delta_{12})H(V_2) + \alpha_3(1-\delta_2)H(V_{12}), \tag{119e}$$
$$R_1+R_2 \le \alpha_1(1-\delta_{12})H(V_1) + \alpha_2(1-\delta_2)H(V_2) + \alpha_3(1-\delta_2)H(V_{12}). \tag{119f}$$

Since we assume $\delta_1 > \delta_2$, the last two constraints are released. From observation, we know there is no loss by letting $H(V_2) = H(V_{12}) = 1$. As α_1 and $H(V_1)$ always appear together, the optimal choice $H(V_1) = 1$ can be justified: for a fixed value of $\alpha_1 H(V_1)$, α_1 increases when $H(V_1)$ decreases which would lead to smaller $(\alpha_2,\ \alpha_3)$ and thus harms achievable rate. The following region is achieved after running FME on the α_1 and α_2, along with $\alpha_1+\alpha_2+\alpha_3 = 1$.

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \le 1, \tag{120a}$$
$$\frac{R_1}{\frac{1-\delta_2}{1-\delta_2+\delta_{12}}} + \frac{R_2}{1-\delta_2} \le 1, \tag{120b}$$
$$\frac{R_1}{\frac{(1-\delta_1)(1-\delta_2+\delta_1)}{1-\delta_2+\delta_{12}}} + \frac{R_2}{1-\delta_2} \le 1, \tag{120c}$$
$$R_1+R_2 \le 1 - \frac{\delta_1\delta_2-\delta_{12}^2}{\delta_1+\delta_2-2\delta_{12}}. \tag{120d}$$

where (120b) is redundant as $\frac{(1-\delta_1)(1-\delta_2+\delta_1)}{1-\delta_2+\delta_{12}} < \frac{1-\delta_2}{1-\delta_2+\delta_{12}}$ given $\delta_1 > \delta_2$. Finally, the proof completes by verifying that (120d) is loose to the combination of (120a) and (120b), which is equivalent to show that the maximum sum-rate by (120a), (120b),

namely R_{sum,CP_1}, is always smaller than that of (120d).

$$1-\frac{\delta_1\delta_2-\delta_{12}^2}{\delta_1+\delta_2-2\delta_{12}} = 1-\delta_{12}\frac{\frac{\delta_1\delta_2}{\delta_{12}}-\delta_{12}}{\delta_1+\delta_2-2\delta_{12}} \tag{121a}$$

$$\geq 1-\delta_{12}\frac{\frac{\delta_1\delta_2}{\delta_{12}}-\delta_{12}-\delta_1(\delta_2-\delta_{12})}{\delta_1+\delta_2-2\delta_{12}-\delta_1(\delta_2-\delta_{12})} \tag{121b}$$

$$= 1+\frac{\delta_{12}^2-\delta_1\delta_2+\delta_1\delta_{12}(\delta_2-\delta_{12})}{\delta_1+\delta_2-2\delta_{12}-\delta_1(\delta_2-\delta_{12})} = R_{sum,CP_1} \tag{121c}$$

where (121b) follows that a function in the form $f(x)=\frac{g-x}{h-x}$ when $0\leq x\leq h$ and $h\leq g$ decreases monotonically as x increases.

4.D PROOF OF COROLLARY 4.7

Due to the symmetry on the channel and on the rate, barely partial constraints are evaluated in this part. First of all, we calculate the concerned mutual information as in previous sections.

$$\alpha_1 I(V_1;Y_1\hat{Y}_{1\to12}\hat{Y}_{1\to13}\hat{Y}_{1\to123}|S^{(1)}Q^{(1)}) = \alpha_{\{1\}}H(V_1) \tag{122a}$$

$$\alpha_1 I(V_3;Y_3\hat{Y}_{1\to13}\hat{Y}_{1\to23}\hat{Y}_{1\to123}|S^{(1)}Q^{(1)}) = \alpha_{\{3\}}(1-\delta^3)H(V_3) \tag{122b}$$

$$\alpha_2 I(V_{1\to12};Y_1\hat{Y}_{2\to123}|S^{(2)}Q^{(2)}) = \alpha_{\{12\}}(1-\delta+\delta^2)H(V_{1\to12}) \tag{122c}$$

$$\alpha_2 I(V_{1\to13};Y_1\hat{Y}_{2\to123}|S^{(2)}Q^{(2)}) = \alpha_{\{13\}}(1-\delta^2)H(V_{1\to13}) \tag{122d}$$

$$\alpha_2 I(V_{1\to13};Y_3\hat{Y}_{2\to123}|S^{(2)}Q^{(2)}) = \alpha_{\{13\}}(1-\delta)(1+\delta^2)H(V_{1\to13}) \tag{122e}$$

$$\alpha_3 I(V_{i\to123};Y_k|S^{(3)}Q^{(3)}) = \alpha_{\{123\},i}(1-\delta)H(V_{i\to123}) \tag{122f}$$

$$\alpha_1\big(I(V_1V_2;\hat{Y}_{1\to12}|S_1^{(1)}S_2^{(1)}Q^{(1)})-I(Y_1;\hat{Y}_{1\to12}|S^{(1)}Q^{(1)})\big) = \alpha_{\{1\}}\delta(1-\delta)H(V_1) \tag{122g}$$

$$\alpha_1\big(I(V_1V_3;\hat{Y}_{1\to13}|S_1^{(1)}S_2^{(1)}Q^{(1)}) - I(Y_1;\hat{Y}_{1\to13}|S^{(1)}Q^{(1)})\big) = \alpha_{\{1\}}\delta^2H(V_1) \tag{122h}$$

$$\begin{aligned}&\alpha_1\big(I(V_1V_3;\hat{Y}_{1\to13}|S_1^{(1)}S_2^{(1)}Q^{(1)}) - I(Y_3;\hat{Y}_{1\to13}|S^{(1)}Q^{(1)})\big)\\ &\quad= \alpha_{\{1\}}\delta^3H(V_1)+\alpha_{\{3\}}\delta^2(1-\delta)H(V_3)\end{aligned} \tag{122i}$$

$$\begin{aligned}&\alpha_1\big(I(V_1V_2V_3;\hat{Y}_{1\to123}|S_1^{(1)}S_2^{(1)}Q^{(1)}) - I(Y_3;\hat{Y}_{1\to123}|S^{(1)}Q^{(1)})\big)\\ &\quad= \alpha_{\{3\}}\delta(1-\delta)^2H(V_3)\end{aligned} \tag{122j}$$

$$\begin{aligned}&\alpha_2\big(I(V_{1\to12}V_{1\to13}V_{1\to23};\hat{Y}_{2\to123}|S_1^{(2)}S_2^{(2)}Q^{(2)}) - I(Y_1;\hat{Y}_{2\to123}|S^{(2)}Q^{(2)})\big)\\ &\quad= \alpha_{\{12\}}\delta^2H(V_{1\to12})+\alpha_{\{13\}}\delta(1-\delta)H(V_{1\to13})\end{aligned} \tag{122k}$$

$$\begin{aligned}&\alpha_2\big(I(V_{1\to12}V_{1\to13}V_{1\to23};\hat{Y}_{2\to123}|S_1^{(2)}S_2^{(2)}Q^{(2)}) - I(Y_3;\hat{Y}_{2\to123}|S^{(2)}Q^{(2)})\big)\\ &\quad= \alpha_{\{12\}}\delta^3H(V_{1\to12})+2\alpha_{\{13\}}\delta^2(1-\delta)H(V_{1\to13})\end{aligned} \tag{122l}$$

where $i=1,2$ and $k=1,2,3$. A simple observation suggest that there is no loss if we assume $H(V_{1\to123})=H(V_{2\to123})=1$. Moreover, we notice that it is the product of $\alpha_{\{1\}}$ and $H(V_1)$ which matters, rather than the value of $\alpha_{\{1\}}$ or $H(V_1)$. For a fixed set of $H(V)$ and a fixed value of $0\leq\alpha_{\{1\}}H(V_1)\leq1$, if $H(V_1)$ is bounded away from 1, $\alpha_{\{1\}}$ will not be set to its minimum which would reduce the remaining α's. In that case, maximum rate cannot be achieved. Therefore,

letting $H(V_1) = 1$ helps when the other parameters are fixed. Similarly, we make $H(V_3) = H(V_{1\to 12}) = H(V_{1\to 13})$ equal to one. Then, the constraints are written as (123).

$$R_{sym} \le \alpha_{\{1\}}, \tag{123a}$$
$$R_3 \le \alpha_{\{3\}}(1-\delta^3), \tag{123b}$$
$$0 \le \alpha_{\{12\}}(1-\delta+\delta^2) - \alpha_{\{1\}}\delta(1-\delta), \tag{123c}$$
$$0 \le \alpha_{\{13\}}(1-\delta^2) - \alpha_{\{1\}}\delta^2, \tag{123d}$$
$$0 \le \alpha_{\{13\}}(1-\delta)(1+\delta^2) - \alpha_{\{1\}}\delta^3 - \alpha_{\{3\}}\delta^2(1-\delta), \tag{123e}$$
$$0 \le \alpha_{\{123\},1}(1-\delta) - \alpha_{\{3\}}\delta(1-\delta)^2, \tag{123f}$$
$$0 \le \alpha_{\{123\},2}(1-\delta) - \alpha_{\{12\}}\delta^2 - \alpha_{\{13\}}\delta(1-\delta), \tag{123g}$$
$$0 \le \alpha_{\{123\},2}(1-\delta) - \alpha_{\{12\}}\delta^3 - 2\alpha_{\{13\}}\delta^2(1-\delta). \tag{123h}$$

In order to obtain the region in Corollary 4.7, we run the FME to the above region along with $2\alpha_{\{1\}} + \alpha_{\{3\}} + \alpha_{\{12\}} + 2\alpha_{\{13\}} + \alpha_{\{123\},1} + \alpha_{\{123\},2} = 1$.

5

Broadcast Relay Channel with State Feedback at the Relay

5.1 OVERVIEW

In this chapter, we investigate the potential performance gain from the use of relay in the presence of CSI feedback from the destinations to the relay. The investigated broadcasting problem can be modeled in a BRC with heterogeneous state feedback at the relay. The main credits of this part are summarized as follows.

- We first propose a decode-compress-and-forward (DCF) relaying scheme that exploits the causal CSI feedback and then derive the corresponding achievable rate region. The proposed scheme is first presented for the general two-user memoryless BRC, and then specified for the fast fading Gaussian noise channel. Our scheme is based on the aforementioned RIA with the help of a relay.

- In particular, we consider two variants of the DCF scheme according to how the source messages are sent and relayed, that is, simultaneous emitted DCF (SE-DCF) and time division DCF (TD-DCF). With both variants, the optimal sum-DoF in the Gaussian BRC is achievable.

- Besides, a major contribution of this chapter lies in the characterization and evaluation of the rate region achieved by the proposed schemes in finite SNR. With numerical simulation focusing on the Gaussian case, we compare our schemes with the conventional DF and CF schemes directly adapted to the fading BRC. Our results reveal that the proposed scheme provides a non-negligible gain over conventional schemes in terms of sum-rate performance, especially in the medium-to-high SNR regime.

It is worth re-emphasizing that the heterogeneity in this chapter includes two aspects:

1. The first heterogeneity lies in network layout that can be easily noticed in both the physical model and its abstract general system model as described by (10). Given that delayed CSI is proven to be useful in the single-tier BC, examining the effectiveness of delayed CSI in such two-tier network (source-relay pair $\leftrightarrow$ macro-cell BS - small-cell BS pair) is quite significant from practical perspective.

2. The second one is the heterogeneity of the CSI availability (no CSI at the source and delayed CSI at the relay) to each node via feedback. This is also worth exploring since it corresponds the pragmatic scenario where only limited resource can be allocated to the feedback links. The feedback to the relay rather than the source can be justified by the economics of geographical layout of the two-tier network as mentioned in Chapter 1.

The rest of the chapter is organized as: Section 5.2 provides the achievable regions in general BRC while Section 5.3 characterizes the regions by introducing the variants of the proposed scheme in the Gaussian case. Numerical results are shown in Section 5.4. Finally, Section 5.5 concludes the chapter. The achievability proofs are presented in the appendix.

5.2 ACHIEVABLE REGIONS FOR MEMORYLESS BRC

In this section, we propose three achievable rate regions for the general two-user memoryless BRC. First, we provide the rate regions attained by the conventional decode-forward and compress-forward schemes, slightly adapted from the single-user relay channel to the current setup. Then, we propose a DCF scheme that combines the relay observation and the state feedback in such a way that the destinations can benefit from the additional side information to decode the messages. An achievable rate region associated with the proposed scheme is presented.

5.2.1 *Conventional schemes*

In the following, we present two achievable rate regions using two conventional schemes. They are adaptations of the existing schemes from the single-user relay channel [57]. Although the regions can easily proved by extending the schemes in [57] and proofs for similar settings can be found, e.g., in [60,61,74], we still include the achievability in the appendix for the sake of completeness. We recall that in the two conventional schemes the relay acts as if the feedback, i.e., the state information from users, is not available.

Decode-Forward (DF)

Based on DF, the source transmits the signal X, that is encoded with U_1 and U_2 (depending on the messages), the relay decodes both messages, i.e. the indices of U_1 and U_2, from its observation Y_r, then re-encodes them and transmits X_1, X_2 via the encoded signal X_r to the destinations.[1] Hence, provided that the source messages are decoded correctly by the relay, the latter can cooperate with the source and form a virtual transmitter with additional resources (e.g., antennas, transmit power). An achievable region is shown below.

[1] In this chapter, we use the notations U_k, X_k, and X_r instead of V_k and V_{12} to avoid confusion across the schemes and to remind the readers that the source and the relay are also functional separate.

Proposition 5.1. (DF inner bound) *An achievable rate region for the two-user memoryless state-dependent BRC by the DF scheme is given by the set of all rate pairs (R_1,R_2) satisfying*

$$R_1 \leq \min\{I(U_1X_1;Y_1|SQ), I(U_1;Y_r|X_1X_2U_2SQ)\}, \tag{124a}$$
$$R_2 \leq \min\{I(U_2X_2;Y_2|SQ), I(U_2;Y_r|X_1X_2U_1SQ)\}, \tag{124b}$$
$$R_1 + R_2 \leq I(U_1U_2;Y_r|X_1X_2SQ), \tag{124c}$$

for all pmf's that verify

$$p(y_1,y_2,y_r,u_1,u_2,x_1,x_2,x,x_r,s,q) = p(y_1,y_2,y_r|x,x_r,s) \times p(x_1)p(x_2)p(u_1|x_1)p(u_2|x_2)p(x|u_1,u_2,q)p(x_r|x_1,x_2,q)p(s)p(q).$$

Proof. The detailed proof is presented in Appendix 5.B. ■

The main ingredients of the above full-duplex DF scheme are the block-Markov superposition coding at the source and backward decoding at the destinations. Note that both transmitted signals at the source and the relay are independent of the instantaneous state of the channel due to the lack of CSIT. This feature will appear repeatedly also in the other schemes in this paper.

Compress-Forward (CF)

Based on CF, the transmitter sends the signal X, that is encoded with U_1 and U_2 (depending on the messages), the relay compresses the received signal Y_r, then encodes the description of the compressed signal $\hat{Y}$ and sends the encoded signal X_r to the destinations. Hence, provided that the description of the compressed signal are decoded correctly at both destinations, the latter can cooperate with the relay and form virtual receivers with additional resources (e.g., antennas, received power). An achievable region is shown below.

Proposition 5.2. (CF inner bound) *An achievable rate region for the two-user memoryless state-dependent BRC with the CF scheme is given by the set of all rate pairs (R_1,R_2) satisfying*

$$R_1 \leq \min\Big\{I(X_rU_1;Y_1|SQ) - I(Y_r;\hat{Y}|X_rU_1Y_1SQ), I(U_1;\hat{Y}Y_1|X_rSQ)\Big\}, \tag{125a}$$
$$R_2 \leq \min\Big\{I(X_rU_2;Y_2|SQ) - I(Y_r;\hat{Y}|X_rU_2Y_2SQ), I(U_2;\hat{Y}Y_2|X_rSQ)\Big\}, \tag{125b}$$

for all pmf's that verify

$$p(y_1,y_2,y_r,\hat{y},x_r,u_1,u_2,x,s,q) = p(y_1,y_2,y_r|x,x_r,s) \times p(\hat{y}|x_r,y_r,q)p(u_1)p(u_2)p(x_r|q)p(x|u_1,u_2,q)p(s)p(q),$$

for $k = 1, 2$.

Proof. The detailed proof is presented to Appendix 5.C. ■

The main ingredients of the above full-duplex CF scheme are the block-Markov superposition coding at the source, Wyner-Ziv compression at the relay, and backward decoding at the destinations. Instead of decoding the original messages, the relay compresses the received signal, encodes the compression index, and then broadcasts the coded signal to the destinations. Each destination decodes the message from its own observation together with the side information provided by the relay.

5.2.2 *Decode-Compress-Forward (DCF) scheme*

We now present a rate region achieved by the proposed DCF scheme. The details of the scheme are rather technical and are deferred to the appendix. The essential idea is to let the relay forward some useful side information based on the decoded source messages and the state feedback from the destinations, in such a way that the interference can be mitigated in an efficient way. Specifically, the source first sends the encoded signal X, while the relay decodes both messages (the indices of U_1 and U_2) from observation Y_r as in the DF scheme. Decoding both messages is necessary provided that in DCF scheme the relay takes partial responsibility of the source (multicasting side information as described in the schemes in Part I) who has full access to the original messages. Then, the relay compresses some function of the two decoded source messages and the CSI feedback from the destinations. Finally, the relay encodes the compression index (associated with $\hat{Y}$) with its own codebook and then transmits the encoded signal X_r to both users as in the CF scheme. Each destination recovers the desired message jointly with the compression index. As in both the DF and CF schemes, the source and the relay transmits simultaneously with block-Markov coding. An achievable rate region is presented below.

Theorem 5.3. (DCF inner bound) *An achievable region of memoryless state-dependent BRC with state feedback at the relay is given by the set of all rates pair (R_1, R_2) satisfying*

$$R_1 \leq \min\Big\{I_{21}, I_{31}, I_{41} + I_{51}\Big\}, \tag{126a}$$

$$R_2 \leq \min\Big\{I_{22}, I_{32}, I_{42} + I_{52}\Big\}, \tag{126b}$$

$$R_1 + R_2 \leq I_1, \tag{126c}$$

where, for $k \neq m \in \{1,2\}$,

$$I_1 \triangleq I(U_1U_2; Y_r|X_rSQ), \tag{127a}$$
$$I_{2k} \triangleq I(U_k; Y_r|U_mX_rSQ), \tag{127b}$$
$$I_{3k} \triangleq I(U_k; Y_k\hat{Y}|X_rSQ), \tag{127c}$$
$$I_{4k} \triangleq I(U_k; Y_k|SQ), \tag{127d}$$
$$I_{5k} \triangleq -I(\hat{Y}; U_m|U_kX_rY_kSQ) + I(X_r; Y_k|U_kSQ), \tag{127e}$$

for all pmf's that verifies

$$\begin{aligned} p(y_1, y_2, y_r, \hat{y}, x, x_r, u_1, u_2, s, q) = \\ p(\hat{y}|x_r, u_1, u_2, s, q)p(y_1, y_2, y_r|x, x_r, s) \\ \times p(x_r|q)p(x|u_1, u_2, q)p(u_1)p(u_2)p(s)p(q). \end{aligned} \tag{128}$$

Sketch of proof. Without going into the details, the rate constraints can be interpreted as follows. First, the constraints $R_1 \leq I_{21}$, $R_2 \leq I_{22}$, and $R_1 + R_2 \leq I_1$ are the conditions required for the relay to recover both messages reliably. Then, $R_k \leq I_{3k}$ is the condition for destination k to decode the desired message with the help of the useful side information, provided that the latter is reliably delivered to destination k. Finally, the condition $R_k \leq I_{4k} + I_{5k}$ can guarantee the successful delivery of the side information at destination k as well as the original message (recovered simply from its observation) simultaneously. Details of the achievability are deferred to the appendix 5.A for better readability. ■

In a single-user single-relay network, decoding completely the source message and then compressing the recovered information is usually considered inefficient. Remarkably, this is not the case in a multi-user network such as our two-user BRC when state feedback is available at the relay. With multiple users, the bottleneck is the interference at the receivers' side, which can be reduced with the help of the relay if the latter can "learn" the interference. It is worth mentioning that our scheme is conceptually different from the conventional DF and CF schemes described at the beginning of the section. Without state feedback, even if the relay can perfectly decode the source messages, it cannot reduce the interference at the receivers' side simply because the interference is a function of the random state that is unknown to the relay. Therefore, the conventional relaying methods in such a context can only be considered as a way to reinforce each user's signal, but not as a way to mitigate interference. As mentioned at the beginning of the paper, we assume that the cost for the relay to obtain state feedback is relatively low in terms of power and bandwidth consumption in our setting since such communication is localized. In contrast, feedback from the receivers to the source as consider in [2] would be much more costly because of the longer distance as compared to the relay-destination link. From the complexity point of view, the main relaying operations of the proposed schemes include both of the DF and CF schemes. Namely, the relay decodes the source messages as in the DF scheme, reconstructs the interferences based on the state feedback, and compresses the interferences as in the CF scheme. The receivers decode the indices of the side information from the relay and use it to decode the source messages, also an

operation comparable to the CF scheme. That being said, the proposed scheme is not significantly more complex than the DF and CF scheme.

It is important to differentiate our scheme from the partial decode and compress forward (pDCF) scheme in the Theorem 7 of Cover and El Gamal's original paper [57]. Essentially, with the pDCF, the relay decodes part of the source message and compresses the rest of the received signal and forward the compression index in a partially coherent way with respect to the source. Therefore, the pDCF scheme contains naturally the DF and CF schemes as special cases, and outperform them in general. The original pDCF was designed for the single-user case, but can be extended straightforward to our setup of broadcast relay channel, just as what we do with the conventional single-user DF and CF schemes in Proposition 1 and 2. Our DCF scheme, although also based on the combination of DF and CF, is different from the pDCF of Cover and El Gamal with a special emphasis on the interference in a multi-user network. The proposed scheme is, by design, intended for the broadcast (thus multi-user) setting and the availability of state feedback at the relay. Specifically, in our DCF scheme, the relay is required to decode the source messages, to reconstruct the interference (thanks to state feedback), and then to compress the interferences experienced by the destinations (instead of compressing the relay's own received signal).

By removing the decoding constraints I_{21}, I_{22}, and I_1 at the relay, we obtain straightforwardly the achievable rate region for the BC with state feedback derived in [38]. Remark that the DCF region does not contains either CF or DF of Proposition 1 and 2 in general. Nevertheless, it is possible to revise the current DCF scheme to have a more general one that always includes both DF and CF as special cases. To that end, it suffices to 1) superimpose U_k over a U_k' at the source as we do with X_k and U_k in the DF case as well as superimposing X_r over (U_1', U_2') to cover the conventional DF; and 2) let $\hat{Y}$ depend on not only X_r, U_1, U_2, but also on Y_r, to cover the conventional CF. With such modifications, we may obtain a rate region that contains the CF and DF regions of Proposition 1 and 2. However, the expressions of such a general region will be so involved that it is hard to evaluate numerically even for the Gaussian case. Since we are mainly interested in the interference-limited regime, we choose to keep the less general but simpler region as it is presented in Theorem 1.

Moreover, note that the presented rate regions for the general memoryless channel depend on the choice of pmf's involving the auxiliary random variables. Unfortunately, little intuition on the choice of "good" pmf's is provided from the mathematical expressions. As a result, it is in general extremely hard to evaluate the optimal rate region that includes all possible pmf choices. In the following, we shall focus on a special case, the fading Gaussian noise BRC, and propose particular choices on the pmf's, which allows us to evaluate the achievable rates.

5.3 THE FADING GAUSSIAN NOISE BRC

For the fading Gaussian noise BRC described by (12), we consider two particular choices of the pmf's of the input variables $(X, X_r, U_1, U_2, \hat{Y}, Q)$. We regard the choices as different transmission schemes within the class of DCF, namely, the

simultaneous-emitted decode-compress-and-forward (SE DCF) and the time-division decode-compress-and-forward (TD DCF). We shall specify the chosen pmf's and the corresponding rate regions and then discuss the operational meaning of these choices.

5.3.1 *Simultaneous-Emitted Decode-Compress-Forward*

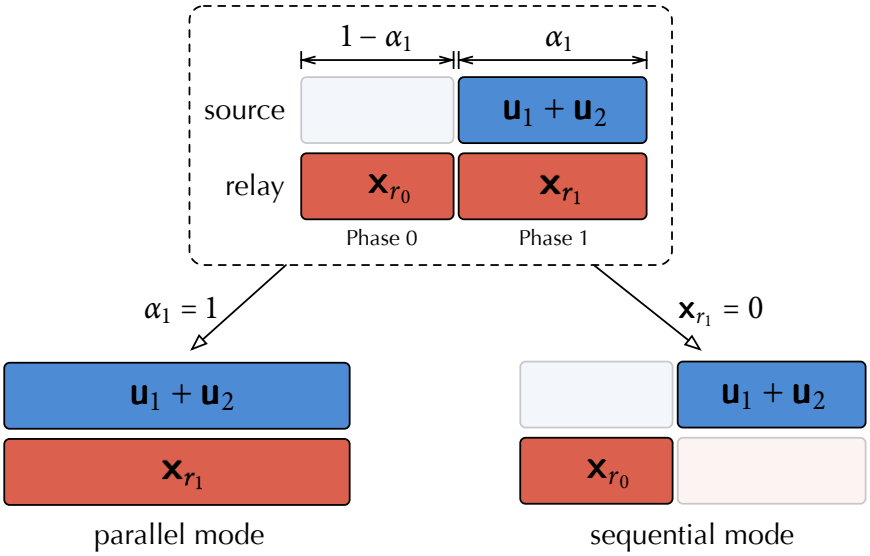

Figure 17: Block structure of the SE-DCF (Simultaneously emitted DCF) in its general form and special modes.

First, we present the SE DCF. The time-sharing RV Q is binary with $P(Q=1)=\alpha_1$ and $P(Q=0)=1-\alpha_1$ with $\alpha_1 \in [0,1]$. The triple $(X, X_r, \hat{Y})$ is controlled by Q as shown in Table 2.

Table 2: RV choice in SE-DCF

	X	X_r	$\hat{Y}$
$Q=0$	0	$\mathbf{x}_{r_0}$	0
$Q=1$	$\mathbf{u}_1+\mathbf{u}_2$	$\mathbf{x}_{r_1}$	$[\boldsymbol{H}_1\mathbf{u}_2\ \boldsymbol{H}_2\mathbf{u}_1]+[\hat{\mathbf{z}}_1\ \hat{\mathbf{z}}_2]$

We choose the related RVs' distributions to be $\mathbf{u}_1 \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{Q}_1)$, $\mathbf{u}_2 \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{Q}_2)$, $\mathbf{x}_{r_1} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{Q}_{r_1})$ and $\mathbf{x}_{r_0} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{Q}_{r_0})$, where the following power constraints are fulfilled,

$$\begin{cases} \alpha_1 \left(\operatorname{tr}(\boldsymbol{Q}_1) + \operatorname{tr}(\boldsymbol{Q}_2)\right) & \leq P_s, \\ \alpha_1 \operatorname{tr}(\boldsymbol{Q}_{r_1}) + (1-\alpha_1)\operatorname{tr}(\boldsymbol{Q}_{r_0}) & \leq P_r. \end{cases}$$

Intuitively, the signals $\mathbf{u}_1$ and $\mathbf{u}_2$ carrying two independent information flows are sent *simultaneously* through X. The interference to both users are then compressed to $\hat{Y}$ with compression noises $\hat{\mathbf{z}}_1 \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{N}_1)$, $\hat{\mathbf{z}}_2 \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{N}_2)$. The relay transmitted signal $\mathbf{x}_{r_0}$ or $\mathbf{x}_{r_1}$, known as the description of compression $\hat{Y}$, is identified with the auxiliary variable X_r depending on the choice of Q. Even though the description of compression is multicast to both users in general, each receiver exploits simply part of the compressed information $\hat{Y}$ and ignore the other part, which in fact makes that the relay "broadcast" part of information in $\hat{Y}$ to its intended user.

With the above RVs' choices, the mutual information expressions involved in the achievable rate region of Theorem 5.3 can be explicitly evaluated as (129a)-(129e).

Corollary 5.4 (SE-DCF Inner bound)**.** *For the fading Gaussian noise BRC with the state feedback of SDC at the relay, a rate pair (R_1, R_2) is achievable if* (126a), (126b), (126c) *with the definitions given in* (129a)-(129e) *are satisfied for some $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_{r_1}, \boldsymbol{Q}_{r_0}, \boldsymbol{N}_1, \boldsymbol{N}_2 \succeq 0$, and $\alpha_1 \in [0,1]$ with $\alpha_1(\mathrm{tr}(\boldsymbol{Q}_1) + \mathrm{tr}(\boldsymbol{Q}_2)) \leq P_s$ and $\alpha_1 \,\mathrm{tr}(\boldsymbol{Q}_{r_1}) + (1-\alpha_1)\,\mathrm{tr}(\boldsymbol{Q}_{r_0}) \leq P_r$.*

$$I_1 = \alpha_1 \mathbb{E} \log\det\left[\boldsymbol{I} + \boldsymbol{H}_r(\boldsymbol{Q}_1 + \boldsymbol{Q}_2)\boldsymbol{H}_r^H\right], \tag{129a}$$

$$I_{2k} = \alpha_1 \mathbb{E} \log\det\left[\boldsymbol{I} + \boldsymbol{H}_r\boldsymbol{Q}_k\boldsymbol{H}_r^H\right], \tag{129b}$$

$$\begin{aligned} I_{3k} = \alpha_1 \mathbb{E} \log\det\Bigg[&\boldsymbol{I} + \boldsymbol{Q}_k\boldsymbol{H}_m^H\boldsymbol{N}_m^{-1}\boldsymbol{H}_m \\ &+ \boldsymbol{Q}_k\boldsymbol{H}_k^H\left(\boldsymbol{I} + \boldsymbol{H}_k\boldsymbol{Q}_m\boldsymbol{H}_k^H\left(\boldsymbol{H}_k\boldsymbol{Q}_m\boldsymbol{H}_k^H + \boldsymbol{N}_k\right)^{-1}\boldsymbol{N}_k\right)^{-1}\boldsymbol{H}_k\Bigg], \end{aligned} \tag{129c}$$

$$I_{4k} = \alpha_1 \mathbb{E} \log\left[\frac{\det(\boldsymbol{I} + \boldsymbol{G}_k\boldsymbol{Q}_{r_1}\boldsymbol{G}_k^H + \boldsymbol{H}_k(\boldsymbol{Q}_1 + \boldsymbol{Q}_2)\boldsymbol{H}_k^H)}{\det(\boldsymbol{I} + \boldsymbol{G}_k\boldsymbol{Q}_{r_1}\boldsymbol{G}_k^H + \boldsymbol{H}_k\boldsymbol{Q}_m\boldsymbol{H}_k^H)}\right], \tag{129d}$$

$$\begin{aligned} I_{5k} = \alpha_1 \mathbb{E} \log&\left[\frac{\det(\boldsymbol{I} + \boldsymbol{G}_k\boldsymbol{Q}_{r_1}\boldsymbol{G}_k^H + \boldsymbol{H}_k\boldsymbol{Q}_m\boldsymbol{H}_k^H)\det\left(\boldsymbol{I} + \boldsymbol{H}_k\boldsymbol{Q}_m\boldsymbol{H}_k^H\right)}{\det(\boldsymbol{H}_k\boldsymbol{Q}_m\boldsymbol{H}_k^H + \boldsymbol{I})\det(\boldsymbol{I} + \boldsymbol{H}_k\boldsymbol{Q}_m\boldsymbol{H}_k^H(\boldsymbol{I} + \boldsymbol{N}_k^{-1}))}\right] \\ &+ (1-\alpha_1)\mathbb{E} \log\left[\det(\boldsymbol{I} + \boldsymbol{G}_k\boldsymbol{Q}_{r_0}\boldsymbol{G}_k^H)\right]. \end{aligned} \tag{129e}$$

Remark 5.1. *Before verifying the high SNR performance of the SE DCF region, we assume $\boldsymbol{Q}_k = P_k\boldsymbol{I}$, $\boldsymbol{Q}_{r_l} = P_{r_l}\boldsymbol{I}$, $P_k = P_s^{\beta_k}$, $P_{r_l} = P_r{}^{\beta_{r_l}}$ for $\beta_k, \beta_l \in [0,1]$, $k \in \{1,2\}$, $l \in \{0,1\}$, and $P_r = P_s = P$ and set $n_{rr} = n_{ts} = n_{tr} = 2$, $n_{r1} = n_{r2} = 1$. Further, $\boldsymbol{N}_1 = \boldsymbol{N}_2 = N\boldsymbol{I}$ where $N = O(P)$ if P tends to infinity. As $P \to \infty$, each of the equations* (129a)-(129e) *is evaluated as follows.*

$$I_1 = 2\alpha_1 \log(P) + O(\log(P)), \tag{130a}$$

$$I_{2k} = 2\alpha_1 \log(P_k) + O(\log(P)) = 2\alpha_1\beta_k \log(P) + O(\log(P)), \tag{130b}$$

$$I_{3k} = 2\alpha_1\beta_k \log(P) + O(\log(P)), \tag{130c}$$

$$\begin{aligned} I_{4k} + I_{5k} = &\alpha_1 \max\{\beta_k - \beta_m, \beta_{r_1} - \beta_m\}\log(P) + (1-\alpha_1)\beta_{r_0}\log(P) \\ &+ O(\log(P)). \end{aligned} \tag{130d}$$

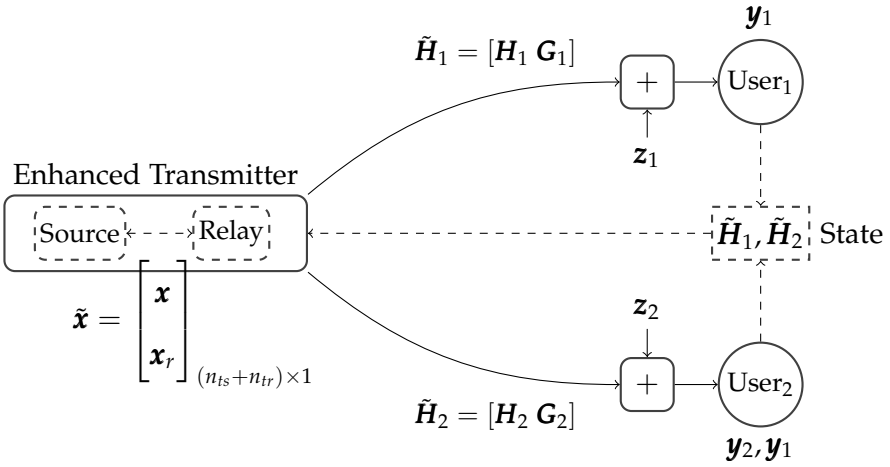

Figure 18: The enhanced channel for outer bound of Gaussian BRC.

We note that I_{2k} and I_{3k} provide the same DoF. In particular, $\lim_{P\to\infty} I_{2k}$ represents the DoF of a real 2-by-2 point-to-point channel while I_{3k} shows the DoF of an artificial 2-by-2 channel with two input antennas at the transmitter's side, a single output antenna that provides user observation and a single virtual output antenna that observes the side information. Thus, it is readily seen that both n_{rr} and n_{ts} should be set at least as large as $n_{r1} + n_{r2}$ to ensure such equality on the DoF by I_{2k} and I_{3k}. Given that the DoF increases with β_{r_0} and β_{r_1} in $I_{4k} + I_{5k}$, the optimal choice of them are $\beta_{r_0} = \beta_{r_1} = 1 \geq \beta_k$. With $\beta_{r_1} = 1$, we know that $\max\{\beta_k - \beta_m, \beta_{r_1} - \beta_m\}$ equals to $\beta_{r_1} - \beta_m = 1 - \beta_m$. Then, the DoF region achieved by SE DCF can be expressed as below.

$$\left\{ (d_1, d_2) \in \mathbb{R}_+, \forall k, m \in \{1,2\}, k \neq m \;\middle|\; \begin{array}{rl} d_1 + d_2 & \leq 2\alpha_1, \\ d_k & \leq 2\alpha_1\beta_k, \\ d_k & \leq 1 - \alpha_1\beta_m, \end{array} \right\} \tag{131}$$

Therefore, the optimal sum DoF $\frac{4}{3}$ can be obtained from the above region by letting $\beta_1 = \beta_2 = \frac{1}{3\alpha_1} = \frac{1}{3}$. The optimality on the sum DoF can be proved by a standard upper bound using genie-aided technique, i.e., generating a degraded BC by allowing one of the users to have the other user's observation for free. In particular, we create an enhanced transmitter who has four antennas and is allocated with $P_s + P_r$ power in total. Such transmitter can be interpreted as combining the source and the relay in BRC with a cost free link where original messages and the delayed CSI can be shared instantaneously, as is shown in the Fig. 18. In the plot, $\tilde{\boldsymbol{x}} = \left[\boldsymbol{x}^T\ \boldsymbol{x}_r^T\right]^T \in \mathcal{C}^{(n_{ts}+n_{tr})\times 1}$ is the new channel input that is subject to $\frac{1}{n}\sum_{i=1}^{n} \|\tilde{\boldsymbol{x}}\|^2 \leq (P_s + P_r)$ and $\tilde{\boldsymbol{H}}_k = [\boldsymbol{H}_k\ \boldsymbol{G}_k] \in \mathcal{C}^{n_{rk}\times(n_{ts}+n_{tr})}$ is the corresponding channel coefficient of enhanced channel k.

Then, we can apply directly the sum DoF upper bound mentioned in [2] to the enhanced two-user MISO (4-by-1) BC, that is, the optimal sum DoF is $\frac{4}{3}$. Note that such "enhancement" can also be used when characterizing the achievable rate region or generalized DoF region in Gaussian BRC.

5.3.2 *Time-Division Decode-Compress-Forward Scheme*

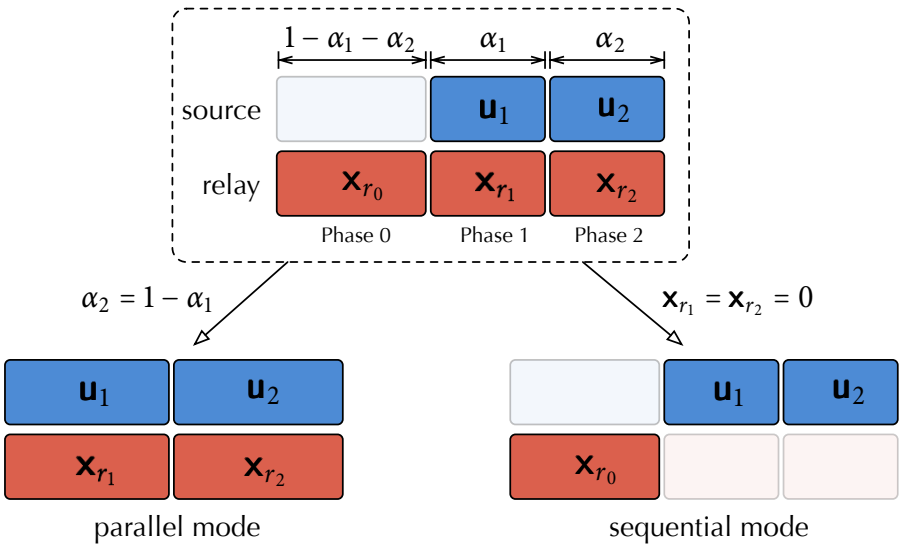

Figure 19: Block structure of the TD-DCF (Time-division DCF) in its general form and special modes.

Next, we consider the TD DCF. The time-sharing RV Q is ternary with $P(Q=1)=\alpha_1$, $P(Q=2)=\alpha_2$, and $P(Q=0)=1-\alpha_1-\alpha_2$ with $\alpha_1,\alpha_2,\alpha_1+\alpha_2\in[0,1]$. The triple $(X, X_r, \hat{Y})$ is controlled by Q as shown in Table 3.

Table 3: RV choice in TD-DCF

	X	X_r	$\hat{Y}$
$Q=0$	0	$\mathbf{x}_{r_0}$	0
$Q=1$	$\mathbf{u}_1$	$\mathbf{x}_{r_1}$	$\boldsymbol{H}_2\mathbf{u}_1+\hat{\mathbf{z}}_1$
$Q=2$	$\mathbf{u}_2$	$\mathbf{x}_{r_2}$	$\boldsymbol{H}_1\mathbf{u}_2+\hat{\mathbf{z}}_2$

We choose the related RVs' distributions to be $\mathbf{u}_1\sim\mathcal{CN}(\mathbf{0},\boldsymbol{Q}_1)$, $\mathbf{u}_2\sim\mathcal{CN}(\mathbf{0},\boldsymbol{Q}_2)$, $\mathbf{x}_{r_q}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{Q}_{r_q})$, $q=0,1,2$, where the following power contraints are satisfied,

$$\begin{cases}\alpha_1\operatorname{tr}(\boldsymbol{Q}_1)+\alpha_2\operatorname{tr}(\boldsymbol{Q}_2) & \le P_s,\\ \alpha_1\operatorname{tr}(\boldsymbol{Q}_{r_1})+\alpha_2\operatorname{tr}(\boldsymbol{Q}_{r_2})+(1-\alpha_1-\alpha_2)\operatorname{tr}(\boldsymbol{Q}_{r_0}) & \le P_r\end{cases}$$

Intuitively, the signals $\mathbf{u}_1$ and $\mathbf{u}_2$ carrying two independent information flows are sent *successively* through X. This is in contrast to the SE DCF in which the two flows are sent simultaneously. Each user overhears the other user's signal, i.e., user k gets the noisy version of $\boldsymbol{H}_k\mathbf{u}_j$ when user j's signal is sent ($Q=j$). The overheard signals are compressed to $\hat{Y}$ with compression noises $\hat{\mathbf{z}}_1\sim\mathcal{CN}(\mathbf{0},\boldsymbol{N}_1)$,

$\hat{\mathbf{z}}_2 \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{N}_2)$. Since $\boldsymbol{H}_k \mathbf{u}_j$ is overheard by destination k and compressed in $\hat{Y}$, this exact part of the side information can be recovered with no cost when SNR tends to infinity. The relay transmitted signal $\mathbf{x}_{r_0}, \mathbf{x}_{r_1}, \mathbf{x}_{r_2}$, known as the description of compression $\hat{Y}$, is identified with the auxiliary variable X_r depending on the choice of Q.

With the above RVs' choices, the mutual information expressions involved in the achievable rate region of Theorem 5.3 can be explicitly evaluated as (132a)-(132e).

Corollary 5.5 (TD DCF Inner bound)**.** *For the fading Gaussian noise BRC with the state feedback of SDC at the relay, a rate pair (R_1, R_2) is achievable if* (126a), (126b), (126c) *with the definitions given in* (132a)-(132e) *are satisfied for some $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_{r_1}, \boldsymbol{Q}_{r_2}, \boldsymbol{Q}_{r_0}, \boldsymbol{N}_1, \boldsymbol{N}_2 \succeq 0$, and $\alpha_1, \alpha_2, \alpha_1 + \alpha_2 \in [0,1]$ with $\alpha_1 \operatorname{tr}(\boldsymbol{Q}_1) + \alpha_2 \operatorname{tr}(\boldsymbol{Q}_2) \leq P_s$ and $\alpha_1 \operatorname{tr}(\boldsymbol{Q}_{r_1}) + \alpha_2 \operatorname{tr}(\boldsymbol{Q}_{r_2}) + (1 - \alpha_1 - \alpha_2) \operatorname{tr}(\boldsymbol{Q}_{r_0}) \leq P_r$.*

$$I_1 = \alpha_1 \mathbb{E} \log\det\left[\boldsymbol{I} + \boldsymbol{H}_r \boldsymbol{Q}_1 \boldsymbol{H}_r^H\right] + \alpha_2 \mathbb{E} \log\det\left[\boldsymbol{I} + \boldsymbol{H}_r \boldsymbol{Q}_2 \boldsymbol{H}_r^H\right], \quad (132a)$$

$$I_{2k} = \alpha_k \mathbb{E} \log\det\left[\boldsymbol{I} + \boldsymbol{H}_r \boldsymbol{Q}_k \boldsymbol{H}_r^H\right], \quad (132b)$$

$$I_{3k} = \alpha_k \mathbb{E} \log\det\left[\boldsymbol{I} + \boldsymbol{Q}_k (\boldsymbol{H}_k^H \boldsymbol{H}_k + \boldsymbol{H}_m^H \boldsymbol{N}_k^{-1} \boldsymbol{H}_m)\right], \quad (132c)$$

$$I_{4k} = \alpha_k \mathbb{E} \log\left[\frac{\det\left(\boldsymbol{I} + \boldsymbol{H}_k \boldsymbol{Q}_k \boldsymbol{H}_k^H + \boldsymbol{G}_k \boldsymbol{Q}_{r_k} \boldsymbol{G}_k^H\right)}{\det\left(\boldsymbol{I} + \boldsymbol{G}_k \boldsymbol{Q}_{r_k} \boldsymbol{G}_k^H\right)}\right], \quad (132d)$$

$$I_{5k} = \alpha_m \mathbb{E} \log\left[\frac{\det\left(\boldsymbol{I} + \boldsymbol{H}_k \boldsymbol{Q}_m \boldsymbol{H}_k^H + \boldsymbol{G}_k \boldsymbol{Q}_{r_m} \boldsymbol{G}_k^H\right)}{\det(\boldsymbol{I} + \boldsymbol{Q}_m (\boldsymbol{H}_k^H (\boldsymbol{I} + \boldsymbol{N}_m^{-1}) \boldsymbol{H}_k)) \det(\boldsymbol{I} + \boldsymbol{G}_k \boldsymbol{Q}_{r_0} \boldsymbol{G}_k^H)}\right]$$
$$+ \alpha_k \mathbb{E} \log\left[\frac{\det\left(\boldsymbol{I} + \boldsymbol{G}_k \boldsymbol{Q}_{r_k} \boldsymbol{G}_k^H\right)}{\det(\boldsymbol{I} + \boldsymbol{G}_k \boldsymbol{Q}_{r_0} \boldsymbol{G}_k^H)}\right] + \mathbb{E} \log\det(\boldsymbol{I} + \boldsymbol{G}_k \boldsymbol{Q}_{r_0} \boldsymbol{G}_k^H). \quad (132e)$$

Remark 5.2. *In order to evaluate the SE DCF region in high SNR regime, we assume $\boldsymbol{Q}_k = P_k \boldsymbol{I}, \boldsymbol{Q}_{r_l} = P_{r_l} \boldsymbol{I}, P_k = P_s^{\beta_k}, P_{r_l} = P_r{}^{\beta_{r_l}}$ for $\beta_k, \beta_l \in [0,1], k \in \{1,2\}, l \in \{0,1,2\}$, and $P_r = P_s = P$ and set $n_{rr} = n_{ts} = n_{tr} = 2$, $n_{r1} = n_{r2} = 1$. Further, $\boldsymbol{N}_1 = \boldsymbol{N}_2 = N\boldsymbol{I}$ where $N = O(P)$ if P tends to infinity. As $P \to \infty$, each of the equations* (132a)-(132e) *is computed as follows.*

$$I_1 = 2\alpha_1\beta_1 \log(P) + 2\alpha_2\beta_2 \log(P) + O(\log(P)), \quad (133a)$$

$$I_{2k} = 2\alpha_k\beta_k \log(P) + O(\log(P)), \quad (133b)$$

$$I_{3k} = 2\alpha_k\beta_k \log(P) + O(\log(P)), \quad (133c)$$

$$I_{4k} + I_{5k} = \beta_{r_0} \log(P) + \alpha_k \max\{\beta_k - \beta_{r_0}, \beta_{r_k} - \beta_{r_0}\} \log(P)$$
$$+ \alpha_m \max\{\beta_m - \beta_{r_0}, \beta_{r_m} - \beta_{r_0}\} \log(P) - \alpha_m \beta_m \log(P)$$
$$+ O(\log(P)). \quad (133d)$$

In the TD DCF, I_1 is redundant while I_{3k} is the same as I_{2k}. With the same reasoning for DoF of SE DCF scheme, we set $\beta_{r_1} = \beta_{r_2} = 1$ and thus $\max\{\beta_k - \beta_{r_0}, \beta_{r_k} - \beta_{r_0}\} = 1 - \beta_{r_0}$ and $\max\{\beta_m - \beta_{r_0}, \beta_{r_m} - \beta_{r_0}\} = 1 - \beta_{r_0}$. Then, we obtain the achievable DoF

region as expressed below.

$$\left\{ (d_1, d_2) \in \mathbb{R}_+ \;\middle|\; \begin{array}{ll} d_k & \leq 2\alpha_k\beta_k, \\ d_k & \leq (1-\alpha_1-\alpha_2)\beta_{r_0} + \alpha_1 + \alpha_2 - \alpha_m\beta_m \end{array} \right\} \tag{134}$$

for $k, m \in \{1, 2\}, k \neq m$.

Given $1 - \alpha_1 - \alpha_2 \geq 0$, *the optimal* β_{r_0} *equals to one. Hence, the sum DoF fulfills*

$$d_1 + d_2 \leq \min\{2(\alpha_1\beta_1 + \alpha_2\beta_2), 1 + \alpha_1\beta_1, 1 + \alpha_2\beta_2, 2 - \alpha_1\beta_1 - \alpha_2\beta_2\}$$

where $\frac{4}{3}$ *DoF is achievable by letting* $\beta_1 = \beta_2 = 1$ *and* $\alpha_1 = \alpha_2 = \frac{1}{3}$. *With these choice of parameters, letting* $\beta_{r_1} = \beta_{r_2} = 0$ *(sequential mode) does not change the achievable sum DoF. Similar to the analysis for SE DCF, TD DCF region is also sum-DoF optimal.*

5.3.3 *Implementation of SE/TD DCF*

In both SE and TD DCF, whenever $Q = 0$, the source remains silent ($X = 0$) while the relay multicast the side information. Note that U_1 and U_2 are independent of Q but are not transmitted when $Q = 0$ since $X = 0$. Therefore, the choice of U_1 and U_2 matters only when $Q \neq 0$. In the following, we explain the practical meaning of the probabilistic time-sharing variable Q. As will be shown later, each block can be structured in a deterministic way with Q indicating the transmission phase within each block. As a result, the probability of $Q = q$ is related to the portion of the length of phase q. Therefore, we can implement SE DCF and TD DCF as is illustrated in Fig. 17 and Fig. 19. By tuning the parameters of the proposed DCF schemes, it is obvious that the performance in terms of throughput and complexity varies. Two extremes of the parameter configuration are the "parallel" mode and "sequential" mode. In the parallel mode, both the source and relay transmit during all the block, i.e., phase 0 is eliminated. In the sequential mode, the source and relay take turns to transmit without overlap. The sequential mode can thus be implemented even with half-duplex relays.

In both SE and TD DCF, the transmissions last for B blocks. The structure of each block b is given in Fig. 17 and Fig. 19. The relay transmits the encoded signal related to the side information for the previous block $b - 1$. The encoded signal is $(x_{r_0},\ x_{r_1})$ in the SE-DCF case and $(x_{r_0},\ x_{r_1},\ x_{r_2})$ in the TD-DCF case, as shown in Fig. 17 and Fig. 19 respectively. In phase 0 of both SE and TD DCF, the source remains silent. Then, the source transmits either the mixture of both information flows encoded in u_1 and u_2 in phase 1 (SE-DCF, Fig. 17), or successively u_1 in phase 1 and u_2 in phase 2 (TD-DCF, Fig. 19). In both cases, the relay listens and then decodes the source messages and generates the side information at the end of the block once the CSI feedback is available from the destinations.

5.3.4 *Relation with Erasure BRC*

We recall that there is similarity between erasure and Gaussian channels as mentioned in Chapter 3, that is, the DoF region in Gaussian BC [2] has the same form as the capacity region of erasure BC [27, 28] for two user case. Additionally, the resemblance can be extended to the secure communication scenario, i.e., secure DoF region in Gaussian BC [29] and secure capacity region in the erasure BC [30]. However, such uniformity is not proved until now for the investigated BRC. In fact, the gap of sum-rate between the lower bound by adapted DCF variants and the upper bound by the aforementioned genie-aided technique (in Section 5.3.1) remains large for the erasure BRC even in the symmetric case. This may be due to the looseness of the selected upper bound in erasure case. Lacking state information at the source makes it unlikely to perform in erasure channel a "joint encoding" between source and relay whereas in genie-aided channel (for upper bound), the delayed state is global and hence the cooperation is available, e.g., the source can even predict what the relay will send. In contrast, the absence of states in Gaussian channel is not as destructive to DoF for the reason that the messages are conveyed at full rate via SRC given $n_{rr} = n_{ts} = 2$.[2] In order to better characterize the capacity in the erasure BRC with state feedback at relay, a stronger upper bound technique is in need.

5.4 NUMERICAL RESULTS

In order to assess the performance of the schemes proposed in Section 5.3, we have already verified the sum DoF of the two variants (SE DCF/TD DCF) in Section 5.3 which turns out to be both optimal within high SNR regime. Hence, we investigate through numerical simulation the sum-rate achieved by DCF's variants in a two-user symmetric Gaussian fading BRC in presence of i.i.d. Rayleigh fading. In the following, we start by introducing the simulation parameters, and then compare the proposed schemes with conventional schemes, followed by some discussions.

5.4.1 *Simulation parameters*

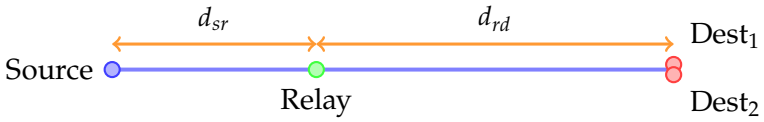

Figure 20: The one-dimensional deployment of a two-user broadcast relay channel used in simulation.

In the following numerical examples, we assume that $n_{ts} = n_{tr} = n_{rr} = 2$ and $n_{r1} = n_{r2} = 1$. In other words, the source-destination and relay-destination

[2] n_{tr} can be set to one since relay multicast the compression index which serves as *one* observation at the receiver.

channels are MISO channels, while the source-relay channel is a MIMO channel. This setup is motivated by the fact that both the macrocell and smallcell BS may have multiple antennas while it is generally harder for the terminals to have more than one antenna. In addition, we take into account a propagation model based only on the pathloss. In particular, pathloss is defined as $L = d^{-n_p}$, where d denotes the distance between communication nodes and n_p stands for the pathloss exponent. We let $n_p = 2$, which corresponds to free-space propagation.[3] Therefore, the channel coefficients are rewritten as $L_{sd}\,\boldsymbol{H}_k$, $L_{sr}\,\boldsymbol{H}_r$, $L_{rd}\,\boldsymbol{G}_k$ for respective states of SDC, SRC, RDC, with L_{sd}, L_{sr}, L_{rd} being the pathloss for the channels and (d_{sd}, d_{sr}, d_{rd}) the corresponding distances of (SDC, SRC, RDC). To simplify the evaluation, we adopt a one-dimensional deployment as shown in Fig. 20, the source, the relay and the destinations are strictly aligned on a straight line where the distance between two users is neglected so that the two receivers are viewed as a single point in the deployment and yet have spatially i.i.d. channels. Hence, the distances between communication nodes are normalized as $(d_{sr} + d_{rd})/d_{sd} = 1$. We let $d_{sd} = 10$ so that 9 integer-valued combinations of (d_{sr}, d_{rd}) are picked. At both source and relay nodes, the same power constraints ($P_s = P_r = P$) are taken. Since the noises are normalized, we define SNR $= \frac{P L_{sd}}{2}$, which is the maximum SNR received at the users from the source.

In particular, in SE DCF, we let $\alpha_1 = \alpha \in [0, 1]$, $\boldsymbol{Q}_k = P_k \boldsymbol{I}_2$, $k = 1, 2$, and $\boldsymbol{Q}_{r_q} = P_{r_q} \boldsymbol{I}_2$, $q = 0, 1$. The power constraint in Corollary 5.4 implies that $\alpha(P_1 + P_2) \leq P/2$ and $(1 - \alpha)P_{r_0} + \alpha P_{r_1} \leq P/2$. In the TD DCF scheme, we equalize the duration of phase 1 and 2 as $\alpha_1 = \alpha_2 = \frac{1}{2}\alpha \in [0, \frac{1}{2}]$, set $\boldsymbol{Q}_k = P_k \boldsymbol{I}_2$, $k = 1, 2$, and $\boldsymbol{Q}_{r_q} = P_{r_q} \boldsymbol{I}_2$, $q = 0, 1, 2$. The power constraint in Corollary 5.5 implies that $\alpha(P_1 + P_2) \leq P$ and $2(1 - \alpha)P_{r_0} + \alpha(P_{r_1} + P_{r_2}) \leq P$. Note that we optimize α and the power allocation individually in the simulations of all variants and modes.

In the simulations, we include three baseline schemes, namely, the BC without relay, the BRC with DF (Proposition 5.1) and the BRC with CF (Proposition 5.2) in the comparisons with SE/TD DCF. It is worth mentioning that only SE/TD DCF exploits state feedback while the reference schemes do not. Furthermore, all schemes are subject to the "no CSIT" constraint. We consider three levels of received SNR: the low-to-medium (7dB), the medium-to-high (27dB) and the high level (47dB). We focus on the sum-rate performance in bits per channel use versus the normalized relay position $\beta = d_{sr}/d_{sd}$.

5.4.2 *DCF schemes versus the baseline schemes*

From Fig. 21, Fig. 22 and Fig. 23, the observations can be obtained followed by the discussion:

- Although not explicit in the figures, the first observation from the results is that each of the three baseline schemes achieves its respective optimal sum-rate with the strategy of the single-user transmission. Indeed, since no CSI is available at any of the transmitters, no gains are provided by serving more than one user at each transmission with the baseline schemes and no

[3] Other values of the pathloss exponent can be considered according to different propagation environments.

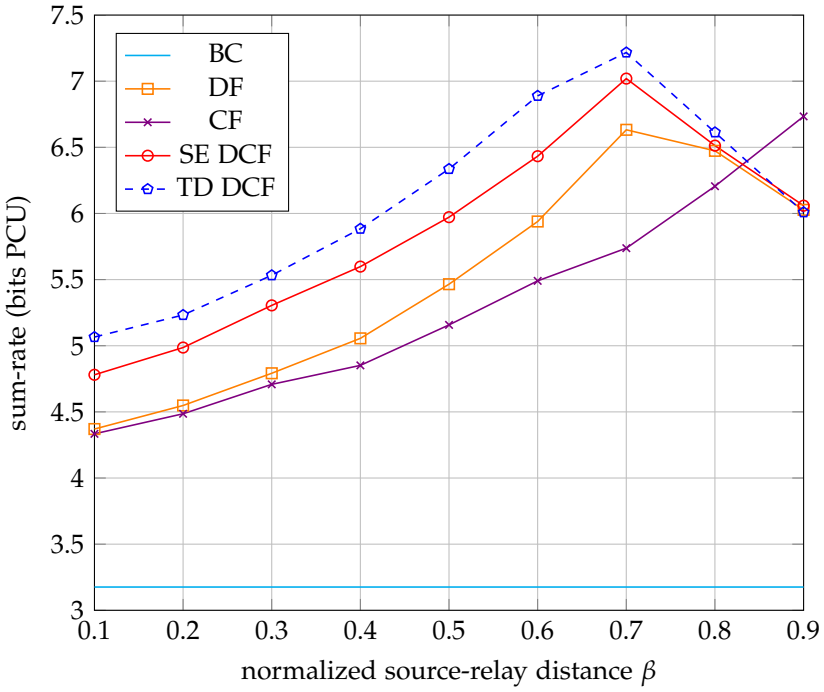

Figure 21: Sum-rate performance of the SE/TD DCF and the baseline schemes (BC/DF/CF) when SNR = 7dB.

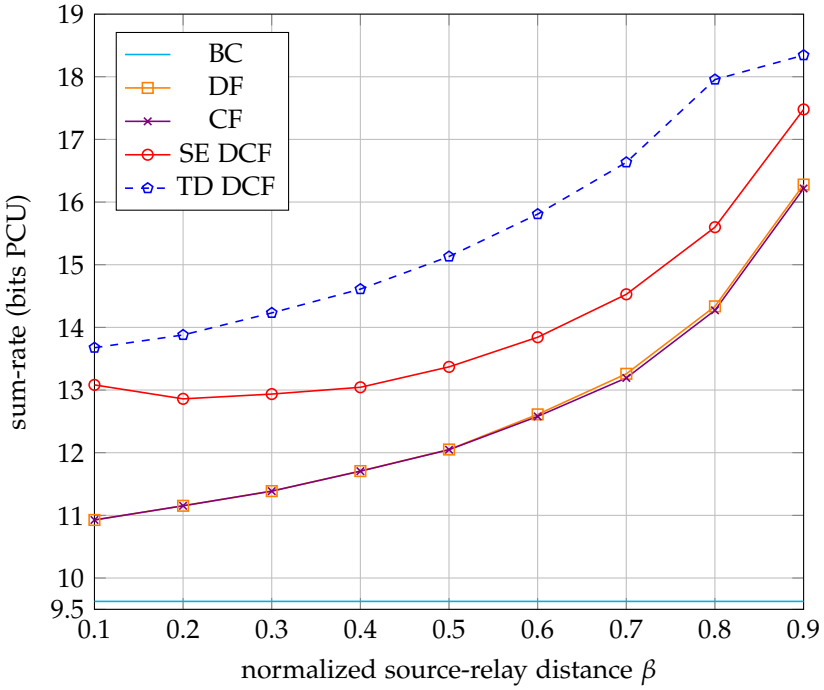

Figure 22: Sum-rate performance of the SE/TD DCF and the baseline schemes (BC/DF/CF) when SNR = 27dB.

downlink multiplexing gain can be exploited.

- Beyond the medium SNR regime, all schemes perform better when the relay

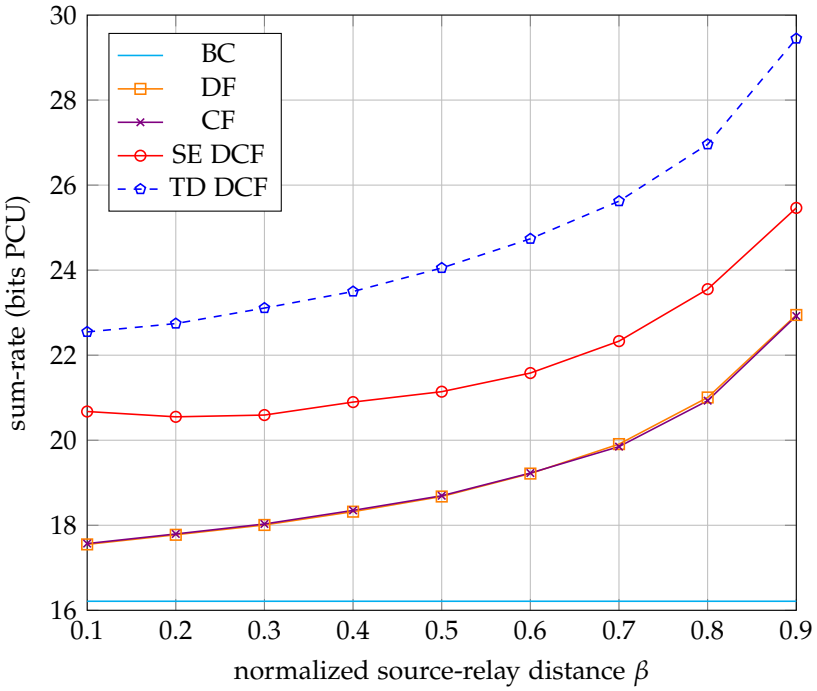

Figure 23: Sum-rate performance of the SE/TD DCF and the baseline schemes (BC/DF/CF) when SNR = 47dB.

gets closer to the destinations, i.e., as β increases. This observation indicates that the bottleneck of the current system lies in the relay-destination link. This is due to the fact that the multiplexing gain provided by the source-relay channel is higher than that of the other channels thanks to multiple antennas at the relay. This gain is not apparent at low-to-medium SNR.

- At low-to-medium SNR (SNR = 7dB), the source-relay link and the other links have similar quality at comparable distances. For both DF and DCF schemes, since the relay has to decode the source messages, there exists an optimal position for the relay between the source and the destinations, where the relay can reliably recover the messages and also help the destinations. This is to be contrasted with the high SNR regime in which the performance is limited by the source-destination links.

- In all regimes the TD DCF outperforms SE DCF and the performance gap increases with the SNR. Intuitively, this is due to the decoding constraint at the relay. Note that both streams are sent simultaneously with SE DCF, which requires the relay to decode both streams after a single transmission. With TD DCF, the relay can decode one stream at a time, which requires a lower source-relay link quality.

- Though it is not shown in the figure, the simulation results suggest that both the variants of the proposed scheme outperform, in terms of the sum-rate performance, the achievable region of broadcast channel with state feedback in [38] even when the transmitter is assigned with a 3dB higher SNR in BC case, as long as the parameters in the proposed schemes' variants,

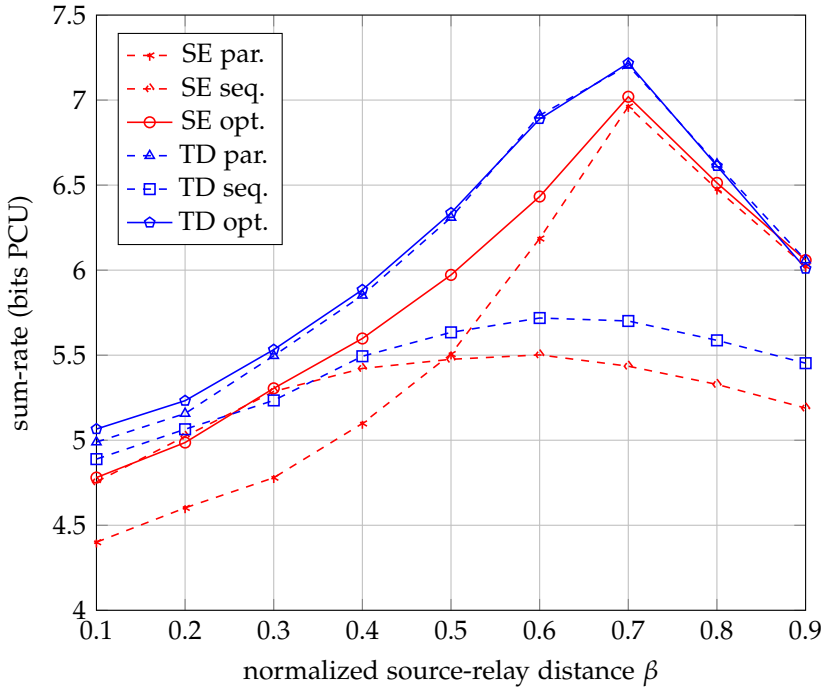

Figure 24: Sum-rate performance of the variants with optimized parameters ("opt.") and the simplified modes: parallel ("par.") and sequential ("seq.") when SNR = 7dB.

i.e., α, β, P_1, P_2, P_{r_0}, P_{r_1}, P_{r_2}, are properly chosen.[4] Therefore, the DCF scheme has a sum-rate gain over the 'convex-hull' of the achievable region mentioned in 5.2.2.

5.4.3 *DCF variants and "parallel/sequential" modes*

From Fig. 24, Fig. 25 and Fig. 26, we focus on variants of DCF schemes and the corresponding modes as follows:

- For both TD and SE DCF, the sequential mode presents a flat $\beta - R$ curve while parallel mode shows a sharp growth as the distance β increases. In fact, in sequential mode, the optimal portion α varies with β. When the relay is close to the source, i.e., β is small, α is large to compensate for a poor relay-destinations channel gain. And α decreases when β increases. Unlike the sequential mode, in the parallel mode, the source and relay transmit concurrently. When β is small, the relay signal is comparable to the source signal at the destinations, which means that the help of the relay comes at the cost of the source signal quality. The overall performance gain from the relay maybe marginal. When β increases, the relay signal becomes stronger than the source signal, the extra information provided by the relay can be

[4]We assume that the relay power is assigned to the transmitter in BC case when comparing the sum-rate in BC case with that of BRC, that is the SNR constraint at the transmitter in BC case is set to 10dB, 30dB, 50dB.

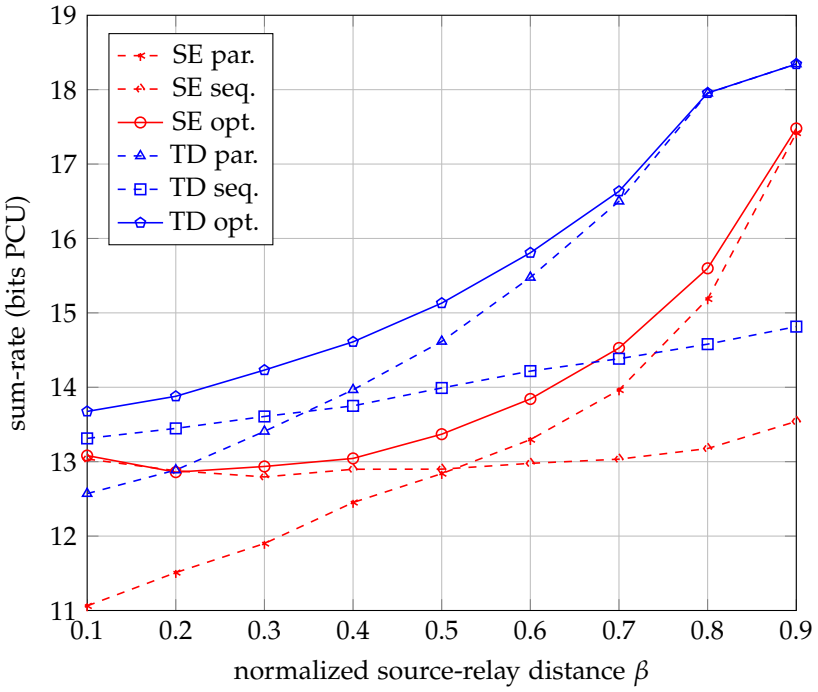

Figure 25: Sum-rate performance of the variants with optimized parameters ("opt.") and the simplified modes: parallel ("par.") and sequential ("seq.") when SNR = 27dB.

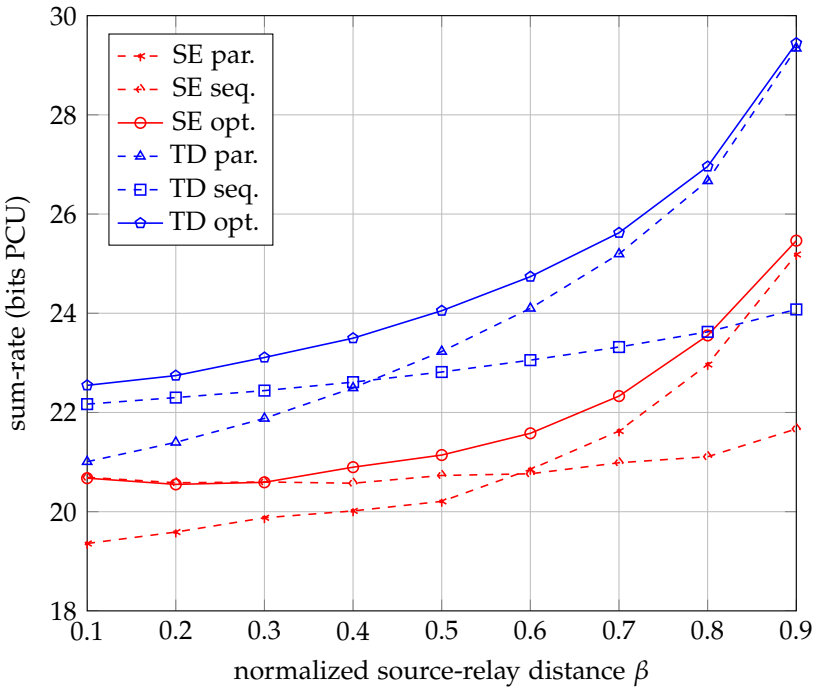

Figure 26: Sum-rate performance of the variants with optimized parameters ("opt.") and the simplified modes: parallel ("par.") and sequential ("seq.") when SNR = 47dB.

obtained by the destinations without hurting the source signal. The overall performance gain becomes larger.

- When the relay is close to the source ($\beta \to 0$), the sequential mode is near optimal, whereas the parallel model is near optimal when the relay is close to the destinations ($\beta \to 0$).[5] For general values of β, neither mode is close to the optimized curve. It implies that there exits a non-trivial value of α to fully exploit the presence of the relay, i.e., a hybrid version of orthogonal and non-orthogonal relaying is needed.
- Although it is not illustrated in the plots, we remark that an asymmetric power allocation ($P_1 \neq P_2$) can attain higher rate than symmetric case for SE DCF and the respective modes. On the one hand, extremely asymmetric power allocations are employed in DF/CF scheme to mitigate the impact of interference and the absence of CSIT. On the other hand, the proposed scheme exploits space-time interference alignment and can cope with certain amount of interference. Therefore, asymmetric power allocation is exploited to alleviate the interference to a low level so that SE DCF can handle it.

5.5 SUMMARY AND CLOSING REMARKS

In this work, we investigated on the downlink transmission of a single-cell multi-user heterogeneous system served by a source (macrocell BS) and a sub-tier relay (smallcell BS), modeled as broadcast relay channel. We first provided a general achievable rate region by the proposed DCF scheme using state feedback from the destinations to the relay. In particular, we focused on two variants of the DCF scheme, namely, the time-division and simultaneously emitted DCF, in a fast fading Gaussian noise BRC and derived the respective achievable regions. Both variant are shown to be optimal in terms of sum DoF.

To validate the theoretical results within finite SNR regime, we set up an one-dimensional deployment model and evaluated the sum-rate performance numerically. We found that our schemes can exploit the state feedback and provide a non-negligible performance gain as compared to the case where feedback is not used. We also revealed that the time-division DCF performs better than the simultaneously emitted DCF thanks to the low source-relay traffic load of the scheme. Remarkably, in both DCF variants, a good design of the scheme parameters depends strongly on the location of the relay.

[5] optimal refers to the variants of the proposed scheme with optimized parameters, i.e., α, P_{r_0}, P_{r_1}, P_{r_2}.

5.A PROOF OF THEOREM 5.3

The schemes (DCF, DF, CF) in this chapter apply the properties of weak typical sequences in the same manner as the one in Chapter 2 and the definition and the analysis of error events are quite standard. Therefore, we avoid to repeat the detailed proof and the readers can refer to Chapter 2 and [74,75] for details.

In this part, we present with details the proof of the achievability (by the DCF scheme) based on block-Markov coding, superposition coding, Wyner-Ziv compression and random binning techniques. The techniques used in the appendices of this chapter are explained in the appendix of Chapter 2 with details, for the sake of space, we provide in the following the essential idea of the achievability proof.

5.A.1 *Codebook generation*

Given a pmf $p(\hat{y}|x_r,u_1,u_2,s,q)p(u_1)p(u_2)p(x|u_1,u_2,q)\ p(x_r|q)p(s)p(q)$ and function $x(u_1,u_2,q)$, the codebooks in each block are generated as follows.

1. Randomly generate 2^{nR_k} independent sequences u_k^n, which is drawn i.i.d. from

$$P_{U_k^n}(u_k^n)=\prod_{i=1}^{n}P_{U_k}(u_{ki}),$$

and index them as $u_k^n(m_k)$ with $m_k\in\{1,2,...,2^{nR_k}\}$, $k=1,2$.

2. Generate randomly a sequence q^n, which is drawn i.i.d. from

$$P_{Q^n}(q^n)=\prod_{i=1}^{n}P_Q(q_i).$$

3. For the generated sequence q^n, randomly and conditionally independently generate 2^{nR_r} independent sequences x_r^n drawn i.i.d from

$$P_{X_r^n|Q^n}(x_r^n|q^n)=\prod_{i=1}^{n}P_{X_r|Q}(x_{ri}|q_i)$$

and index them as $x_r^n(r)$ with $r\in\{1,2,...,2^{nR_r}\}$.

4. For each x_r^n sequence, given the generated sequence q^n and the state sequence s^n feedback from receivers, randomly generate $2^{n\hat{R}}$ conditionally independent sequences $\hat{y}^n$ drawn i.i.d from

$$P_{\hat{Y}^n|X_r^n,S^n,Q^n}(\hat{y}^n|x_r^n,s^n,q^n)=\prod_{i=1}^{n}P_{\hat{Y}|X_r,S,Q}(\hat{y}_i|x_{r,i},s_i,q_i)$$

and index them as $\hat{y}^n(l,r)$ with $l\in\{1,2,...,2^{n\hat{R}}\}$.

$b=1$	$b=2$	$\cdots$	$b=B-1$	$b=B$	$b=B+1$	$\cdots$	$b=B+L$
$u_k^n(m_{k1})$	$u_k^n(m_{k2})$	$\cdots$	$u_k^n(m_{k(B-1)})$	$u_k^n(m_{kB})$	$u_k^n(1)$	$\cdots$	$u_k^n(1)$
$x_r^n(1)$	$x_r^n(r_1)$	$\cdots$	$x_r^n(r_{B-2})$	$x_r^n(r_{B-1})$	$x_r^n(r_B)$	$\cdots$	$x_r^n(r_B)$
$\hat{y}^n(l_1,1)$	$\hat{y}^n(l_2,r_1)$	$\cdots$	$\hat{y}^n(l_{B-1},r_{B-2})$	$\hat{y}^n(l_B,r_{B-1})$	$\hat{y}^n(l_B,r_B)$	$\cdots$	$\hat{y}^n(l_B,r_B)$

Table 4: Encoding of the DCF scheme

5. Partition the sets $\{l|l \in [1 : 2^{n\hat{R}}]\}$ into 2^{nR_r} equal sized bins $\mathcal{B}(r)$ with $2^{n(\hat{R}-R_r)}$ elements in each bin, while $r \in [1 : 2^{nR_r}]$ denotes the bin index.

6. Generate the x^n sequence as each element can be obtained from the function $x(u_1, u_2, q)$.

7. Provide the codebooks and bins to the communication nodes.

5.A.2 *Encoding*

The transmission is performed in $B+L$ blocks, where B blocks are dedicated to convey the messages while the L blocks assure the successful decoding of last index of auxiliary variable w.p. 1 ($L \to \infty$, $\frac{L}{B} \to 0$). See Table 4 for encoding details.

Next, we focus on the source and the relay encoding as below.

- At each block $b \in [1 : B]$, the transmitter communicates to each destination the messages (m_{1b}, m_{2b}) based on $u_1^n(m_{1b}), u_2^n(m_{2b})$ and time sharing random variable q^n. To be specific, encoder randomly picks and then sends $x^n(b)$ as function of $(u_1^n(m_{1b}), u_2^n(m_{2b}), q^n(b))$. For $b \in [B+1 : B+L]$, the source transmits dummy messages with $m_{1b} = m_{2b} = 1$.

- At the end of each block $b = [1 : B]$, we assume that the relay receives the state sequence $s^n(b)$ via causal feedback from the destinations, and perfectly recovers the source messages (m_{1b}, m_{2b}). The constraints for the successful decoding of (m_{1b}, m_{2b}) are presented in Section 5.A.3. Then, the relay generates the codebook of $\hat{y}^n$ using the state sequence $s^n(b)$ as described at the beginning of the section, and searches for at least one index l_b such that

  $$\big(u_1^n(m_{1b}), u_2^n(m_{2b}), \hat{y}^n(l_b, r_{b-1}), x_r^n(r_{b-1}), s^n(b), q^n(b)\big) \in \mathcal{T}_\delta^n\big(U_1, U_2, \hat{Y}, X_r, S, Q\big).$$

  The probability that such l_b exists goes to one as n tends to infinity if

  $$\hat{R} \geq I(\hat{Y}; U_1, U_2 | X_r, S, Q). \tag{135}$$

  The relay obtains the bin index r_b such that $l_b \in \mathcal{B}(r_b)$, which is sent at block $b+1$ in the form of $x_r^n(r_b)$. By convention, $r_0 = 1$. For $b = [B+1 : B+L-1]$,

at the block $b+1$, the relay repeats the last description of compression, namely r_B. For $b = [B : B+L]$, $l_b = l_B$ (see Table 4) assures that the joint typicality in relay's encoding always holds.

5.A.3 *Decoding and error events analysis*

We assume that the sequence q^n in each block is known to the transmitter, the relay, and both the receivers due to the fact that q^n are chosen from the codebook according to the channel parameters/statistics and the transmitting protocol, where both are known to the communication nodes.

Then, the decoding strategies at the relay and the destinations are shown as below, followed by respective error events analysis.

1) At the end of each block $b = [1 : B]$, we assume that the relay receives the state sequence $s^n(b)$ via causal feedback from the destinations. The relay decoder searches for the unique (m_{1b}, m_{2b}) such that

$$\left(u_1^n(m_{1b}), u_2^n(m_{2b}), x_r^n(r_{b-1}), y_r^n(b), s^n(b), q^n(b)\right) \in \mathcal{T}_\delta^n(U_1, U_2, X_r, Y_r, S, Q).$$

Thus, (m_{1b}, m_{2b}) can be decoded with error probability goes to zero (as n goes to infinity) provided that,

$$R_1 \leq I(U_1; Y_r | X_r U_2 S Q), \tag{136a}$$
$$R_2 \leq I(U_2; Y_r | X_r U_1 S Q), \tag{136b}$$
$$R_1 + R_2 \leq I(U_1 U_2; Y_r | X_r S Q). \tag{136c}$$

2) At every destination (let us take destination 1 as an example), r_B is decoded at the end of block $B+L$ such that

$$(x_r^n(r_B), y_1^n(b), s^n(b), q^n(b)) \in \mathcal{T}_\delta^n(X_r, Y_1, S, Q)$$

and the error probability goes to zero when n tends to infinity, $b = [B+1 : B+L]$ provided that,

$$R_r \leq L \cdot I(X_r; Y_1 | S Q), \tag{137}$$

which holds with high probability as L goes to infinity ($L = O(B)$).

Hereafter we discuss two strategies for the rest of the decoding according to the relation between I_{51} and zero. Quite some resources need to be allocated for decoding the relay message, under the condition $I_{51} < 0$ ($I_{41} + I_{51} < I_{41}$), so that decoding the source message with relay's information will be even more costly than without the side information from relay. For instance, the system might not benefit much from activating the relay when the RDC link has a really low SNRs in the Gaussian case. Thus, it is reasonable to shut the relay down and return to the BC case. Hence whenever $I_{51} < 0$, the destination 1 simply ignores the existence of the relay and decode directly the source

message such that

$$(u_1^n(\hat{m}_{1b}), y_1^n(b), s^n(b), q^n(b)) \in \mathcal{T}_\delta^n(U_1, Y_1, S, Q)$$

are jointly typical. To ensure the correctness of recovered message, its rate should fulfill $R_1 \leq I(U_1; Y_1|SQ) = I_{41}$ which, together with the previous encoding/decoding constraints, formulates a rate region. As long as this region is achievable, a reduced region limited by $R_1 \leq \min\{I_{21}, I_{31}, I_{41} + I_{51}\}$ and $I_{51} < 0$ is achievable since $\min\{I_{21}, I_{31}, I_{41} + I_{51}\} \leq I_{41} + I_{51} \leq I_{41}$.

In the sequel, we present the decoding strategy for $I_{51} \geq 0$ and prove that

$$R_1 \leq \min\{I_{21}, I_{31}, I_{41} + I_{51}\}$$

is as well achievable in this case, which yields a consistent rate region regardless of the relation between I_{51} and zero.

With the assumption that r_b is correctly recovered, a *joint* decoding is performed in a *backward* manner, i.e., for b from B to 1, we find jointly $(\hat{r}_{b-1}, \hat{l}_b, \hat{m}_{1b})$ such that

$$\big(u_1^n(\hat{m}_{1b}), \hat{y}^n(\hat{l}_b, \hat{r}_{b-1}), x_r^n(\hat{r}_{b-1}), y_1^n(b), s^n(b), q^n(b)\big) \in \mathcal{T}_\delta^n(U_1 \hat{Y} X_r Y_1 S Q).$$

The entire decoding procedure is designed backwardly, i.e., the indices $(\hat{r}_{b-1}, \hat{l}_b, \hat{m}_{1b})$ are decoded *a priori*, before the recovery of the indices $(\hat{r}_{b-2}, \hat{l}_{b-1}, \hat{m}_{1(b-1)})$ for $b \in \{2, 3, \cdots, B\}$.

The error events are analyzed as below.

i. The probability of error event $(\hat{r}_{b-1} = r_{b-1}, \hat{l}_b = l_b, \hat{m}_{1b} \neq m_{1b})$ tends to zero if

$$R_1 \leq I(U_1; Y_1 \hat{Y}|X_r SQ). \tag{138}$$

ii. The error event such that $(\hat{r}_{b-1} \neq r_{b-1}, \hat{l}_b = l_b, \hat{m}_{1b} = m_{1b})$ could be avoided if

$$\hat{R} \leq I(X_r \hat{Y}; U_1 Y_1|SQ). \tag{139}$$

Note that the backward decoding is performed given a perfectly recovered r index from last decoding. As long as an error occurs on r_{b-1}, such prerequisite does not hold any more and thus, due to the wrong bin index r_{b-1}, an incorrect $\hat{y}^n(\hat{l}_{b-1}, \hat{r}_{b-2})$ sequence is called in next decoding. Therefore, even though l_b is correct, we compute (139), (141) as if l is wrong.

iii. The probability of error event $(\hat{r}_{b-1} = r_{b-1}, \hat{l}_b \neq l_b, \hat{m}_{1b} = m_{1b})$ goes to zero if

$$\hat{R} - R_r \leq I(\hat{Y}; Y_1 U_1|X_r SQ). \tag{140}$$

iv. The probability of error event $(\hat{r}_{b-1} \neq r_{b-1}, \hat{l}_b = l_b, \hat{m}_{1b} \neq m_{1b})$ tends to

zero if

$$R_1 + \hat{R} \leq I(U_1 X_r; Y_1|SQ) + I(\hat{Y}; Y_1 U_1|X_r SQ). \tag{141}$$

v. The probability of error event ($\hat{r}_{b-1} = r_{b-1}, \hat{l}_b \neq l_b$, $\hat{m}_{1b} \neq m_{1b}$) tends to zero if

$$R_1 + \hat{R} - R_r \leq I(U_1; Y_1|X_r SQ) + I(\hat{Y}; Y_1 U_1|X_r SQ). \tag{142}$$

vi. The probability of error event ($\hat{r}_{b-1} \neq r_{b-1}, \hat{l}_b \neq l_b$, $\hat{m}_{1b} = m_{1b}$) tends to zero if

$$\hat{R} \leq I(X_r \hat{Y}; Y_1 U_1|SQ). \tag{143}$$

vii. The probability of error event ($\hat{r}_{b-1} \neq r_{b-1}, \hat{l}_b \neq l_b$, $\hat{m}_{1b} \neq m_{1b}$) tends to zero if

$$R_1 + \hat{R} \leq I(X_r U_1; Y_1|SQ) + I(\hat{Y}; Y_1 U_1|X_r SQ). \tag{144}$$

3) Finally, with the same reasoning, we can obtain symmetric constraints for user 2 by exchanging indices in the constraints listed above. Moreover, the constraints (135)-(136c) and (138)-(144) are simplified by applying the FME to R_r and $\hat{R}$ successively. In the end, the redundant constraints are removed.

5.B PROOF OF PROPOSITION 5.1

In this part, we present with details the proof of the achievability (by the extended version of DF scheme for relay-aided channel) based on the block-Markov coding, superposition coding techniques.

Codebook generation:

Fix the pmf $p(x_1)p(x_2)p(u_1|x_1)p(u_2|x_2)p(x|u_1,u_2,q)p(x_r|x_1,x_2,q)p(s)p(q)$ and function $x(u_1,u_2,q)$ that attain the rate region in Proposition 5.1. Then, the codebooks in each block are generated as follows.

1. Randomly and conditionally independently generate 2^{nR_k} independent sequences x_k^n drawn i.i.d from

$$P_{X_k^n}(x_k^n) = \prod_{i=1}^{n} P_{X_k}(x_{ki})$$

and index them as $x_k^n(m_k)$ with $m_k \in \{1, 2, ..., 2^{nR_k}\}$.

2. For each m_k, randomly generate 2^{nR_k} independent sequences u_k^n, which is

$b=1$	$b=2$	$\cdots$	$b=B-1$	$b=B$	$b=B+1$	$\cdots$	$b=B+L$
$u_k^n(1,m_{k1})$	$u_k^n(m_{k1},m_{k2})$	$\cdots$	$u_k^n(m_{k(B-2)},m_{k(B-1)})$	$u_k^n(m_{k(B-1)},m_{kB})$	$u_k^n(m_{kB},m_{kB})$	$\cdots$	$u_k^n(m_{kB},m_{kB})$
$x_k^n(1)$	$x_k^n(m_{k1})$	$\cdots$	$x_k^n(m_{k(B-2)})$	$x_k^n(m_{k(B-1)})$	$x_k^n(m_{kB})$	$\cdots$	$x_k^n(m_{kB})$

Table 5: Encoding of the DF scheme

drawn i.i.d. from

$$P_{U_k^n}(u_k^n)=\prod_{i=1}^{n}P_{U_k}(u_{ki}),$$

and index them as $u_k^n(m_k,m_k')$ with $m_k'\in\{1,2,...,2^{nR_k}\}$, $k=1,2$.

3. Generate randomly a sequence q^n, which is drawn i.i.d. from

$$P_{Q^n}(q^n)=\prod_{i=1}^{n}P_Q(q_i).$$

4. Generate the x^n sequence as each element can be obtained from the function $x(u_1,u_2,q)$.

5. Provide the codebooks and bins to the communication nodes.

Encoding:

The transmission is performed in $B+L$ blocks, where B blocks are dedicated to convey the messages while the L blocks assure the successful decoding of last index of m_{kB} w.p. 1 ($L\to\infty$, $\frac{L}{B}\to 0$). See Table 5 for encoding details.

Next, we focus on the source and the relay encoding as below.

- At each block $b\in[1:B]$, the transmitter communicates to each destination the messages (m_{1b},m_{2b}) based on the generated codewords $u_1^n(m_{1(b-1)},m_{1b})$, $u_2^n(m_{2(b-1)},m_{2b})$ and time sharing random variable q^n. We set $m_{k0}=1$ by convention. To be specific, encoder randomly picks and then sends $x^n(b)$ as function of $(u_1^n(m_{1b}),u_2^n(m_{2b}),q^n(b))$.For $b\in[B+1:B+L]$, the source transmits the messages $m_{kb}=m_{kB}$.

- The relay's encoding is to some extent linked with relay's decoding in DF scheme. At the end of each block $b=[1:B]$, the relay searches for at least one index m_{kb} such that the following joint typicality holds.

$$\Big(u_1^n(m_{1(b-1)},m_{1b}),u_2^n(m_{2(b-1)},m_{2b}),x_1^n(m_{1(b-1)}),x_2^n(m_{2(b-1)}),$$
$$y_r^n(b),s^n(b),q^n(b)\Big)\in\mathcal{T}_\delta^n\big(U_1,U_2,X_1,X_2,Y_r,S,Q\big).$$

The probability that such m_{kb} exists goes to one as n tends to infinity if the

sequel conditions are fulfilled.

$$R_1 \leq I(U_1; Y_r|X_1X_2U_2SQ), \tag{145a}$$
$$R_2 \leq I(U_2; Y_r|X_1X_2U_1SQ), \tag{145b}$$
$$R_1 + R_2 \leq I(U_1U_2; Y_r|X_1X_2SQ). \tag{145c}$$

Through this search, the relay obtains the source messages m_{kb} which is encoded in the form of $x_k^n(m_{kb})$. For $b = [B+1 : B+L-1]$, at the block $b+1$, the relay repeats the message m_{kB}. Then, relay encoder randomly picks and sends $x_r^n(b)$ such that

$$x_r^n(b) \in \mathcal{T}_\delta^n\big(X_r^n(b)|X_1^n(m_{1b}), X_2^n(m_{2b}), Q^n(b)\big).$$

The relay repeats $m_{kb} = m_{kB}$ for $b \in [B+1 : B+L]$.

Decoding and error events analysis:

The decoding strategies at the destinations are shown below, followed by respective error events analysis.

1) At every destination (let us take destination 1 as an example), m_{1B} is decoded at the end of block $B+L$ such that

$$(x_1^n(m_{1B}), u_1^n(m_{1B}, m_{1B}), y_1^n(b), s^n(b), q^n(b)) \in \mathcal{T}_\delta^n(X_1, U_1, Y_1, S, Q)$$

and the error probability goes to zero when n tends to infinity, $b = [B+1 : B+L]$ provided that,

$$R_1 \leq L \cdot I(X_1U_1; Y_1|SQ), \tag{146}$$

which holds with high probability as L goes to infinity ($L = O(B)$).

2) With a correctly decoded $m_{1(b+1)}$, the receiver 1 performs a backward decoding to search for index m_{1b} such that

$$(x_1^n(m_{1b}), u_1^n(m_{1b}, m_{1(b+1)}), y_1^n(b+1), s^n(b+1), q^n(b+1)) \in \mathcal{T}_\delta^n(X_1, U_1, Y_1, S, Q)$$

for $b = [1 : B-1]$ and the error probability goes to zero when n tends to infinity, provided that,

$$R_1 \leq I(X_1U_1; Y_1|SQ), \tag{147}$$

3) Finally, with the same reasoning, we can obtain symmetric constraints for user 2 by exchanging indices in the constraints listed above. In the end, the constraints in Proposition 5.1 are obtained.

5.C PROOF OF PROPOSITION 5.2

In this part, we present with details the proof of the achievability (by the CF scheme) based on block-Markov coding, Wyner-Ziv compression and random binning techniques.

Codebook generation:

Fix the pmf $p(\hat{y}|x_r, y_r, q)p(u_1)p(u_2)p(x|u_1, u_2, q)p(x_r|q)p(s)p(q)$ and function $x(u_1, u_2, q)$ that attain the rate region in Proposition 5.2. Then, the codebooks in each block are generated as follows.

1. Randomly generate 2^{nR_k} independent sequences u_k^n, which is drawn i.i.d. from

$$P_{U_k^n}(u_k^n) = \prod_{i=1}^{n} P_{U_k}(u_{ki}),$$

   and index them as $u_k^n(m_k)$ with $m_k \in \{1, 2, ..., 2^{nR_k}\}$, $k = 1, 2$.

2. Generate randomly a sequence q^n, which is drawn i.i.d. from

$$P_{Q^n}(q^n) = \prod_{i=1}^{n} P_Q(q_i).$$

3. For the generated sequence q^n, randomly and conditionally independently generate 2^{nR_r} independent sequences x_r^n drawn i.i.d from

$$P_{X_r^n|Q^n}(x_r^n|q^n) = \prod_{i=1}^{n} P_{X_r|Q}(x_{ri}|q_i)$$

   and index them as $x_r^n(r)$ with $r \in \{1, 2, ..., 2^{nR_r}\}$.

4. For each sequence pair x_r^n and q^n conceived, randomly generate $2^{n\hat{R}}$ conditionally independent sequences $\hat{y}^n$ drawn i.i.d from

$$P_{\hat{Y}^n|X_r^n,Q^n}(\hat{y}^n|x_r^n, q^n) = \prod_{i=1}^{n} P_{\hat{Y}|X_r,Q}(\hat{y}_i|x_{r,i}, q_i)$$

   and index them as $\hat{y}^n(l, r)$ with $l \in \{1, 2, ..., 2^{n\hat{R}}\}$.

5. Generate the x^n sequence as each element can be obtained from the function $x(u_1, u_2, q)$.

6. Partition the sets $\{l|l \in [1 : 2^{n\hat{R}}]\}$ into 2^{nR_r} equal sized bins $\mathcal{B}(r)$ with $2^{n(\hat{R}-R_r)}$ elements in each bin, while $r \in [1 : 2^{nR_r}]$ denotes the bin index.

7. Provide the codebooks and bins to the communication nodes.

$b=1$	$b=2$	$\cdots$	$b=B-1$	$b=B$	$b=B+1$	$\cdots$	$b=B+L$
$u_k^n(m_{k1})$	$u_k^n(m_{k2})$	$\cdots$	$u_k^n(m_{k(B-1)})$	$u_k^n(m_{kB})$	$u_k^n(1)$	$\cdots$	$u_k^n(1)$
$x_r^n(1)$	$x_r^n(r_1)$	$\cdots$	$x_r^n(r_{B-2})$	$x_r^n(r_{B-1})$	$x_r^n(r_B)$	$\cdots$	$x_r^n(r_B)$
$\hat{y}^n(l_1,1)$	$\hat{y}^n(l_2,r_1)$	$\cdots$	$\hat{y}^n(l_{B-1},r_{B-2})$	$\hat{y}^n(l_B,r_{B-1})$	$\hat{y}^n(l_B,r_B)$	$\cdots$	$\hat{y}^n(l_B,r_B)$

Table 6: Encoding of the CF scheme

Encoding:

The transmission is performed in $B+L$ blocks, where B blocks are dedicated to convey the messages while the L blocks assure the successful decoding of last index of auxiliary variable w.p. 1 ($L\to\infty$, $\frac{L}{B}\to 0$). See Table 6 for encoding details.

Next, we focus on the source and the relay encoding as below.

- At each block $b\in[1:B]$, the transmitter communicates to each destination the messages (m_{1b},m_{2b}) based on $u_1^n(m_{1b}),u_2^n(m_{2b})$ and time sharing random variable q^n. To be specific, encoder randomly picks and then sends $x^n(b)$ as function of $(u_1^n(m_{1b}),u_2^n(m_{2b}),q^n(b))$.For $b\in[B+1:B+L]$, the source transmits dummy messages with $m_{1b}=m_{2b}=1$.

- After the transmission of block $b=[1:B]$, the relay searches for at least one index l_b such that

$$\big(y_r^n(b),\hat{y}^n(l_b,r_{b-1}),x_r^n(r_{b-1}),s^n(b),q^n(b)\big)\in\mathcal{T}_\delta^n\big(Y_r,\hat{Y},X_r,S,Q\big).$$

  The probability that such l_b exists goes to one as n tends to infinity if

$$\hat{R}\geq I(\hat{Y};Y_r|X_r,S,Q). \tag{148}$$

  The relay obtains the bin index r_b such that $l_b\in\mathcal{B}(r_b)$, which is sent at block $b+1$ in the form of $x_r^n(r_b)$. By convention, $r_0=1$. For $b=[B+1:B+L-1]$, at the block $b+1$, the relay repeats the last description of compression, namely r_B.

Decoding and error events analysis:

The decoding strategies at the destinations are shown below.

1) At every destination (take destination 1 for example), r_B is decoded at the end of block $B+L$ such that $(x_r^n(r_B),y_1^n(b),s^n(b),q^n(b))\in\mathcal{T}_\delta^n(X_r,Y_1,S,Q)$ and the error probability goes to zero when n tends to infinity, $b=[B+1:B+L]$

provided that,

$$R_r \leq L \cdot I(X_r; Y_1|SQ), \tag{149}$$

which holds with high probability as L goes to infinity ($L = O(B)$).

2) As it is explained in Appendix 5.A, we only present the decoding strategy with a prerequisite $I(X_r; Y_k|U_kSQ) - I(Y_r; \hat{Y}|X_rU_kY_kSQ) \geq 0$ and claim that a consistent rate region is achievable no matter whether $I(X_r; Y_k|U_kSQ) - I(Y_r; \hat{Y}|X_rU_kY_kSQ)$ is positive or not.

   With the assumption that r_b is correctly recovered, a backward and joint decoding is performed, i.e., for b from B to 1, we find jointly $(\hat{r}_{b-1}, \hat{l}_b, \hat{m}_{1b})$ such that

$$\left(u_1^n(\hat{m}_{1b}), x_r^n(\hat{r}_{b-1}), \hat{y}^n(\hat{l}_b, \hat{r}_{b-1}), y_1^n(b), s^n(b), q^n(b)\right) \in \mathcal{T}_\delta^n(U_1X_r\hat{Y}Y_1SQ).$$

   The error probabilities all tend to zero if

$$R_1 \leq I(U_1; Y_1\hat{Y}|X_rSQ). \tag{150a}$$
$$\hat{R} \leq I(X_r\hat{Y}; U_1Y_1|SQ). \tag{150b}$$
$$\hat{R} - R_r \leq I(\hat{Y}; Y_1U_1|X_rSQ). \tag{150c}$$
$$R_1 + \hat{R} \leq I(U_1X_r; Y_1|SQ) + I(\hat{Y}; Y_1U_1|X_rSQ). \tag{150d}$$
$$R_1 + \hat{R} - R_r \leq I(U_1; Y_1|X_rSQ) + I(\hat{Y}; Y_1U_1|X_rSQ). \tag{150e}$$

3) Finally, we exchange the subscripts to obtain symmetric constraints for user 2. In the end, the constraints in Proposition 5.2 are achieved with the help of FME.

Part III

Conclusions

6 Conclusions and Perspectives

In this chapter, we first revisit some results, summarize the contributions exhibited in the previous chapters and then present some perspectives about possible future work.

6.1 CONCLUSIONS AND SOME COMMENTS

In this thesis, we investigated one of the popular approaches of feedback enabled interference mitigation for downlink communication problem, that is, RIA allowed by state feedback. In particular, in Part I of the thesis we have focused on the homogeneous outdated state feedback while in Part II of the thesis we have explored the heterogeneous delayed CSIT in two basic models: the broadcast channel and the broadcast relay channel. The study was conducted with information theoretical tools.

6.1.1 *Broadcasting with homogeneous state feedback*

The first part of the thesis focused on the homogeneous delayed CSIT in multiple user broadcast channel. Though both chapters in this part dealt with the similar problem, they have distinct emphasis, that is, inner and outer bounds in two user case (Chapter 2) and relatively simple inner bounds in K user channels (Chapter 3).

In two user model, we attempted to characterize the gain of state feedback at finite SNR as its outstanding benefit with infinite SNR has long been proven in [2, 12, 13, 19] and references therein. In the effort to assess it, we derived an inner bound by using the prominent tools of block Markov coding and Wyner-Ziv compression [61, 73, 74]. The transmission strategy hinges on leveraging the previously encountered overheard information, i.e., rebuild and encode the misplaced observations and then multicast a quantized version. Moreover, a standard outer bound is also provided as a means of comparison and then tighten in Gaussian case by an extremal inequality and some algebraic manipulations. Later on, it is shown that numerically optimized symmetric rate outperforms conventional strategy. Nevertheless, even with the numerical optimizer, it is not efficient to pick out a gainful set of the parameters $(P_k,\ N)$ in the proposed region, which spurs the work of Chapter 3 as a partial solution.

In K user model, we not only aim at finding a coding method to endorse the finite SNR gain in K user Gaussian channel, but also wish to get rid of the

input alphabet size constraint in EBC. With this in mind, we excluded the block Markov and binning techniques and introduced a distributed lossy coding and thus the scheme presented in Chapter 3 is sketched as being general and simple. Consequently, the tractable region we obtained was shown to achieve a superior performance than TDMA and MAT-related schemes for beyond two-user GBC and to attain capacity with non-restricted input for certain subclasses of EBC, simultaneously. As stated previously, thanks to the time division structure, we convert the power allocation optimization problem with P's inside the expectation to a slot distributing issue solvable by analysis, which is followed by optimizing compression noise variance via simulation. To sum up, the effectiveness of delayed CSIT at finite SNR is justified in these two chapters.

6.1.2 *Broadcasting with heterogeneous state feedback*

In this part of the thesis, we concentrate on the state feedback subjects to heterogeneity.

State feedback from partial receiver(s)

In Chapter 4, we considered for the most part the scenario where only one receiver feedback the state. In this model, we tailored the schemes prepared in Chapter 2, 3 and derived the corresponding regions. It is worth re-emphasizing that our adapted schemes, as opposed to some linear methods (e.g., MAT [2], packet combining [27,28,43]) in the literature that exploit obsolete state, pivot on non-linear coding by themselves. Though it is unclear and also not trivial to determine whether non-linear scheme can have a superior performance to the existing linear methods in general, both analytical evaluation and numerical examples provided a positive answer in two user EBC. Moreover, we revealed where such advantage lies in. Specifically, our adapted schemes provide more decodable bits that are useful at the no-feedback user than the XOR based linear codes, and therefore allocate more flows to the no-feedback user. Quite remarkably, one constraint of the regions by non-linear coding and that of outer bound collapse, which seems to imply the looseness of a symmetric outer bound. In addition, we specified one region to three user case and demonstrated that incorporating a no-feedback user into two-user EBC is indeed beneficial for certain class of EBC.

State feedback to relay in a BRC

In Chapter 5, we studied a particular model where an infrastructure relay is added into the BC. In this channel, a more realistic restriction is imposed on the feedback link that outdated state is provided to the relay only. We first obtain in the general two user BRC an inner bound derived from the Decode-Compress-and-Forward scheme. Both the decoding and compressing operations in the conventional relaying strategies (DF, CF) are merged in DCF scheme such that the relay decodes all the messages, encodes them with state, and then compresses the encoded signals to make them suitable for multicasting. Then in the Gaussian fading BRC,

depending on how the source message is conveyed and relayed, the inner bound is specified to two variants, namely, TD-DCF and SE-DCF. After verifying that both variants are DoF-optimal, we further evaluated the corresponding rate regions at finite SNR. The numerical results confirm the performance gain from the use of relay in the presence of the heterogeneous state feedback (from the destinations to the relay). The improvement is actually in two aspects: 1) a remarkable gain over no-feedback rate in BRC is observed when compared with the adapted DF and CF scheme; 2) an enhancement against the rate with homogeneous delayed CSIT of BC (no relay) is proved implicitly by comparing with the rates in Chapter 2, 3. It is also revealed that TD-DCF leads to a better performance than SE-DCF in the investigated setting.

In the traditional RC, either CF or DF schemes is chosen according to the links' qualities. For instance, the CF strategy outperforms in general the DF scheme when the SRC link is weaker than SDC or the receiver has a sufficiently strong connection to the relay. Provided that, mixing the two strategies is often considered ineffective in that the rate constraint on the SRC released by applying compression on the output at relay will remain if we impose the relay to decode the message. Nevertheless, decoding and then compressing can be indeed advantageous in multiple user relay networks. From the perspective of combating interference, decoding the message is almost a prerequisite for the reconstruction of interference (overheard outputs) in state-dependent channel while compressing the encoded signal accommodates side information to multicasting.

In spite of the distinction on how the transmission strategies are performed, one may not ignore the resemblance between the RVs choices of some of the schemes proposed in this thesis. In particular, the superposition-structured block Markov scheme resembles the parallel mode of DCF coding while the Distributed Compression method reminds us of the sequential mode of DCF coding in BRC. More specifically, both the parallel mode of SE-DCF and the superposition-structured block Markov scheme let the transmitter(s) send out signals related to private messages and common auxiliary message simultaneously. Similarly, sequential mode of TD-DCF shares exactly the same choice on RVs with DC scheme if we substitute $V_{1\to\{12\}}$ for X_r as both allocate different time slots for messages with different object(s) such that the multicasting and unicasting will not interfere each other. However, the preferable RVs are not identical. For two user BC, time-sharing seems to be more favorable while the best mode in BRC varies with the channel condition, e.g., the distance between each communication nodes.

6.2 PERSPECTIVES ON FUTURE WORK

Owing to the limitation upon the duration of a Ph.D. program, several possible topics have not been investigated. Among them, we allude in the sequel some potentially compelling directions for future work.

Assuming temporally i.i.d. channel leads to tractable regions, Nonetheless, with respect to practice, the real system might not be captured/characterized by a channel model where the receivers are always subject to the identical channel

process. To that end, encompassing statistical state information will be helpful to an uncertain extent. The easiest way to test its efficiency is to put it in the BC and make it work with state feedback. As stated in Chapter 2, the gap between inner and outer bounds seems to be large due to looseness of outer bound. Hence, another interesting direction for future work could be deriving a tighter outer bound and characterizing the respective gap.

We remind the readers that the capacity of two-user EBC with heterogeneous feedback is not fully determined possibly due to the looseness of outer bound used. The classic tool for outer bound in general BC is the genie-aided technique along with the fact that feedback does not enlarge capacity of degraded channel. This tool relies essentially on the degradedness of the channel, which is not the case when heterogeneous delayed CSIT is studied. The fact that the current input is correlated to previous states and inputs renders it quite involved to find the maximal pmf and optimal encoder and prevents us from obtaining a single letter expression of outer bound. Thus, finding a powerful tool for respective outer bound accounts for an important direction of our future investigation.

Two possible extensions of Chapter 5 will be considering the impact of imperfect feedback by introducing uplink noise and the influence of feeding back state of single user to the relay. One of the main focus of this thesis is how effective the state feedback is in different scenarios where system resources (power, feedback availability, etc.) are limited. Analogous to the case investigated in the manuscript, the state contaminated by feedback noise is another restriction possibly occurred in real system. Therefore, jointly consider it in BC, BRC would take us a step closer to the implementation. Since the usefulness of non-linear coding tools have already been demonstrated in Chapter 4, combining it with BRC appears to be profitable.

Incorporating the feedback-aided BRC with cache deployed at the relay and receivers sounds to be a reasonable direction to go. As proved in [82], caching is compatible with the state feedback enabled interference mitigation. However, the user mobiles may not be equipped with large size cache due to the physical limitation of the portable devices. In this regard, deploying the infrastructure relay with cache can serve as a helper that alleviates interference (Chapter 5) and as an alternative/backup cache. In fact, the cache at relay will also relieve part of relay's burden on decoding both messages from the source.

A

Weak Typical Sequences

Following [75], we use *weak(ly) typical sets* in this thesis. Some useful facts are recalled here. Let X and Y be RVs on some finite or countably infinite sets $\mathcal{X}$ and $\mathcal{Y}$, respectively. We denote the joint probability distribution of (X,Y) (resp. conditional distribution of Y given X, and marginal distribution of X) by $p_{X,Y}$ (resp. $p_{Y|X}$, and p_X).

Definition A.1 (Typical sequence and typical set)**.** *A sequence $x^n \in \mathcal{X}^n$ with pmf $p_{X^n}(x^n) = \prod_{i=1}^n p_X(x_i)$ is called (weakly) δ-typical with respect to X if there always exists an $\delta > 0$ such that*

$$|-\frac{1}{n}\log_2(p_{X^n}(x^n)) - H(X)| \leq \delta$$

The set of all such sequences is called (weakly) typical set and is denoted by $\mathcal{T}_\delta^n(X)$. The typical set (resp. sequence) is always referred to weakly typical set (resp. sequence) in this thesis unless it is clearly stated.

Definition A.2 (Joint typical sequence and typical set)**.** *A sequence $(x^n, y^n) \in (\mathcal{X} \times \mathcal{Y})^n$ with pmf $p_{X^n,Y^n}(x^n, y^n) = \prod_{i=1}^n p_{X,Y}(x_i, y_i)$ is called joint (weakly) δ-typical with respect to (X, Y) if there always exists an $\delta > 0$ such that*

$$\begin{aligned} |-\frac{1}{n}\log_2(p_{X^n}(x^n)) - H(X)| &\leq \delta, \\ |-\frac{1}{n}\log_2(p_{Y^n}(y^n)) - H(Y)| &\leq \delta, \\ |-\frac{1}{n}\log_2(p_{X^n,Y^n}(x^n, y^n)) - H(X,Y)| &\leq \delta. \end{aligned}$$

The set of all such sequences is called joint (weakly) typical set and is denoted by $\mathcal{T}_\delta^n(X,Y)$.

Lemma A.1 ([75, Theorem 3.1.2])**.** *There exists a sequence $\epsilon_n \to 0$ as $n \to \infty$ such that,*

$$Pr(x^n \in \mathcal{T}_\delta^n(X)) \geq 1 - \epsilon_n.$$

Similarly, there exists a sequence $\epsilon_n' \to 0$ as $n \to \infty$ such that, for the sequence (x^n, y^n) drawn i.i.d. according to $p_{X^n,Y^n}(x^n, y^n) = \prod_{i=1}^n p(x_i, y_i)$,

$$Pr((x^n, y^n) \in \mathcal{T}_\delta^n(X,Y)) \geq 1 - \epsilon_n'.$$

Lemma A.2 ([75, Theorem 3.1.2], [75, Theorem 15.2.1]). *There exists a sequence $\epsilon_n \to 0$ as $n \to \infty$ such that,*

$$|\frac{1}{n}\log_2(|\mathcal{T}_\delta^n(X)|) - H(X)| \leq \epsilon_n,$$
$$|\frac{1}{n}\log_2(|\mathcal{T}_\delta^n(Y)|) - H(Y)| \leq \epsilon_n,$$
$$|\frac{1}{n}\log_2(|\mathcal{T}_\delta^n(X,Y)|) - H(X,Y)| \leq \epsilon_n.$$

Lemma A.3 (Asymptotic equipartition property). *There exists a sequence $\epsilon_n \to 0$ as $n \to \infty$ such that, for each pair of sequences $(x^n, y^n) \in \mathcal{T}_\delta^n(X,Y)$,*

$$|-\frac{1}{n}\log_2(p_{X^n}(x^n)) - H(X)| \leq \epsilon_n,$$
$$|-\frac{1}{n}\log_2(p_{Y^n}(y^n)) - H(Y)| \leq \epsilon_n,$$
$$|-\frac{1}{n}\log_2(p_{X^n,Y^n}(x^n,y^n)) - H(X,Y)| \leq \epsilon_n.$$

Lemma A.4 (Conditional typicality lemma). *Let $(X,Y) \sim p(x,y)$. Let us assume that $x^n \in \mathcal{T}_\delta^n(X)$ and $Y^n \sim p(y^n|x^n) = \prod_{i=1}^n p_{Y|X}(y_i|x_i)$. Then, for every $\delta' > \delta$ and $0 < \epsilon < 1$,*

$$\lim_{n\to\infty} P\{(x^n, Y^n) \in \mathcal{T}_{\delta'}^n(X,Y)\} \leq 1 - \epsilon.$$

Lemma A.5 (Joint typicality lemma). *We let $(X,Y,Z) \sim p(x,y,z)$. If the (differential) entropy $I(y;z|x) < \infty$, and $\delta' < \delta$, there exists $\epsilon(\delta) > 0$ that goes to zero as δ tends to 0 that makes the following statements to be true when n is sufficiently large.*

- *For a pair of arbitrary sequences $(\tilde{x}^n, \tilde{y}^n)$ and $\tilde{Z}^n \sim \prod_{i=1}^n p_{Z|X}(\tilde{z}_i|\tilde{x}_i)$, we have*

$$P\{(\tilde{x}^n, \tilde{y}^n, \tilde{Z}^n) \in \mathcal{T}_\delta^n(X,Y,Z)\} \leq 2^{-n(I(Y;Z|X)-\epsilon(\delta))}.$$

- *For a pair of joint typical sequence $(x^n, y^n) \in \mathcal{T}_\delta^n(X,Y)$ and $\tilde{Z}^n \sim \prod_{i=1}^n p_{Z|X}(\tilde{z}_i|x_i)$, we have*

$$P\{(x^n, y^n, \tilde{Z}^n) \in \mathcal{T}_\delta^n(X,Y,Z)\} \geq 2^{-n(I(Y;Z|X)+\epsilon(\delta))}.$$

Lemma A.6 (Packing lemma). *The RVs (X,U,Y) follows the distribution $p(x,u,y)$. We assume $(\tilde{U}^n, \tilde{Y}^n)$ follows the distribution $p_{\tilde{U}^n,\tilde{Y}^n}(\tilde{u}^n, \tilde{y}^n)$ which is not necessarily i.i.d. distributed. $l \in \mathcal{L}$ is the index for random sequence $X^n(l)$*

with the cardinality of alphabet $|\mathcal{L}| \leq 2^{nR}$. Each of the sequence is distributed following $\prod_{i=1}^{n} p_{X|U}(x_i|\tilde{u}_i)$. The sequence $X^n(l)$ is pairwise conditionally independent of $\tilde{Y}^n$ given $\tilde{U}^n$. In addition, such sequence can be arbitrarily dependent on other X^n sequences. In this case, if $R < I(X;Y|U) - \epsilon(\delta)$ where $\epsilon(\delta)$ tends to zero as δ goes to zero, we have

$$\lim_{n\to\infty} P\{(X^n(l), \tilde{U}^n, \tilde{Y}^n) \in \mathcal{T}_\delta^n(X, \tilde{U}, \tilde{Y}) \text{ for some } l \in \mathcal{L}\} = 0.$$

Lemma A.7 (Covering lemma)**.** *Suppose $(X^n, Y^n) \sim p(x^n, y^n)$ is a pair of random sequences that are jointly typical w.p. 1 when n tends to infinity, i.e., $\lim_{n\to\infty} P\{(X^n, Y^n) \in \mathcal{T}_\delta^n(X, Y)\} = 1$. We consider random sequences $\hat{Y}^n(l)$ where $l \in \mathcal{L}$ with $|\mathcal{L}| \geq 2^{nR}$. These sequences $\hat{Y}^n(l)$, each distributed following $\prod_{i=1}^{n} p_{\hat{Y}|X}(\hat{y}_i|x_i)$, are conditionally independent of each other and of Y^n given X^n. In this case, if $R > I(\hat{Y};Y|X) + \epsilon(\delta)$ where $\epsilon(\delta)$ tends to zero as δ goes to zero, we have*

$$\lim_{n\to\infty} P\{(X^n, Y^n, \hat{Y}^n(l)) \notin \mathcal{T}_\delta^n(X, Y, \hat{Y}) \text{ for all } l \in \mathcal{L}\} = 0.$$

The conditional typicality lemma, the joint typicality lemma, the packing lemma, and the covering lemma are proved in [77]. In fact, the authors of [77] proved these lemmas for a more general setup of typicality, with strong/weak/robust typicality included. However, one should be aware that their work has not solved the problem that Markov lemma does not hold in general when weak typicality is concerned (it holds for some bounded cases). Please refer to [77] for detailed proofs. All previous definitions and properties can be extended to any finite number of variables. As shown in [77], the definition and the corresponding properties of weak typical sequences, e.g., asymptotic equipartition property (AEP) and other lemmas, are justified when the differential entropy is well-defined, that is, bounded away from infinity.

B

Resumé en Français

Dans ce chapitre, nous préludons à la motivation de notre travail, présentons plusieurs modèles que nous avons étudiés, et puis vous montrons certains résultats fondamentaux.

B.1 LA MOTIVATION

L'un des objectifs essentiels de chercheurs en communication est de fournir des techniques efficaces répondant aux exigences énorm pour l'accès à internet. Il y a trois raisons pour ces exigences de connexion considérables : 1) le nombre d'appareils mobiles est en croissance exponentielle au cours des dernières années; 2) le nombre d'applications de données sans fil dans chaque terminal augmente de façon considérable; 3) la quantité de données requises par chaque application devient immense. Pour tenir compte de ces demandes de trafic sans précédent, d'énormes appareils de types différents sont servis simultanément dans une système avec plusieurs BSs. Par conséquent, les concepteurs de systèmes auront à faire face à au moins trois questions fondamentales qui rendent le problème assez difficile, c'est-à-dire., *interférences*, *puissance limitée*, et *hétérogénéité* (hétérogénéité défini à la Section 1.1.3). En raison des terminaux d'utilisateurs massifs dans le réseau, le canal de communication est transféré du limité par bruit au limité par interférence. Pour la même raison, la ressource de puissance allouée à l'information destinée à chaque utilisateur est fortement limitée compte tenu d'une contrainte de puissance totale. Comme les techniques sans fil évoluent au fil des générations, les multiples BSs pourraient être soumis à des normes différentes. En tant que tel, l'incorporation de multiples BSs hétérogènes dans un réseau hétérogène est en effet difficile car elle peut les empêcher de formuler une coopération cohérente. De plus, l'hétérogénéité n'existe pas seulement du point de vue du BSs. Par exemple, lorsque des mobiles distincts souhaitent retourner certaines informations à la BSs pour aider l'réduction de l'interférence, les débits de la voie de retour de certains mobiles pourraient être négligeables en raison du budget de puissance limité au dispositif, de la fluctuation intense sur la qualité des canaux de la voie de retour, de l'absence de lien de ligne de visée pour une transmission au bon qualité, etc. Dans ce cas, il y a de l'hétérogénéité de la voie de retour car il est préférable d'arrêter ces liens de la voie de retour faibles pour avoir d'efficacité. Les trois aspects susmentionnés peuvent être modélisés et traités séparément, cependant, dans le système réel, ils se produisent principalement en même temps.

Il est bien connu à ce jour que la connaissance de CSIT est cruciale pour les systèmes de réseau sans fil susmentionnés avec multi-utilisateurs [1]. En

particulier, CSIT permet d'réduire l'interférence entre les flux d'informations, qui est le obstacle principal à des débits plus hauts dans les communications multi-utilisateurs. Néanmoins, CSIT opportun et précis est soit difficile, si possible du tout, soit coûteux à obtenir. Bien que la qualité de CSIT puisse être finalement améliorée avec les ressources suffisant (par exemple, estimation de canal, la voie de retour pour CSIT, etc.), l'opportunité dépend fortement de la nature du canal physique. En fait, CSI précis peut aussi devenir obsolète lorsqu'il est envoyé par les récepteurs à l'émetteur dans un environnement à mobilité élevée où le canal varie rapidement.

Dans le travail récent [2] sur le GBC, Maddah Ali et Tse ont proposé un schéma linéaire (alias MAT) et révélé qu'avec MAT, même complètement obsolète CSIT peut être exploité pour augmenter sensiblement le débit des deux utilisateurs MISO BC. Dans ce cas, il obtient un sum-DoF (identifié comme étant le facteur pré-log) de $\frac{4}{3}$ par rapport à 1 dans le cas où le CSIT est absent. Dans le même travail, le MAT est montré de pouvoir obtenir le DoF optimal pour le cas MISO avec K utilisatuers lorsque le nombre d'antennes d'émission n'est pas inférieur au nombre de récepteurs K. Cette conclusion reste valable tant que l'émetteur peut recevoir la voie de retour de CSI (avec un retard arbitraire mais fini) de tous les récepteurs et que l'erreur CSI s'évanouit avec un débit de $O(\mathsf{SNR}^{-1})$. Plus précisément, leur schéma est construit sur l'alignement espace-temps à l'aide du CSIT retardé [1], et est montré d'être optimal en termes de la DoF dans le cas du SNR haut. En fait, le même principe était déjà utilisé auparavant par Dueck dans [3]. L'optimalité est simple à prouver sachant que un borne supérieure peut être obtenu par ajouter un génie dans le système qui fournit la sortie de canal d'un utilisateur à l'autre et réduit ainsi le problème à celui du *physically-degraded* BC avec la voie de retour [4]. Après [2], plusieurs généralisations ont été effectuées, telles que, entre autres, les MIMO IC [5], $K-$user SISO IC/X channel [6], MIMO-BC/IC avec CSIT [7–9] MISO-BC avec CSIT [10], MISO-BC alternant CSIT [11], certains trois-utilisateur MIMO-BC [12, 13], communication sécurisée de MIMO-BC [14], d'autres travaux [15–18] et leurs références. Comme ces travaux soient percutants du point de vue de l'espace-signal, c'est-à-dire l'alignement espace-temps d'interférence, la plupart d'entre eux se concentrent sur DoF. Nous pouvons citer l'atténuation d'interférence aidé par l'état (avec le même principe) par MAT ou RIA (présentée pour la première fois dans [19]) pour la diversité.

B.1.1 *CSIT retardé: SNR fini*

Bien que l'atténuation de l'interférence activée par le CSIT retardé ait attiré une attention remarquable en raison de sa performance exceptionnelle dans le régime avec SNR haut, vérifiée par le gain de DoF, est-ce que le CSIT retardé est toujours efficace dans le GBC avec fading sous condition de un SNR *fini* semble être une question plus intéressante d'côté de la pratique dans les systèmes sans fil. Comme nous le savons, le schéma de MAT a une structure fixe construite sur un argument de comptage de dimension, l'optimalité DoF quand le SNR est haut ne peut pas assurer son efficacité quand le SNR est fini à cause de son insensibilité à tels paramètres. En fait, certains résultats sur les bornes inférieures sur la capacité du

[1] Le CSIT retardé et la voie de retour de l'état sont utilisés alternativement dans cette thèse.

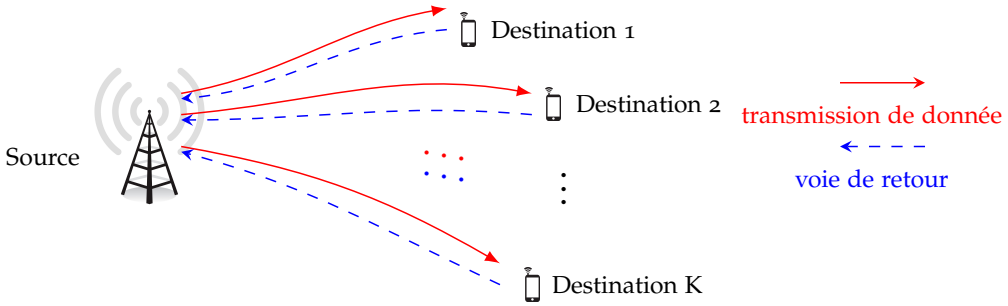

Figure 27: Un canal avec les voies de retour homogène

BC sans mémoire ont déjà été présentés, par exemple, avec une voie de retour généralisée dans les canaux généraux [20–22] et les références rélatifs. De plus, il y avait eu plusieurs autres résultats de régions débit réalisables obtenus par des schémas qui ont spécifiés pour le cas de SNR fini et ont basés sur la structure MAT susmentionnée, par exemple dans [23–26]. Dans [23], les auteurs ont développé pour K-utilisateur deux méthodes de précodage qui conçoivent les précodéreurs de telle sorte qu'une certaine quantité d'interférence à un des utilisateurs est autorisée afin que le signal reçu à l'autre est amélioré. Quand $K = 2, 3$, son gain de performance contre le MAT a été révélé lorsqu'un type spécifique de décodeur est sélectionné. De sorte à réduire la pénalité de multidiffusion[2] dans le schéma MAT, les auteurs de [24] ont inclus la quantification sur le signal qui porte l'auxiliaire message commun. Dans le cas de$K$-utilisateur sous Rayleigh fading, ils ont démontré qu'un écart entre la borne inférieure correspondante et une borne supérieure où certaines utilisateurs sont assisté par une genie, en termes de débit symétrique, est inférieur à $2\log_2(K+2)$ qui augumente sous-linéairement avec K. Plus récemment, [25] a analysé un scénario dans lequel les CSIs statistiques et l'information obsolète de la réalisations de canaux sont disponibles à l'émetteur. Il a été démontré par des exemples numériques que les CSIs statistiques peuvent agrandir la région débit quand les canaux sous fading de la type Rayleigh sont corrélée temporellement. Les auteurs de [26] ont étudié les performances d'interruption de service[3] dans le GBC avec un schéma adapté du MAT. Il est important de mentionner que ces schémas exploitent soit le codage linéaire tout seul, soit le codage linéaire avec une quantification à la structure de longueur de transmission fixe qui est utilisée dans le schéma de MAT. Malgré que un gain de performance contre MAT est possible avec certains des schémas, les débits dans le régime de SNR *moyen-à-bas* sont dominés par celui de la stratégie TDMA qui est optimal pour le cas sans CSIT. Par conséquent, trouver un schéma qui exploite le potentiel de CSIT retardé sous condition de un SNR *fini*/avec *puissance limitée* est encore un problème intéressant.

B.1.2 *CSIT retardé dans EBC et GBC*

Sur les travaux indépendantes, la région de capacité de EBC avec la voie de retour causale de l'état a été entièrement déterminée dans le cas de trois utilisateurs et partiellement caractérisée dans les cas où $K > 3$. L'idée principale derrière les schémas proposés dans ces références est fondamentalement la même: la source envoie d'abord les paquets non-codés, puis génère, selon l'état, des "combinaisons linéaire" adéquates des paquets perdus par certains récepteurs mais entendus par quelques autres récepteurs et multidiffuse ces combinaisons linéaires à certains récepteurs correspondants. Étant donné que suffisamment de combinaisons linéairement indépendantes sont disponibles à chaque récepteur, les paquets souhaités (messages souhaités) sont toujours prêts à être récupérés. Cependant, les schémas dans [27, 28] ne fonctionnent que pour un alphabet de paquet de taille 2^q avec $q \in \mathbb{N}$ tel que $2^q \geq K$. Ceci permet de garantir l'existence d'un nombre désiré de vecteurs qui sont linéairement indépendants dans l'espace vectoriel correspondant. Par conséquence, la région de capacité reste inconnu pour les autres tailles d'alphabet d'entrée de canal. Essentiellement, il y a une similitude frappante entre le schéma de MAT et les schémas proposés dans [27, 28]. Plus précisément, le schéma de MAT suit la routine comme: le émetteur envoi d'abord du signal non-codé, puis crée selon la voie de retour de CSI et multidiffuse des combinaisons linéaires des signaux entendus par hasard[4], afin que les messages peuvent être facilement récupérés, par exemple, par élimination Gaussienne, avec une erreur de décodage évanouie si suffisamment d'équations linéairement indépendants sont disponibles à chaque utilisateur. De plus, il existe une ressemblance entre la région DoF de GBC et la région de capacité de EBC. En fait, une telle similarité s'est également produite entre les régions de DoF/débit de GBC/EBC pour les communication sécurisées, comme indiqué dans [29, 30].

B.1.3 *La voie de retour hétérogène*

Comme nous l'avons déjà dit, les techniques d'atténuation des intéferences ne pourraient pas être efficaces dans les canaux physiques si nous excluons les limitations concernant une puissance limitée et/ou l'hétérogénéité. Motivé par ceci, nous considérons le RIA sous la condition de voie de retour hétérogène d'état. Pour faciliter la compréhension, nous énonçons brièvement la définition de la voie de retour d'état homogène/hétérogène spécifiquement pour cette thèse. Tous les noeuds de la communication participent au processus de feedback lorsque nous supposons une voie de retour homogène, par exemple, comme montré dans la Fig 27. En revanche, la voie de retour d'état hétérogène est définie comme seule une partie des noeuds de communication participe au processus de feedback; Tandis que le feedback d'état homogène signifie que le CSI retardé est transmis de toutes les destinations à tous les émetteurs. Pour être précis, nous donnons deux exemples pour illustrer deux types de la voie de retour d'état hétérogène. Le première exemple est que l'hétérogénéité de la voie de retour se trouve du

[2] multicast
[3] outage
[4] overheard signals

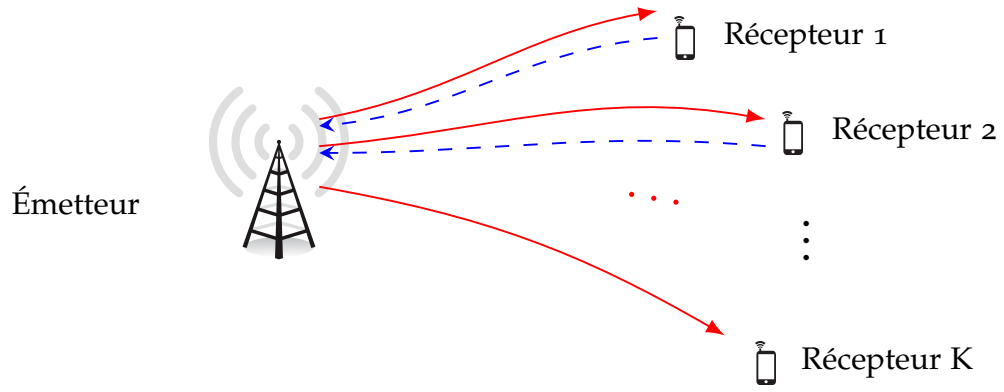

Figure 28: Un example où la hétérogénéité de la voie de retour est dans le côté des récepteurs

côté des récepteurs, c'est-à-dire que seulement certains récepteurs (une partie de tous) envoient leur CSI obsolète à tous les émetteurs alors qu'aucune CSI vient des autres récepteurs n'est détectée, par exemple, comme montré dans le Fig 28. Dans ce cas, la source ne sait pas l'état de toutes les connexions de la système et doit concevoir sa transmission selon l'état partiel/hétérogène. Le deuxième exemple est que l'hétérogénéité de la voie de retour se situe du côté des émetteurs, c'est-à-dire que, bien que tous les utilisateurs renvoient leurs états, peu d'émetteurs (parmi tous) sont capables de collecter les informations obtenus par les voies de retour à cause soit de la mauvaise qualité du lien de voie de retour ou soit de la désign du protocole.[5] L'un des exemples est montré dans le Fig 29. Même si les hétérogénéités peuvent se produire simultanément aux deux côtés de la communication, pour les raisons de clarté, nous omettons cette définition.

La voie de retour d'état de récepteur(s) partiel

Le CSIT hybride représente que les états des SDCs disponible à l'émetteur sont aux qualités différentes, par exemple, CSIT retardé ('D'), sans CSIT (N), CSIT parfait et instantané à la source (P). Comme il est omniprésent et général, le CSIT hybride, y compris la voie de retour hétérogène d'état comme cas particulier, est important en point de vue de pratique et a été étudié en profondeur avec un accent particulier sur DoF [7, 11, 31–37]. Bien que nous avons vérifié dans [38–40] le gain de performance avec l'état en qualité DD à presque tous les SNRs applicables, l'amélioration est montrée à disparaître même dans les régimes de SNR haut et le DoF s'effondre à l'unité lorsque le système soumit à l'état en qualité PN [41], et encore moins le cas DN. En d'autres termes, du point de vue de DoF, l'inclusion d'un utilisateur dans un canal point à point avec un CSIT parfait/retardé n'est pas bénéfique si le CSI du nouvel utilisateur est manquant. Des résultats similaires ont été rapportés dans [36] pour les cas de plus de deux utilisateurs. Le DoF total de un BC ayant trois utilisateurs et avec l'état DDN et PDN est limitée à $\frac{4}{3}$ et à $\frac{3}{2}$ qui coïncide avec le DoF optimal de BC ayant deux

[5] L'hétérogénéité peut être crée délibérément (par le concepteur du système) telle que les utilisateurs transmettents l'information des canaux seulement à quelques émetteurs pour faire face au problème de manque de puissance de feedback.

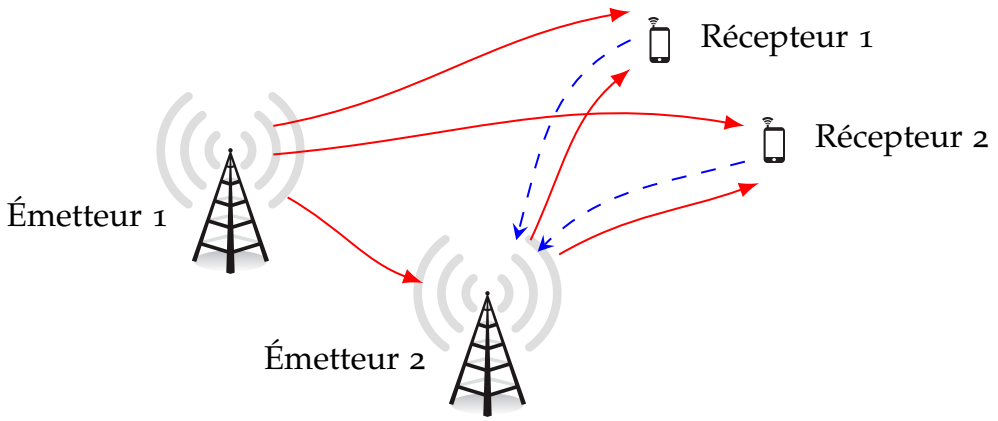

Figure 29: Un example où la hétérogénéité de la voie de retour est dans le côté des émetteurs

utilisateurs et avec l'état DD [2] et PD [42], respectivement.

En raison de la similarité (ou nous pouvons l'appeler uniformité) entre la région DoF de GBC et la capacité d'EBC dans le cas de l'état DD, il semble raisonnable de prédire une capacité dégradée dans l'EBC avec l'état DN compte tenu de la performance dégradée de DoF avec les même qualités d'état. De façon assez surprenante, Lin et Wang ont démontré dans [43, 44] que le CSI obtenu via la voie de retour d'un seul utilisateur peut toujours être utile lorsqu'on étudie un EBC qui est spatialement indépendant et en même temps est symétrique. Dans leur travaux, ils ont proposé un schéma multiphasique qui est capable d'exploiter un gain *opportuniste* de codage par codage réseau linéaire[6]. L'ingrédient essentiel de l'approche réside dans le fait que l'émetteur fait *XOR*-er les bits désirés mais effacés et les bits indésirés mais observés à un utilisateur (l'un qui fournit l'état retardé), les envoie comme l'information secondaire et laisse les deux utilisateurs décoder de façon opportuniste les bits voulus sachant les bits entendu précédemment. Encouragé par cela, nous avons étudié le cas quand l'hétérogénéité de voie de retour est dans le côté des récepteurs. En fait, les auteurs de [43, 44] ne sont pas les seuls chercheurs qui remarquent le potentiel d'une voie de retour d'état hétérogène. En particulier, [41, 45, 46] a pris le cas de un BC dont la taille de alphabet d'état est fini et qui a le AWGN comme contre-exemple de la conjecture que la performance est dégradée pour le pmf général, c'est-à-dire, un DoF total strictement supérieur de l'unité est montré dans [47, 48] d'être réalisable avec l'état DN. Indépendamment de la différence sur les outils de codage utilisés dans [43, 44, 47, 48], l'idée derrière est cohérente. La source code les informations secondaires seulement en fonction de l'état partiel alors que le récepteur connaît les deux états de sorte qu'il peut décider quelle partie à décoder selon que la réalisation d'état correspondante est conforme à celle du schéma qui marche pour le cas de DD, par exemple le récepteur 2 décide de décoder les bits auxiliaires dont les états sont $(S_1, S_2) = (1, 0)$ parce que ces bits sont souhaités par l'utilisateur 2 et sont codés dans les informations secondaires.

[6] linear netword coding

State feedback to partial transmitter(s)

Étant donné que des BSs dense qui soumettents à des normes distinctes sont déployées dans le réseau, cela est intuitif d'inclure la coopération entre des BS à plusieurs niveaux via un réseau câblé (un système d'antenne distribuée avec backhaul) ou sans fil (relais) pour garantir une transmission cohérente. Ainsi, nous considérons une configuration de réseau hétérogène typique dans laquelle une macrocell BS dessert deux utilisateurs qui sont à proximité d'un BS de une cellule petite (par exemple, femto/pico BS). Une question directe est de savoir si et comment smallcell BS pourrait améliorer les performances de liaison descendante[7] du réseau macrocell. Essentiellement, le smallcell BS dans cette système peut être considéré comme un relais, et le canal global devient un BRC.

Nous rappellons l'idée principale de MAT, qui consiste en deux phases: 1) la phase de diffusion dans laquelle le signal contenant des nouveaux messages est envoyé aux récepteurs, et 2) la phase de multidiffusion dans laquelle l'information secondaire basé sur l'information de canal passé et les messages envoyés est diffusée. Il est donc instinctif de penser à l'émetteur comme deux émetteurs virtuels distincts: l'un prend en charge dans la première phase sans avoir besoin de CSIT, alors que l'autre est responsable de la deuxième phase et utilise le CSIT retardé et le message transmis. Le deuxième émetteur virtuel ressemble beaucoup à un relais qui "apprend" les messages et les coefficients de canal et "renvoie" les informations secondaires nécessaires pour "aider" les récepteurs à décoder leurs propres messages.

Avec le raisonnement ci-dessus, dans le chapitre 5, nous déployons un relais d'infrastructure qui permet d'accomplir le RIA *a la MAT* avec l'état hétérogène de tous les mobiles vers le smallcell BS (relais) au lieu du macrocell BS (source). Notre réglage sur la voie de retour est principalement inspiré par le fait que l'envoie le CSI par la voie de retour avec précision au macrocell BS est coûteuse en pratique. Le coût est en termes à la fois de la largeur de bande de liaison montante[8] et de la puissance d'émission de liaison montante liée à la voie de retour. Puisque le relais est supposé être plus proche des destinations que la source, l'envoie au relais peut être plus faisable. En d'autres termes, la communication de retour est localisée et la puissance de cette communication pourrait être réduite grâce à un bon qualité de canal.

Dans les sections suivantes, nous presentons trois types de modèles de canal, c'est-à-dire, BC avec une voie de retour d'état homogène, BC avec une voie de retour d'état hétérogène et BRC avec une voie de retour d'état au relais. Dans ces modèles, nous fournissons certains des résultats que nous avons obtenus sans déranger les lecteurs avec des détails. En ce qui concerne la longueur d'un résume, l'examen de tous les résultats ne semble pas possible. Par conséquent, nous présentons les résultats les plus intéressants.

[7] downlink
[8] uplink

B.2 LE CANAL DE DIFFUSION AVEC DES VOIES DE RETOUR HOMOGÈNE D'ÉTAT

Comme une première étape à vérifier l'amélioration potentielle du RIA avec SNR fini, nous choisissons le modèle le plus simple possible dans le problème de codage du réseau, à savoir, le BC. Le canal est un BC sans mémoire avec K-utilisateur dépendant de l'état, par lequel la source souhaite communiquer avec l'aide de la voie de retour de l'état de toutes les destinations, dans n canal utilise[9], K messages indépendants aux K récepteurs, respectivement. Le canal peut être décrit par le pmf joint,

$$p_{\mathbb{Y}_1,\cdots,\mathbb{Y}_K|\mathbb{X},\mathbb{S}}(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_K|\boldsymbol{x},\boldsymbol{s})p_{\mathbb{S}}(\boldsymbol{s}) = \prod_{i=1}^{n} p_{Y_1,\cdots,Y_K|X,S}(y_{1i},\ldots,y_{Ki}|x_i,s_i)p_S(s_i), \quad (151)$$

où $\boldsymbol{x} \in \mathcal{X}^n$, $\boldsymbol{y}_k \in \mathcal{Y}_k^n$, et $\boldsymbol{s} \in \mathcal{S}^n$ pour tout $k \in \mathcal{K} \triangleq \{1, 2, \cdots, K\}$ sont les séquences de l'entrée du canal, k-ème sortie du canal, et l'état de canal, respectivement. $\mathcal{X}$, $\mathcal{Y}$, $mathcalS$ sont respectivement les alphabets de l'entrée du canal, de la sortie du canal et de l'état. L'information d'état de canal global est connue instantanément à tous les récepteurs, par exemple, chaque utilisateur a accès à tous les événements d'erreur de EBC ou à tous les coefficients de canal de GBC immédiatement après les transmissions. Du côté de l'émetteur, l'état de canal est connu strictement causalement sans erreur, par exemple via une voie de retour sans bruit. Sauf s'il est indiqué clairement, le coût du retour de l'état n'est pas pris en compte dans toute la dissertation car nous supposons que des ressources suffisantes sont allouée pour garantir un débit illimité pour la voie de retour. Par souci de brièveté, nous supposons que l'information d'état est fournie à l'émetteur avec un retard de un canal utilise et que le canal lui-même est i.i.d. temporellement. Les détails du modèle de canal sont montrées sur Fig 30. Comme nous le voyons sur Fig 30, l'information d'état des premiers $i-1$ canaux utilises, à savoir s^{i-1}, est disponible à l'émetteur uniquement à la fin du $i-$ème canal utilise. Pour tout sous-ensemble $\mathcal{U} \subseteq \mathcal{K}$, nous définissons $S_{\mathcal{U}}$ et $Y_{\mathcal{U}}$ comme les combinaisons d'états et de sorties de canaux des utilisateurs dans l'ensemble $\mathcal{U}$, respectivement.

Soit $M_k \in \mathcal{M}_k \triangleq \{1, 2, \cdots, 2^{nR_k}\}$ le message pour l'utilisateur k choisi dans l'ensemble de message $\mathcal{M}_k$, pour tous les $k \in \mathcal{K}$. Nous disons que le tuple de débits $(R_1, \ldots, R_K)$ est réalisable s'il existe

- une séquence des fonctions de encodage $\{f_i : \mathcal{M}_1 \times \cdots \times \mathcal{M}_K \times \mathcal{S}^{i-1} \to \mathcal{X}\}_{i=1}^{n}$,
- K fonctions de decodage $\{g_k : \mathcal{Y}_k^n \times \mathcal{S}^n \to \mathcal{M}_k\}_{k=1}^{K}$,

telle que $\max_k \Pr\big(g_k(Y_k^n, S^n) \neq M_k\big) \to 0$ quand $n \to \infty$. Le débit symétrique R_{sym} est réalisable si le tuple de débits $(R_{\text{sym}}, \ldots, R_{\text{sym}})$ est réalisable.

[9]channel use

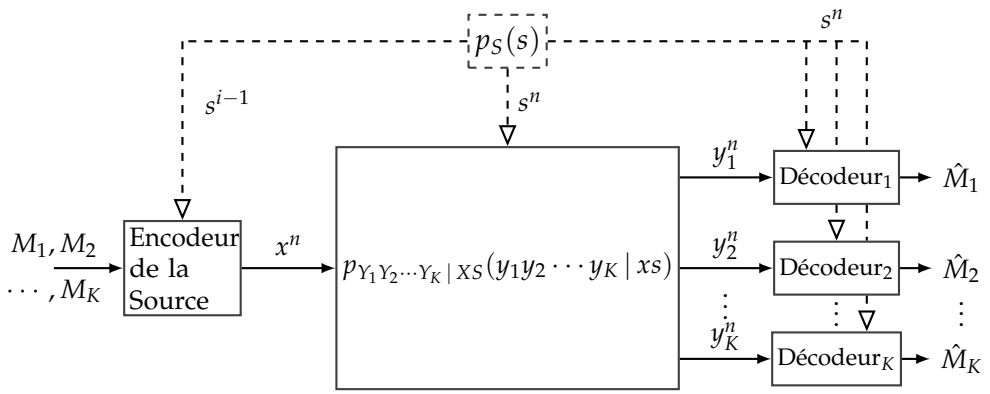

Figure 30: Modèle de un système general de BC avec K-utilisateur et la voie de retour d'état.

B.3 BORNES INFÉRIEURES ET SUPÉRIEURES DANS CANAL GÉNERAL

Dans le canal que nous présentons, nous avons obtenus quelque bornes inférieures et supérieures. Nous vous montrons les bornes ci-dessous et les détails et les explication des scémas sont fournis uniquement dans les chapitres correspondantes sachant que l'espace pour ce résumé est limité.

Theorem B.1 (Borne Inférieure de Markov de bloc)**.** *Une région de débit réalisable de un BC sans mémoire dépendant de l'état avec des voies de retour d'état est donnée par l'ensemble de toutes les paires de débits (R_1, R_2) satisfaisant:*

$$R_1 \leq \min\left\{I(V_1; Y_1\hat{Y} \mid V_{12}SQ), I(V_1V_{12}; Y_1|SQ) - I(\hat{Y}; V_2X|V_1Y_1V_{12}SQ)\right\}, \tag{152a}$$

$$R_2 \leq \min\left\{I(V_2; Y_2\hat{Y} \mid V_{12}SQ), I(V_2V_{12}; Y_2|SQ) - I(\hat{Y}; V_1X|V_2Y_2V_{12}SQ)\right\}, \tag{152b}$$

pour tous les pmf vérifiant

$$\begin{aligned} p(y_1, y_2, x, v_{12}, v_1, v_2, s, q) &= p(y_1, y_2|x, s)p(\hat{y} \mid x, v_{12}, v_1, v_2, s) \\ &\times p(x|v_{12}, v_1, v_2, q)p(v_1|q)p(v_2|q)p(v_{12}|q)p(s)p(q). \end{aligned} \tag{153}$$

Démonstration. Voir l'annexe 2.A. ■

Theorem B.2 (Borne Supérieure)**.** *Toute paire de débit réalisable (R_1, R_2) pour le canal de diffusion avec le retour d'état doit satisfaire*

$$R_1 \leq \min\{I(U_1; Y_1 \mid S),\ I(X; Y_1, Y_2 \mid U_2, S)\} \tag{154a}$$

$$R_2 \leq \min\{I(U_2; Y_2 \mid S),\ I(X; Y_1, Y_2 \mid U_1, S)\} \tag{154b}$$

pour quelques pmf $p(y_1, y_2, s \mid x)p(x, u_1, u_2)$.

Démonstration. Voir la section 2.2. ■

Theorem B.3 (Bornes Inférieure de DC). *Un tuple des débits $(R_1, \ldots, R_K)$ est réalisable avec le schéma DC dans un BC de K−utilisateur avec les voies de retour d'état causal pourvu que*

$$R_k \leq \alpha_1 I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \mid S^{(1)}, Q^{(1)}), \tag{155a}$$

$$0 \leq \min_{\substack{i,j,k,\mathcal{J}:\\ i<j,k\in\mathcal{J}}} \Big\{ \alpha_j I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \mid S^{(j)}, Q^{(j)}) - \alpha_i I(\{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \mid Y_k^{(i)}, S^{(i)}, Q^{(i)}) \Big\}, \tag{155b}$$

pour certains K-tuple $(\alpha_1, \ldots, \alpha_K) \in \mathbb{R}_+^K$ avec $\sum_k \alpha_k = 1$, et certains pmfs mentionnés dans le chapitre 3.

Démonstration. Voir la section 3.3. ■

Dans les deux sous-sections suivantes, deux canaux de sous-classes de BC sont définis dans lesquels nous illustrons comment la voie de retour d'état homogène est exploitée dans ces canaux.

B.3.1 *Le canal de diffusion Gaussien avec fading*

Le canal gaussien est utilisé fréquemment dans la recherche car il capture les propriétés principales des communications sans fil. La région de capacité du GBC est connue lorsque l'état du canal est déterministe [51, 52]. Lorsque le canal est sujet à une fading inconnue à l'émetteur (pas de CSIT), la capacité est toujours inconnue sauf pour certains processus de fading particuliers comme le cas où les deux utilisateurs subissent exactement la même distribution de fading [53]. Pour les cas plus généraux, seules les caractérisations avec SNR haut en termes de DoF sont disponibles dans la littérature.

Le GBC sous la contrainte de fading est un canal de AWGN dépendant de l'état qui est avec n_t antennes de transmission et $n_{r,k}$ antennes de réception à l'utilisateur k, définies par

$$\boldsymbol{y}_{k,i} = \boldsymbol{H}_{k,i}\boldsymbol{x}_i + \boldsymbol{z}_{k,i}, \quad k = 1, \ldots, K;\ i = 1, \ldots, n \tag{156}$$

où $\boldsymbol{z}_{k,i} \in \mathbb{C}^{n_{\mathrm{r},k}\times 1}$, $\boldsymbol{x}_i \in \mathbb{C}^{n_{\mathrm{t}}\times 1}$, $\boldsymbol{y}_{k,i} \in \mathbb{C}^{n_{\mathrm{r},k}\times 1}$, et $\boldsymbol{H}_{k,i} \in \mathbb{C}^{n_{\mathrm{r},k}\times n_{\mathrm{t}}}$. De plus, nous supposons que le bruit est $\boldsymbol{z}_{k,i} \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{n_{\mathrm{r},k}})$ et l'entré de canal est $\boldsymbol{x}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{Q})$. L'entré de canal est sujet aux contraintes de puissance $\frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{x}_i\|^2 \leq P = \mathrm{tr}(\boldsymbol{Q})$ pour toutes les séquences d'entré $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. l'indice $i \in \{1, \cdots, n\}$ est omis en géneral et sera spécifié si nécessaire. Le SNR est definié comme $\mathsf{SNR} \triangleq \frac{P}{n_{\mathrm{t}}\sigma^2}$. Sauf s'il est indiqué, nous supposons que les matrices de canal sont indépendants spatialement.

Nous identifions l'état S_k avec la matrice de canal $\boldsymbol{H}_k$. De la même façon que pour $S_{\mathcal{U}}$, nous utilisons $\boldsymbol{H}_{\mathcal{U}}$ pour réprésenter une matrice à partir d'une

concaténation verticale de tous les matrices $\{\boldsymbol{H}_k\}_{k\in\mathcal{U}}$, la même notation s'applique pour $\boldsymbol{Y}_\mathcal{U}$ et $\boldsymbol{Z}_\mathcal{U}$. Par conséquent, il s'ensuit que $\boldsymbol{Y}_\mathcal{U} = \boldsymbol{H}_\mathcal{U}\boldsymbol{X} + \boldsymbol{Z}_\mathcal{U}$.

Definition B.1. *Un BC Gaussien sous fading comme indiqué ci-dessus est symétrique si la matrice de canaux $\boldsymbol{H}_k$ suit une distribution identique pour tout $k \in \mathcal{K}$, c'est-à-dire, $f(\boldsymbol{H}_k = \boldsymbol{H}) = f(\boldsymbol{H}_{k'} = \boldsymbol{H})$, $k, k' \in \mathcal{K}$, où $f(\cdot)$ définit la fonction de densité de probabilité.*

Corollary B.4. *Dans un GBC symétrique sous fading, le schéma de Markov de bloc atteind un débit symétrique au sous,*

$$R_{sym} = \max_{\beta_i \geq 0} \left(K + \sum_{j=2}^{K} \binom{K}{j} \prod_{t=2}^{j} \frac{\sum_{l\leq t} b_{l,t}}{a_t} \right)^{-1} a_1 \tag{157}$$

où, pour $t = 1, \ldots, K$,

$$a_t \triangleq \mathbb{E}_{\boldsymbol{H}_\mathcal{T}} \log\det\left(\boldsymbol{I} + \mathsf{SNR}\, \boldsymbol{H}_\mathcal{T}^H \boldsymbol{\Lambda}_t \boldsymbol{H}_\mathcal{T}\right) \tag{158a}$$

$$b_{l,t} \triangleq \mathbb{E}_{\boldsymbol{H}_1,\boldsymbol{H}_l} \log\det\left(\boldsymbol{I} + \frac{\mathsf{SNR}}{\beta_t} \boldsymbol{H}_l (\boldsymbol{I} + \mathsf{SNR}\, \boldsymbol{H}_1^H \boldsymbol{H}_1)^{-1} \boldsymbol{H}_l^H\right) \tag{158b}$$

avec $\mathcal{T} \triangleq \{1\} \cup \{t+1, \ldots, K\}$, $\boldsymbol{\Lambda}_t \triangleq \mathrm{diag}\big(\boldsymbol{I}_{n_r}, \beta_t^{-1}\boldsymbol{I}_{(K-t)n_r}\big)$.

B.3.2 *Le canal de diffusion d'effacement*

Le canal d'effacement joue lui-même un rôle important dans la communication en raison de son utilité sur la modélisation de l'erreur de décodage de paquets à partir d'une couche supérieure au-delà de la couche physique. Un EBC est un BC dont les sorties sont sujettes à d'éliminations possibles.

Le EBC est un canal déterministe dépendant de l'état défini par

$$Y_k = \begin{cases} X, & S_k = 1,\ w.p.\ 1-\delta_k \\ E, & S_k = 0,\ w.p.\ \ \delta_k \end{cases}, \tag{159}$$

qui est complètement caractérisé par l'ensemble des probabilités de la réalisation de l'état

$$\mathrm{P}_{\mathcal{U},\bar{\mathcal{U}}} \triangleq \Pr(S_\mathcal{U} = \mathbf{0},\ S_{\bar{\mathcal{U}}} = \mathbf{1}), \quad \mathcal{U} \subseteq \{1, \ldots, K\}, \tag{160}$$

avec $\sum_\mathcal{U} \mathrm{P}_{\mathcal{U},\bar{\mathcal{U}}} = 1$. Tout au long de la thèse, nous utilisons $S_\mathcal{U} = \mathbf{0}$ (*resp.* $S_\mathcal{U} = \mathbf{1}$) à dire que $S_k = 0$ (*resp.* $S_k = 1$), $\forall\ k \in \mathcal{U}$. Par souci de simplicité, nous désignons $\Pr(S_\mathcal{F} = 0)$ avec $\delta_\mathcal{F}$ et $\Pr(S_\mathcal{F} = 0, S_\mathcal{T} = 1)$ avec $\mathrm{P}_{\mathcal{F},\mathcal{T}}$ pour tous les $\mathcal{F}$ et $\mathcal{T}$ satisfant $\mathcal{F} \cap \mathcal{T} = \phi$ et $\mathcal{F} \cup \mathcal{T} \subseteq \mathcal{K}$, qui correspondent à l'événement que des effacements se produisent à tous les utilisateurs dans l'ensemble $\mathcal{F}$ et l'événement que chacun dans l'ensemble $\mathcal{F}$ rencontre un effacement tandis que chaque utilisateur dans l'ensemble $\mathcal{T}$ reçoit parfaitement le signal d'entrée, respectivement. Dans le chapitre 3 et 4, nous utilisons $\mathcal{I}$ et $\mathcal{J}$ pour certains sous-ensembles d'utilisateurs

avec des contraintes implicites de taille $|\mathcal{I}| = i$ et $\|\mathcal{J}| = j$, respectivement. Les contraintes sont explicitées lorsque cela est nécessaire. En outre, $\mathcal{F}$, $\mathcal{T}$, $\mathcal{U}$ sont également supposés être les sous-ensembles d'utilisateurs mais sans contraintes de taille particulières. Nous récitons le Lemme 10 dans [28] ici qui montre que $\mathrm{P}_{\mathcal{F},\mathcal{T}}$ et δ sont connectés par,

$$\mathrm{P}_{\mathcal{F},\mathcal{T}} = \sum_{\mathcal{U}\subseteq\mathcal{T}} (-1)^{|\mathcal{U}|}\delta_{\mathcal{F}\cup\mathcal{U}}, \tag{161}$$

Pour plus de commodité, $\delta_{\{k\}}$ est écrit comme δ_k. Nous définissons en outre deux sous-classes de canal d'effacement, symétrique EBC et spatialement indépendant EBC.

Definition B.2. *On dit qu'un EBC est symétrique si pour tous les $\mathcal{J}$, le probabilité d'effacement $\delta_{\mathcal{J}}$ est déterminée uniquement par le cardinality de l'ensemble $\mathcal{J}$, c'est-à-dire, $\delta_j = \delta_{\mathcal{J}} = \delta_{\mathcal{J}'}$ tient toujours pour tous les ensembles $\mathcal{J}$ et $\mathcal{J}'$ telle que $|\mathcal{J}| = |\mathcal{J}'| = j$.*

Definition B.3. *On dit qu'un EBC est spatialement indépendant tant que les effacements à travers les utilisateurs sont i.i.d.. Pour l'EBC spatialement indépendant, nous avons:*

$$\delta_{\mathcal{U}} = \prod_{k\in\mathcal{U}} \delta_k, \quad \mathcal{U} \subseteq \mathcal{K}, \tag{162}$$

qui donne

$$\mathrm{P}_{\mathcal{F},\mathcal{T}} = \prod_{k\in\mathcal{F}} \delta_k \prod_{k'\in\mathcal{T}} (1-\delta_{k'}), \quad \mathcal{F}\cap\mathcal{T} = \phi,\ \mathcal{F}\cup\mathcal{T} \subseteq \mathcal{K}. \tag{163}$$

Sans perte de généralité, nous supposons que les probabilités d'effacement marginales remplissent $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_K$. Nous réaffirmons la définition de 'one-sided fairness' ou 'one-sidedly fair condition' (définie pour la première fois dans [27, 28]) comme suit.

Definition B.4. *On dit qu'un tuple de débit $(R_1, \cdots, R_K)$ satisfe one-sided fairness si pour tout les $\delta_i \geq \delta_j$ et tout les $i \leq j$, $R_i\delta_i \geq R_j\delta_j$ est correct.*

Corollary B.5. *Pour tout alphabet d'entrée $|\mathcal{X}|$, le débit $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ est réalisable avec le schéma de DC dans un EBC avec les voies de retour si la suite est satisfé,*

$$R_k \leq \alpha_1 P(Q_2^{(1)} = k)(1-\delta_{\mathcal{K}})\log|\mathcal{X}|, \tag{164a}$$

$$0 \leq \min_{\substack{j,k,\mathcal{J}:\\ k\in\mathcal{J}}} \Big\{ \alpha_j P(Q_2^{(j)} = \mathcal{J})(1-\delta_{\mathcal{K}\setminus\mathcal{J}\cup\{k\}}) - \sum_{i=1}^{j-1} \alpha_i \sum_{\mathcal{I}\subset\mathcal{J},\mathcal{I}\ni k} P(Q_2^{(i)} = \mathcal{I})\mathrm{P}_{\mathcal{K}\setminus\mathcal{J}\cup\{k\},\mathcal{J}\setminus\mathcal{I}} \Big\}, \tag{164b}$$

pour quelques K-tuple $(\alpha_1, \ldots, \alpha_K) \in \mathbb{R}_+^K$ avec $\sum_k \alpha_k = 1$.

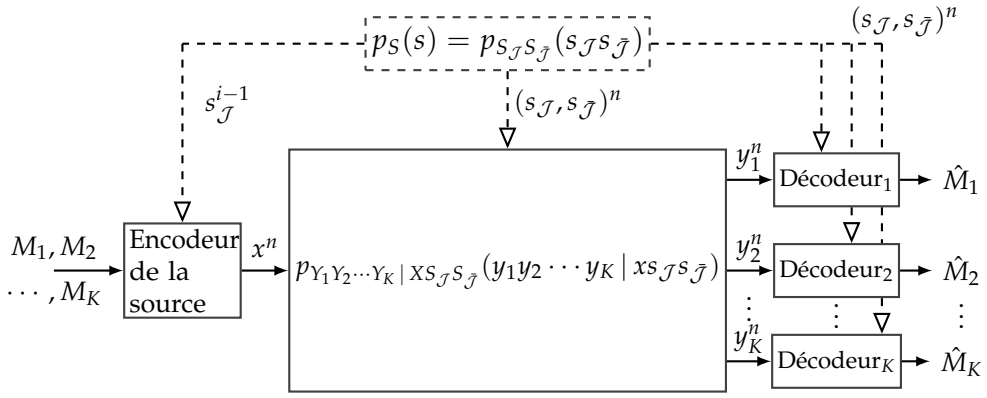

Figure 31: Le modèle de un système géneral: un BC multi-utilisatuer avec les voies de retour pratielles d'état.

B.4 LE CANAL DE DIFFUSION AVEC LES VOIES DE RETOUR HÉTÉOGÈNE D'ÉTAT

Nous avons étudié le BC avec une voie de retour hétérogène dans le chapitre 4. Le modèle qui nous intéresse est semblable à celui défini dans la section B.2 dans le sens que les deux systèmes se composent d'une source unique, de destinations multiples, et d'un BC via que K messages indépendants de la source sont envoyés vers la destination désirée. La différence réside dans les informations d'état connues à la source. Dans ce canal, nous divisons l'ensemble d'état en deux parties comme $S = (S_{\mathcal{J}}, S_{\bar{\mathcal{J}}})$ où $\bar{\mathcal{J}} \triangleq \mathcal{K} \setminus \mathcal{J}$ et $S_{\bar{\mathcal{J}}}$ est le complément de l'ensemble d'état $S_{\mathcal{J}}$.

Les informations d'état sont disponibles *hétérogènement* aux nœuds distincts. En particulier, toutes les destinations connaissent instantanément les informations d'état globales. Au contraire, l'émetteurs apprend seulement une *partie* du CSI obsolète, à savoir $S_{\mathcal{J}}$, par exemple, les utilisateurs dans l'ensemble $\mathcal{J}$ retournent leurs états à la source. Le modèle détaillé de système est montré sur Fig. 31.

Dans ce canal, les fonctions de encodage et decodage sont modifiés comme suit.

- Une séquence de fonctions de encodage affecte $X_i(M_1, M_2, \cdots, M_K, S_{\mathcal{J}}^{i-1})$ aux messages et à une partie de la séquence d'état $S_{\mathcal{J}}^{i-1} \in {\mathcal{S}_{\mathcal{J}}}^{i-1}$ comme indiqué ci-dessous,

$$\{\varphi_{s,i} : \mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_K \times {\mathcal{S}_{\mathcal{J}}}^{i-1} \to \mathcal{X}\}_{i=1}^n.$$

- Chacune des K fonctions de décodage affecte un index $\hat{M}_k$ sur la base des sorties de canal qui sont observée $(y_k^n, s_{\mathcal{J}}^n, s_{\bar{\mathcal{J}}}^n)$.

$$\phi_k : \mathcal{Y}_k^n \times \left(\mathcal{S}_{\mathcal{J}}, \mathcal{S}_{\bar{\mathcal{J}}}\right)^n \to \mathcal{M}_k.$$

On dit que un tuple de débit $(R_1, \ldots, R_K)$ est réalisable s'il existe une séries de

fonction de codage et décodage tel que

$$\limsup_{n\to\infty} \max_k \mathbb{E}_{S^n_{\mathcal{J}}, S^n_{\bar{\mathcal{J}}}} \left[\Pr\left(\phi_k(Y^n_k, S^n_{\mathcal{J}}, S^n_{\bar{\mathcal{J}}}) \neq M_k\right)\right] \to 0.$$

Sans perte de généralité, nous supposons que les utilisateurs dans $\mathcal{J}$ qui retournent les états sont indexés avec $1, 2, \cdots, j$ et donc le reste qui ne fournit aucune information d'état à la source est indexé avec $j+1, \cdots, K$. Quand $\mathcal{J} = \mathcal{K}$, il tombe dans le scénario qui est été étudié dans le chapitre 3, c'est-à-dire, une voie de retour d'état homogène. Pour plus de commodité de notation, nous identifions le CSIT en qualité différente: le CSIT instantané/parfait, le CSIT retardé, ou non connu avec P, D, N [11, 33]. Bien que notre motif essentiel est d'explorer l'impact de la voie de retour hybride d'état DN, les notations PDN sont toujours définies pour les comparer.

Definition B.5. *Dans le BC avec $K-$utilisateur, on dit que la voie de retour d'état est dans le cas de $D_jN_lP_{K-j-l}$ si le CSI retardé des utilisateurs $1, \cdots, j$ sont obtenus par l'émetteur via une voie de retour parfait, tandis que les utilisateurs $j+1, \cdots, j+l$ n'informent aucune d'état à la source, alors que la source connaît l'état des destinations $j+l+1, \cdots, K$ parfaitement.* [10]

B.4.1 *Un codage non-lineaire dans un EBC*

Theorem B.6 (Bornes Inférieure $\mathcal{R}_{BM}$)**.** *La paire de débit (R_1, R_2) est réalisable pour un BC avec deux utilisatuers dans le cas de DN (D_1N_1) avec une adaptation du schéma de Markov de bloc si (R_1, R_2) sont dans la région $\mathcal{R}_{BM}$ qui est défini par les contraintes dessous,*

$$\begin{aligned} R_1 \leq \min\{&I(V_1; Y_1\hat{Y}|V_{12}S_1S_2Q), I(V_1; Y_1\hat{Y}|V_{12}S_1S_2Q) + I(V_{12}; Y_1|S_1S_2Q) \\ &+ I(\hat{Y}; Y_1|V_{12}S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q)\}, \qquad (165a) \\ R_2 \leq \min\{&I(V_2; Y_2\hat{Y}|V_{12}S_1S_2Q), I(V_2; Y_2\hat{Y}|V_{12}S_1S_2Q) + I(V_{12}; Y_2|S_1S_2Q) \\ &+ I(\hat{Y}; Y_2|V_{12}S_1S_2Q) - I(\hat{Y}; V_1V_2|V_{12}S_1Q)\}, \qquad (165b) \end{aligned}$$

pour quelques pmf's qui vérfie,

$$\begin{aligned} p(y_1, y_2, \hat{y}, x, v_1, v_2, v_{12}, s_1, s_2, q) =& p(y_1, y_2|x, s_1, s_2)p(\hat{y}|x, v_{12}, v_1, v_2, s_1) \\ &\times p(x|v_{12}, v_1, v_2, q)p(v_1|q)p(v_2|q)p(v_{12}|q)p(s_1)p(s_2)p(q). \end{aligned}$$

Démonstration. Voir la section 4.2. ■

Corollary B.7 (Bornes Inférieure $\mathcal{R}'_{BM}$)**.** *Pour un EBC ayant deux utilisateurs avec le CSI à l'émetteur obtenu par la voie de retour de une seul utilisatuer, la paire de débit (R_1, R_2) est réalisable tant que il est dans la région $\mathcal{R}'_{BM}$ qui vérifie,*

[10] Omettre les indices (sous lettre) de l'état (P, D, N) est autorisé pour la simplicité quand aucun utilisateur ne tombe dans la catégorie correspondante, $D_2N_0P_0$ est écrit comme D_2.

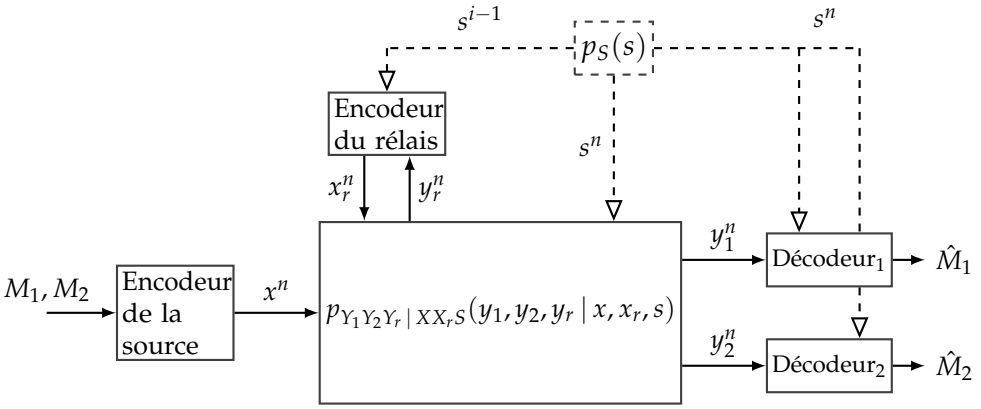

Figure 32: Le modèle d'un système géneral: un BRC de deux-utilisateur avec la voie de retour d'état au relais.

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1, \tag{166a}$$

$$\frac{R_1}{\frac{1-\delta_2}{1-\delta_2+\delta_{12}}} + \frac{R_2}{1-\delta_2} \leq 1. \tag{166b}$$

Démonstration. Voir la section 4.2. ■

Theorem B.8 (Borne Supérieure)**.** *Toutes les paire de débit réalisable* (R_1, R_2) *pour le EBC avec le retour d'état de un seul utilisatuer* (D_1N_1) *doit satisfaire*

$$\frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_{12}} \leq 1, \tag{167a}$$

$$\frac{R_1}{1-\delta_{12}} + \frac{R_2}{1-\delta_2} \leq 1. \tag{167b}$$

Démonstration. Voir la section 4.2. ■

La même région est également réalisable avec une autre schéma adaptée du schéma DC. En outre, nous avons également des résultats sur le cas où il y a plus de trois utilisateurs, néanmoins, en raison de la limitation de l'espace, nous les omettons dans cette partie et veuillez vous référer au chapitre 4 pour plus de détails.

B.5 LE CANAL RÉLAIS DE DIFFUSION AVEC LA VOIE DE RETOUR D'ÉTAT AU RÉLAIS

Le canal de liaison descendante est dans une classe plus large de un BRC sans mémoire de deux-utilisateur, comme montré dans Fig. 32, qui est défini par le

pmf joint

$$p(y_1^n, y_2^n, y_r^n | x^n, x_r^n, s^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}, y_{ri} | x_i, x_{ri}, s_i) \tag{168}$$

où x^n et x_r^n désignent la séquence de n symboles d'entrée à la source et au relais, respectivement; Y_1^n,, y_2^n, et y_r^n correspondent respectivement aux sorties des canaux du récepteur 1, du récepteur 2 et du relais intermédiaire; s^n définit l'état du canal qui détermine le comportement du canal étant donné que certaines entrées de x^n et x_r^n sont choisis. Comme indiqué dans (168), le canal est temporellement i.i.d. et ainsi les sorties sont décidées exclusivement par les entrées du canal et l'état actuel. À la $i$ème plage horaire, seul l'état des plages horaire précédents s^{i-1} est disponible au relais alors que l'état instantané s^i est connu par décodeur lors de la réception.

L'encodeur de la source mappe chaque paire de messages M_1 et M_2 à une séquence de symboles d'entrée x^n, où $x \in \mathcal{X}$; [11] le codeur du relais mappe ses observations passées (y_r^{i-1}, s^{i-1}), à partir de la source et des destinations via le lien de voie de retour, à un symbole $x_{r,i}$ à chaque plage horaire i, où $y_r^{i-1} \in \mathcal{Y}_r^{i-1}$, $s^{i-1} \in \mathcal{S}^{i-1}$ et $x_{r,i} \in \mathcal{X}_r$; chaque destination k décode son propre message, $\hat{M}_k$, à partir du signal reçu $y_k^n \in \mathcal{Y}_k^n$ et la séquence d'état s^n, pour $k = 1, 2$. Les codeurs et décodeurs sont définis comme suit.

- L'encodeur de la source: $\varphi_s : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}^n$;
- L'encodeur du rélais: $\{\varphi_{r,i} : \mathcal{Y}_r^{i-1} \times \mathcal{S}^{i-1} \to \mathcal{X}_r\}_{i=1}^n$;
- L'décodeur k: $\phi_k : \mathcal{Y}_k^n \times \mathcal{S}^n \to \mathcal{M}_k$.

On dit qu'une paire de débit (R_1, R_2) est réalisable si il existe une séries de tel fonctions d'encodage et d'décodage telle que le probabilité maximal d'erreur de décoder l'un ou l'autre message arrive à zéro quand $n \to \infty$, comme définis dans la Section B.2.

B.5.1 *Le schéma de Decode-Compress-Forward (DCF)*

Nous présentons maintenant une région de débit obtenue par le schéma de DCF proposé. Les détails du schéma sont plutôt techniques et sont fournis à l'annexe de la chapitre 5. L'idée essentielle est de laisser le relais transmettre certaines informations sécondaires utiles qui sont basées sur les messages de la source décodés et les retours d'état des destinations, de telle manière que l'interférence puisse être atténuée de manière efficace. En particulier, la source envoie d'abord le signal codé X, tandis que le relais décode les deux messages (les indices de message lié à U_1 et U_2) de l'observation Y_r comme dans le schéma de DF. Le décodage des deux messages est nécessaire pourvu que dans le schéma DCF, le relais prend partiellement la responsabilité de la source (multidiffuser

[11] La spécification de la cardinalité de l'ensemble d'entrée du canal doit être adapté aux canaux distincts, par exemple, infini dans le canal analogique, binaire dans le canal binaire d'effacement/symétrique, etc.

l'information sécondaire comme dans les schémas de la Partie I) qui a un accès à toutes les messages. Ensuite, le relais comprime une fonction des deux messages de la source décodés et le retour d'état des destinations. Enfin, le relais code l'index de compression (associé à $\hat{Y}$) avec son propre codebook et transmet le signal codé X_r aux deux utilisateurs comme dans le schéma de CF. Chaque destination récupère le message souhaité conjointement avec l'indice de compression. Comme dans les schémas DF et CF, la source et le relais transmettent simultanément avec le codage de Markov de bloc. Une région de débit réalisable est présentée ci-dessous.

Theorem B.9. (Bornes Inférieure DCF) *Une région réalisable pour un BRC sans mémoire dépendant sur l'état avec voie de retour d'état au rélais, est défini par l'ensemble de toutes les paires de débit (R_1, R_2) satisfant*

$$R_1 \leq \min\Big\{I_{21}, I_{31}, I_{41} + I_{51}\Big\}, \tag{169a}$$

$$R_2 \leq \min\Big\{I_{22}, I_{32}, I_{42} + I_{52}\Big\}, \tag{169b}$$

$$R_1 + R_2 \leq I_1, \tag{169c}$$

où, pour $k \neq m \in \{1, 2\}$,

$$I_1 \triangleq I(U_1U_2; Y_r|X_rSQ), \tag{170a}$$

$$I_{2k} \triangleq I(U_k; Y_r|U_mX_rSQ), \tag{170b}$$

$$I_{3k} \triangleq I(U_k; Y_k\hat{Y}|X_rSQ), \tag{170c}$$

$$I_{4k} \triangleq I(U_k; Y_k|SQ), \tag{170d}$$

$$I_{5k} \triangleq -I(\hat{Y}; U_m|U_kX_rY_kSQ) + I(X_r; Y_k|U_kSQ), \tag{170e}$$

pour tout les pmf's qui verifient

$$\begin{aligned} p(y_1, y_2, y_r, \hat{y}, x, x_r, u_1, u_2, s, q) = \\ p(\hat{y}|x_r, u_1, u_2, s, q)p(y_1, y_2, y_r|x, x_r, s) \\ \times p(x_r|q)p(x|u_1, u_2, q)p(u_1)p(u_2)p(s)p(q). \end{aligned} \tag{171}$$

Démonstration. Sans entrer dans les détails, les contraintes de débit peuvent être interprétées comme suit. Tout d'abord, les contraintes $R_1 \leq I_{21}$, $R_2 \leq I_{22}$ et $R_1 + R_2 \leq I_1$ sont les conditions requises pour que le relais récupère les deux messages de manière fiable. Ensuite, $R_k \leq I_{3k}$ est la condition pour assurer que la destination k décode le message désiré à l'aide des informations secondaires utiles, pourvu que cet information soit décodé de manière fiable à la destination k. Enfin, la condition $R_k \leq I_{4k} + I_{5k}$ peut garantir une transmission réussi d'envoyer simultanée des informations secondaires et le message original à destination k (récupéré simplement de son observation). Les détails de la réalisabilité sont fournis à l'annexe 5.A pour une meilleure lisibilité. ■

B.6 CONCLUSIONS ET QUELQUES COMMENTAIRES

Dans cette thèse, nous avons étudié l'une des approches populaires de voie de retour qui permette l'atténuation des interférences pour le problème de communication descendante, c'est-à-dire, RIA autorisé par la voie de retour de l'état. En particulier, dans la partie I de la thèse nous nous sommes concentrés sur la voie de retour d'état homogène, tandis que dans la partie II de la thèse nous avons exploré le CSIT retardé quand il est hétérogène mais dans deux modèles différentes: le canal de radiodiffusion et le canal de radiodiffusion avec un rélais. L'étude a été menée avec des outils de théorie information.

B.6.1 *Radiodiffusion avec la voie de retour homogène d'état*

La première partie de la thèse a porté sur le CSIT retardé homogène dans le canal de diffusion avec multi-utilisateurs. Bien que les deux chapitres de cette partie traitent du problème similaire, ils ont une accent distincte, c'est-à-dire des bornes inférieures et supérieures dans le cas de deux-utilisateur (Chapitre 2) et des bornes inférieures relativement simples pour les canaux avec $K-$utilisateur (Chapitre 3).

Dans le modèle de deux-utilisateur, nous avons tenté de caractériser le gain d'avoir la voie de retour d'état avec SNR fini car son avantage remarquable avec SNR infini a depuis longtemps été prouvé dans [2, 12, 13, 19]. Dans l'effort de l'évaluer, nous avons dérivé une borne inférieure en utilisant les outils connus du codage Markov de bloc[12] et Wyner-Ziv compression [61, 73, 74]. La stratégie de transmission dépend de levier des informations entendus précédemment, c'est-à-dire, reconstruire et encoder les observations mal placées, puis multidiffuse une version de l'information quantifiée. De plus, une borne supérieure standard est également fournie comme moyen de comparaison et est ensuite serrée dans le cas Gaussien par une inégalité extrême et quelques manipulations algébriques. Plus tard, nous montrons que le débit symétrique optimisé numériquement surpasse celui-là de la stratégie conventionnelle. Néanmoins, même avec l'optimiseur numérique, il n'est pas efficace de choisir un ensemble de paramètres utiles $(P_k,\ N)$ dans la région proposée, et donc ce fait stimule le travail du Chapitre 3 comme une solution partielle.

Dans le modèle de $K-$utilisateur, nous cherchons non seulement à trouver une méthode de codage pour endosser le gain avec SNR fini dans le canal Gaussien de $K-$utilisateur, mais aussi pour éliminer la contrainte sur le taille de l'alphabet d'entrée d'canal dans EBC. Dans cette optique, nous avons exclu les techniques de codage Markov de bloc et de binning et avons présentés un codage distribué avec perte[13] et donc le schéma présenté dans le chapitre 3 est général et simple. Par conséquence, la région de débit obtenue qui est docile a été démontrée d'obtenir une performance supérieure à celles des schémas liés à TDMA et MAT pour au-delà de deux utilisateurs GBC et d'atteindre les capacité de certains sous-classes de EBC sans constraintes sur le taille d'entrée de canal, simultanément. Comme indiqué précédemment, grâce à la structure de division de temps, nous

[12] block Markov code
[13] distributed lossy coding

convertissons le problème d'optimisation d'allocation de puissance où les P's sont dans l'expectation au un problème d'optimisation d'allocation de plage horaire qui est résoluble par analyse suivant par l'optimisation de la variation de bruit de compression par simulation. En résumé, l'efficacité de CSIT retardé avec SNR fini est justifiée dans ces deux chapitres.

B.6.2 *Radiodiffusion avec la voie de retour hétérogène d'état*

Dans cette partie de la thèse, nous nous concentrons sur les sujets de voie de retour d'état à l'hétérogénéité.

Voie de retour d'état des récepteur(s) partiel

Dans le chapitre 4, nous avons considéré pour la plupart le scénario où un seul récepteur retourne l'état. Dans ce modèle, nous avons adapté les schémas préparés dans le chapitre 2, 3 et avons dérivé les régions correspondantes. Il convient de souligner que nos schémas adaptés, par opposition à certaines méthodes linéaires (par exemple, MAT [2], paquets combinant [27, 28, 43]) dans la littérature qui exploitent l'état obsolète, sont des codages non-linéaire par eux-mêmes. Bien qu'il soit incertain et non trivial de déterminer si un schéma non-linéaire peut avoir en général une performance supérieure aux méthodes linéaires existantes, à la fois les évaluations analytique et numériques ont fourni une réponse positive dans EBC avec deux-utilisateur. De plus, nous avons révélé où cet avantage est venu. Plus précisément, nos schémas adaptés fournissent plus des bits décodables qui sont utiles à l'utilisateur qui n'a pas la voie de retour que les codes linéaires (XOR), et donc allouent plus de flux à l'utilisateur sans voie de retour. Remarquablement, une contrainte des régions par le codage non-linéaire et celle de la borne supérieure s'effondrent, qui semble impliquer le relâchement d'une borne supérieure qui est spécial pour le cas symétrique. De plus, nous avons spécifié une région au cas de trois-utilisateur et avons démontré que l'incorporation d'un utilisateur sans retourner son état dans le EBC avec deux-utilisateur est en effet bénéfique pour certaines classes de EBC.

Voie de retour d'état au relais dans un BRC

Dans le chapitre 5, nous avons étudié un modèle particulier dans lequel un relais d'infrastructure est ajouté dans le BC. Dans ce modèle, une restriction plus réaliste est imposée à la liaison de voie de retour que l'état obsolète est disponible uniquement au relais. Nous obtenons d'abord dans le BRC général avec deux-utilisateur une borne inférieure dérivée de la schéma Décoder-Comprimer-et-Transmettre[14]. Les opérations de décodage et de compression dans les stratégies pour les canaux relais classiques (DF, CF) sont fusionnées dans le schéma de DCF de sorte que le relais décode tous les messages, les encode avec l'état obtenu via les voies de retour, puis comprime les signaux codés pour les adapter à la

[14]Decode-Compress-and-Forward

multidiffusion. Ensuite, dans le BRC Gaussien avec fading, selon la façon dont le message de la source est transmis et relayé, la borne inférieure est spécifiée à deux variantes, TD-DCF et SE-DCF. Après avoir vérifié que toutes les deux variantes sont optimales en terme de DoF, nous avons ensuite évalué les régions de débit correspondantes dans le cas de SNR fini. Les résultats numériques confirment le gain de performance de l'utilisation du relais en présence de la voie de retour hétérogène d'état (des destinations au relais). L'amélioration est en fait dans deux aspects: 1) un gain remarquable par rapport au débit du cas sans avoir les voies de retour dans BRC, dans lequelle nous avons utilsés les schémas adapté de DF et CF, est observé; 2) une amélioration par rapport au débit du cas sans avoir un rélais mais avoir le CSIT homogène à la source est prouvée implicitement en comparant le débit d'ici avec les débits obtenus dans le chapitre 2, 3. Il est également révélé que le TD-DCF apporte une meilleure performance que le SE-DCF dans le scénario étudié.

Dans le RC traditionnel, soit le schéma CF soit DF est choisi en fonction des qualités des connexions. Par exemple, la stratégie CF surpasse en général le schéma de DF lorsque le lien SRC est plus faible que SDC ou si le récepteur a une connexion suffisamment forte avec le relais. Pourvu que le mélange des deux stratégies est souvent considéré comme inefficace parceque la contrainte de vitesse sur le lien SRC libérée par l'application de la compression sur l'observation du relais restera si nous imposons le relais à décoder le message. Néanmoins, le décodage et la compression ensemble peuvent être bénefique dans un réseau avec relais s'il y a plusieurs utilisateurs. Du point de vue de l'attenuation d'interférence, le décodage du message est pratiquement une condition préalable à la reconstruction des interférences (sorties entendus par hasard) dans un canal dépendant de l'état tandis que la compression du signal encodé faciliter la multidiffusion des informations secondaires.

Malgré la distinction sur la façon dont les stratégies de transmission sont réalisées, nous ne pouvons pas ignorer la ressemblance entre les choix RVs de certains des schémas proposés dans cette thèse. En particulier, le schéma de Markov de bloc qui utilise une structure de superposition ressemble au mode parallèle du schéma DCF alors que la méthode de compression distribuée nous rappelle le mode séquentiel du schéma DCF. Plus précisément, à la fois le mode parallèle de SE-DCF et le schéma de Markov de bloc de structure superposée permettent au(aux) émetteur(s) d'envoyer simultanément des signaux relatifs aux messages privés et aux messages auxiliaires communs. De même, le mode séquentiel de TD-DCF partage exactement le même choix sur RVs avec le schema DC si nous substituons $V_{1\to\{12\}}$ pour X_r comme tous les deux schémas allouent des intervalles de temps différents pour les messages aux objets différents telle que la multidiffusion et l'unidiffusion[15] ne s'interfèrent pas entre elles. Cependant, les RVs préférables ne sont pas identiques. Pour un BC avec deux-utilisateur, le schéma de temps-partagé semble être plus favorable alors que le meilleur mode dans BRC varie selon la condition de canal, par exemple, la distance entre les noeuds de communication.

[15]unicast

Bibliography

[1] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2845–2866, 2010.

[2] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4418–4431, 2012.

[3] G. Dueck, "Partial feedback for two-way and broadcast channels," *Information and Control*, vol. 46, no. 1, pp. 1–15, 1980.

[4] A. E. Gamal, "The feedback capacity of degraded broadcast channels (Corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 379–381, 1978.

[5] C. S. Vaze and M. K. Varanasi, "The Degrees of Freedom region and interference alignment for the MIMO interference channel with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4396–4417, 2012.

[6] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the Degrees of Freedom of K-user SISO interference and X channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6542–6561, 2013.

[7] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The Degrees of Freedom region of temporally-correlated MIMO networks with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 494 – 514, 2014.

[8] J. Xu, J. G. Andrews, and S. A. Jafar, "MISO broadcast channels with delayed finite-rate feedback: Predict or observe?" *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1456–1467, 2012.

[9] P. de Kerret, X. Yi, and D. Gesbert, "On the Degrees of Freedom of the K-user time correlated broadcast channel with delayed CSIT," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July. 2013, pp. 624–628.

[10] J. Chen and P. Elia, "Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8336–8356, 2013.

[11] R. Tandon, S. A. Jafar, S. Shamai Shitz, and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO BC," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4106–4128, 2013.

[12] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the Degrees of Freedom of three-user MIMO broadcast channel with delayed CSIT," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July. 2011, pp. 209–213.

[13] C. S. Vaze and M. K. Varanasi, "The Degrees of Freedom regions of two-user and certain three-user MIMO broadcast channels with delayed CSIT," *arXiv preprint arXiv:1101.0306*, Dec. 2011.

[14] S. Yang, M. Kobayashi, P. Piantanida, and S. Shamai, "Secrecy Degrees of Freedom of MIMO broadcast channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5244–5256, 2013.

[15] Z. Wang, M. Xiao, C. Wang, and M. Skoglund, "Degrees of Freedom of multi-hop MIMO broadcast networks with delayed CSIT," *IEEE Wireless Communications Letters*, vol. 2, no. 2, pp. 1–4, 2013.

[16] ——, "Degrees of Freedom of two-hop MISO broadcast networks with mixed CSIT," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6982–6995, 2014.

[17] C. He, S. Yang, and P. Piantanida, "An achievable rate region of broadcast relay channel with state feedback," in *IEEE International Conference on Communications (ICC)*, June. 2015, pp. 4193–4198.

[18] M. A. M. Ali and A. S. Avestimehr, "Approximate capacity region of the MISO broadcast channels with delayed CSIT," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2913 – 2924, 2016.

[19] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 228–240, 2012.

[20] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1329–1345, 2013.

[21] Y. Wu and M. Wigger, "Coding schemes with rate-limited feedback that improve over the no feedback capacity for a large class of broadcast channels," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 2009–2033, 2016.

[22] R. Venkataramanan and S. Pradhan, "An achievable rate region for the broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6175–6191, 2013.

[23] X. Yi and D. Gesbert, "Precoding methods for the MISO broadcast channel with delayed CSIT," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1–11, 2013.

[24] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Approximate capacity of the two-user MISO broadcast channel with delayed CSIT," in *51st Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2013, pp. 1136–1143.

[25] J. Wang, M. Matthaiou, S. Jin, and X. Gao, "Precoder design for multiuser MISO systems exploiting statistical and outdated CSIT," *IEEE Transactions on Communications*, vol. 61, no. 11, pp. 4551–4564, 2013.

[26] B. Clerckx and D. Gesbert, "Space-time encoded MISO broadcast channel with outdated CSIT: An error rate and diversity performance analysis," *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1661–1675, 2015.

[27] C.-C. Wang, "On the capacity of 1-to-K broadcast packet erasure channels with channel output feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 931–956, 2012.

[28] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser broadcast erasure channel with feedback-capacity and algorithms," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5779–5804, 2013.

[29] S. Yang and M. Kobayashi, "Secure communication in K-user multi-antenna broadcast channel with state feedback," in *IEEE International Symposium on Information Theory (ISIT)*, June. 2015, pp. 1976–1980.

[30] L. Czap, V. M. Prabhakaran, C. Fragouli, and S. N. Diggavi, "Secret communication over broadcast erasure channels with state-feedback," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4788–4808, 2015.

[31] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of Freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 315–328, 2013.

[32] K. Mohanty, C. S. Vaze, and M. K. Varanasi, "The Degrees of Freedom region of the MIMO interference channel with hybrid CSIT," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 1837–1848, 2015.

[33] S. Amuru, R. Tandon, and S. Shamai, "On the Degrees-of-Freedom of the 3-user MISO broadcast channel with hybrid CSIT," in *IEEE International Symposium on Information Theory (ISIT)*, July. 2014, pp. 2137–2141.

[34] S. Lashgari, R. Tandon, and A. S. Avestimehr, "A general outer bound for MISO broadcast channel with heterogeneous CSIT," in *IEEE International Symposium on Information Theory (ISIT)*, June. 2015, pp. 401–405.

[35] S. Lashgari, R. Tandon, and S. Avestimehr, "Three-user MISO broadcast channel: how much can CSIT heterogeneity help?" in *IEEE International Conference on Communications (ICC)*, June. 2015, pp. 4187–4192.

[36] ——, "MISO broadcast channel with hybrid CSIT: beyond two users," *arXiv preprint arXiv:1504.04615*, 2015.

[37] B. Rassouli, C. Hao, and B. Clerckx, "DoF analysis of the MIMO broadcast channel with alternating/hybrid CSIT," *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1312–1325, 2016.

[38] C. He, S. Yang, and P. Piantanida, "On the capacity of the fading broadcast channel with state feedback," in *IEEE International Symposium on Communications, Control and Signal Processing (ISCCSP)*, May. 2014.

[39] ——, "A non-linear transmission scheme for the *K*-user broadcast channel with state feedback," in *International Symposium on Information Theory and Its Applications (ISITA), to be presented*, Nov. 2016.

[40] ——, "A non-linear transmission strategy for the *K*-user BC with delayed CSIT," *submitted to IEEE Transactions on Wireless Communications*, 2016.

[41] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of Degrees of Freedom under finite precision CSIT," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5603–5618, 2016.

[42] R. Tandon, M. A. Maddah-Ali, A. Tulino, H. V. Poor, and S. Shamai, "On fading broadcast channels with partial channel state information at the transmitter," in *IEEE International Symposium on Wireless Communication Systems (ISWCS)*, 2012, pp. 1004–1008.

[43] S.-C. Lin and I.-H. Wang, "Single-user CSIT can be quite useful for state-dependent broadcast channels," in *IEEE International Symposium on Information Theory (ISIT)*, July. 2016, pp. 160–164.

[44] ——, "Bursty broadcast channels with hybrid CSIT," *http://www-o.ntust.edu.tw/ sclin/paper/sc16Mixed.pdf*.

[45] C. Hao, B. Rassouli, and B. Clerckx, "Degrees-of-Freedom region of MISO-OFDMA broadcast channel with imperfect CSIT," *arXiv preprint arXiv:1310.6669*, 2013.

[46] C. Hao and B. Clerckx, "MISO broadcast channel with imperfect and (un)matched CSIT in the frequency domain: DoF region and transmission strategies," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2013, pp. 1–6.

[47] M. A. Maddah-Ali, "On the degrees of freedom of the compound miso broadcast channels with finite states," in *2010 IEEE International Symposium on Information Theory*. IEEE, 2010, pp. 2273–2277.

[48] T. Gou, S. A. Jafar, and C. Wang, "On the Degrees of Freedom of finite state compound wireless networks," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3286–3308, 2011.

[49] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.

[50] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback (Corresp.)," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 667–671, 1984.

[51] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[52] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.

[53] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On Degrees of Freedom region of MIMO networks without channel state information at transmitters," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 849–857, 2012.

[54] H. Kim, Y.-K. Chia, and A. El Gamal, "A note on the broadcast channel with stale state information at the transmitter," *IEEE Transactions on Information Theory*, vol. 61, no. 7, pp. 3622–3631, 2015.

[55] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback-capacity and algorithms," in *IEEE Workshop on Network Coding, Theory, and Applications*, 2009, pp. 54–61.

[56] E. C. Van Der Meulen, "Three-terminal communication channels," *Advances in applied Probability*, vol. 3, pp. 120–154, 1971.

[57] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, 1979.

[58] M. N. Khormuji, A. El Gamal, and M. Skoglund, "State-dependent relay channel: Achievable rate and capacity of a semideterministic class," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2629–2638, 2013.

[59] I. Estella Aguerri and D. Gunduz, "Capacity of a class of state-dependent relay channels," *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1280 – 1295, 2016.

[60] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, 2005.

[61] A. Behboodi and P. Piantanida, "Cooperative strategies for simultaneous and broadcast relay channels," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1417–1443, 2013.

[62] L. Chen, "Parallel relay broadcast channels," in *IEEE International Symposium on Information Theory (ISIT)*, 2010, pp. 585–589.

[63] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 900–928, 2007.

[64] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3517–3535, 2007.

[65] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3132–3152, 2011.

[66] J. Han and C.-C. Wang, "Linear network coding capacity region of the smart repeater with broadcast erasure channels," in *IEEE International Symposium on Information Theory (ISIT)*, July. 2016.

[67] J. Chen, S. Yang, A. Özgür, and A. Goldsmith, "Outdated CSIT can achieve full DoF in heterogeneous parallel channels," in *IEEE International Symposium on Information Theory (ISIT)*, 2014, pp. 2564–2568.

[68] G. Atia and A. F. Molisch, "Cooperative relaying with imperfect channel state information," in *IEEE Global Telecommunications Proceedings (GLOBECOM)*, 2008, pp. 1–6.

[69] Y. Cui, V. K. Lau, and R. Wang, "Distributive subband allocation, power and rate control for relay-assisted OFDMA cellular system with imperfect system state knowledge," *IEEE Transactions on Wireless Communication*, vol. 8, no. 10, pp. 5096–5102, 2009.

[70] W. Xu, X. Dong, and W.-S. Lu, "MIMO relaying broadcast channels with linear precoding and quantized channel state information feedback," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5233–5245, 2010.

[71] C. He, S. Yang, and P. Piantanida, "Exploiting heterogeneous delayed CSIT in broadcast channel: linear or non-linear coding?" *in preparation*.

[72] ——, "On the Gaussian fading broadcast relay channel with causal state feedback," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2797–2807, 2016.

[73] A. Behboodi and P. Piantanida, "Mixed noisy network coding and cooperative unicasting in wireless networks," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 189–222, 2015.

[74] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[75] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2006.

[76] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5011–5023, 2009.

[77] J. Jeon, "A generalized typicality for abstract alphabets," in *Information Theory (ISIT), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 2649–2653.

[78] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1469–1482, 2006.

[79] C. S. Vaze and M. K. Varanasi, "The Degree-of-Freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5354–5374, 2012.

[80] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," *IEEE/ACM Transactions on Networking (ToN)*, vol. 16, no. 3, pp. 497–510, 2008.

[81] A. Tulino, S. Verdú, G. Caire, and S. Shamai, "The Gaussian erasure channel," in *IEEE International Symposium on Information Theory (ISIT)*, 2007, pp. 1721–1725.

[82] A. Ghorbel, M. Kobayashi, and S. Yang, "Content delivery in erasure broadcast channels with cache and feedback," *accepted* in *IEEE Transactions on Information Theory*, 2016.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Radiodiffusion avec CSIT retardée: analyse de SNR fini et voie de retour hétérogène

Mots clefs : canal diffusion, canal relai, CSIT retardée, voie de retour hétérogène, SNR fini

Résumé : Cette thèse explore, sous certains paramètres réalistes, l'une des techniques clés pour les réseaux sans fil de demain, i.e., la réduction des interférences permis par la voie de retour (feedback). Nous nous concentrons sur la voie de retour du type d'état, également connu sous le nom de CSIT retardé, qui aide les récepteurs à profiter des observations indésirables par créer des dimensions de signaux supplémentaires. Afin de vérifier l'utilité de la CSIT retardé dans des situations sévères, nous l'étudions avec SNR fini et / ou avec la hétérogénéité de la voie de retour dans une configuration de communication de diffusion, qui est largement utilisé pour modéliser la transmission de liaison descendante dans les systèmes cellulaires. Tout au long de la thèse, nous utilisons des outils de théorie information, par exemple, le codage lossy distribué, bloc Markov codage, la technique de compression (Wyner-Ziv), etc.

Dans la première partie de cette thèse, nous sommes surtout intéressés par la performance de CSIT retardée avec SNR fini et l'uniformité à travers les résultats dans le canal Gaussien et dans le canal d'effacement. Plusieurs schémas relativement simples sont proposés pour des canaux de diffusion multi-utilisateur (sans mémoire) dans le cas où les états sont supposés être entièrement connus à la destination, mais causalement à l'émetteur. Lors de l'analyse des régions correspondantes les cas Gaussien/ d'effacement, nous avons caractérisé des améliorations en termes de débits symétriques plus élevés et plus d'options de l'alphabet d'entré.

Dans la deuxième partie de cette thèse, les algorithmes adaptés aux hétérogénéités différentes de la voie de retour sont ciblés, c'est-à-dire, seulement une partie des nœuds de communication sont impliqués dans le processus de feedback. En particulier, nous nous concentrons sur le canal de 1) diffusion supporté par les voies de retour des récepteurs partiels, 2) diffusion avec relais et voie de retour au relais. Étant donné que (tous /partiel) états retardés sont accessibles à des émetteurs (tous/partiels), les approches proposées, bien qu'ils emploient les méthodes de codages visant à réduire les interférences à tous les récepteurs, doivent soit donner la priorité aux utilisateurs qui fournissent les états et compter sur un gain de codage opportuniste pour les autres, soit forcer chaque source à prendre la responsabilité partielle de transmission. Les améliorations sur les débits réalisables sont justifiées dans des cas avec analyse et quelques exemples.

Les résultats et les évaluations de cette thèse, qui donnent quelques indications sur comment le retour d'état peut être exploité dans la transmission de liaison descendante, montrent qu'une bonne performance de débit pourrait être atteinte avec le CSIT retardé même lorsque la puissance de transmission est limitée et lorsque le retour d'état est disponible de façon hétérogène.

Title : Broadcasting with delayed CSIT: finite SNR analysis and heterogeneous feedback

Keywords : broadcast channel, relay channel, delayed CSIT, heterogeneous feedback, finite SNR

Abstract : This dissertation explores one of the key techniques for future wireless networks, namely feedback enabled interference mitigation, under some realistic settings. We focus on the state-type feedback, also known as delayed CSIT, which helps leverage receivers' overheard observations to create extra signal dimensions. In order to verify the usefulness of delayed CSIT in harsh situations, we investigate it with finite SNR and/or feedback heterogeneity in a broadcast communication setup, which is widely utilized to model downlink transmission in cellular systems. Throughout the thesis, we use some information theoretical tools, e.g., distributed lossy source coding, block Markov coding, Wyner-Ziv compression technique, e.t.c.

In the first part of this dissertation, we are mainly interested in the finite SNR performance of delayed CSIT and the uniformity across the results in the Gaussian broadcast channel and in the erasure channel. Several relatively simple schemes are proposed in multi-user memoryless broadcast channels when states are assumed to be fully known at the destinations but only strictly causally at the transmitter. Enhancements in terms of higher symmetric rates and more input alphabet options are then characterized when analyzing the corresponding regions in Gaussian/erasure cases.

In the second part of this dissertation, algorithms adapted to distinct feedback heterogeneities are targeted as only part of the communication nodes are involved in the feedback process. In particular, we concentrate on 1) broadcast channel with feedback from part of all receivers; and 2) broadcast relay channel with feedback at the relay. Given that (partial) delayed states are accessible at (partial) transmitters, the proposed approaches, though employ coding methods aiming at mitigating interference at all receivers, have to either give priority to the users who feedback and rely on opportunistic coding gain for the others or force each source to take limited responsibility in the transmission. Improvements on achievable rates are justified in either cases with analysis and some examples.

The results and their evaluations in this thesis, which give some insights on how to exploit the state feedback in downlink transmission, show that good rate performance can be achieved with delayed CSIT even when transmission power is limited and when the state feedback is heterogeneous.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France