Feedback and Partial Message Side-Information on the Semideterministic Broadcast Channel

Annina Bracher and Michèle Wigger, *Senior Member, IEEE*

Abstract— The capacity of the semideterministic discrete memoryless broadcast channel (SD-BC) with partial message side-information (P-MSI) at the receivers is established. In the setting without a common message, it is shown that P-MSI to the stochastic receiver alone can increase capacity, whereas P-MSI to the deterministic receiver can only increase capacity if also the stochastic receiver has P-MSI. The latter holds only for the setting without a common message: if the encoder also conveys a common message, then P-MSI to the deterministic receiver alone can increase capacity. These capacity results are used to show that feedback from the stochastic receiver can increase the capacity of the SD-BC without P-MSI and the sum-rate capacity of the SD-BC with P-MSI at the deterministic receiver. The link between P-MSI and feedback is a feedback code, which—roughly speaking—turns feedback into P-MSI at the stochastic receiver, and hence helps the stochastic receiver mitigate experienced interference. For the case, where the stochastic receiver has full MSI (F-MSI) and can thus fully mitigate experienced interference also in the absence of feedback, it is shown that feedback cannot increase capacity.

Index Terms— Channel capacity, feedback, partial message side-information, semideterministic broadcast channel.

I. INTRODUCTION

WE DERIVE the capacity region of the semideterministic discrete memoryless broadcast channel (SD-BC) with *partial message side-information (P-MSI)* (Theorem 2). In this setting each receiver knows part of the message intended for the other receiver already before the transmission begins. Our capacity result generalizes that of [1], [2] for the SD-BC without MSI. The capacity region of the general BC with *full MSI (F-MSI)*, where each receiver knows the entire message intended for the other receiver, was established in [3] and [4]. The work of Kramer and Shamai [4] also considers P-MSI and establishes the capacity region of the BC with P-MSI and degraded message sets. The three-receiver BC with P-MSI is studied in [5]. Independently of our work, Asadi, Ong, and Johnson proposed a coding scheme for general two-receiver BCs with P-MSI [6]. One can show that— for a judicious choice of the auxiliary random variables—their

scheme achieves the capacity region of the SD-BC. Their work does not, however, provide a converse.[1]

Generally speaking, P-MSI reduces the effect of self-interference on the BC and hence enables more efficient communication (see, e.g., [4]). More specifically, in the current paper we show that on the SD-BC P-MSI affects the capacity region as follows:

- P-MSI at the deterministic receiver can increase capacity if, and only if, one of the following two holds: 1) also the stochastic receiver has P-MSI; or 2) the encoder conveys also a common message (Remark 7).
- P-MSI at the stochastic receiver can increase capacity (Remark 8); and this holds irrespective of whether or not the deterministic receiver has P-MSI or the encoder conveys a common message.

To establish these findings we use our capacity result for the SD-BC with P-MSI. Of particular interest to us is the latter finding, which we shall use to design a feedback code for the SD-BC with or without P-MSI that can improve over the channel's no-feedback capacity.

Feedback on the BC was first studied in [10], where it is shown that even perfect feedback does not increase the capacity region of the physically-degraded BC [10]. It was later proved that feedback can, however, increase the capacity region of several BCs that are not physically degraded [11]–[16]; and achievable rate regions for the BC with feedback were established in [12]–[16]. An intuition for the gain due to feedback is that feedback allows the transmitter to create a common message that is useful to both receivers [11], [14]. Typically, transmitting one common message is more efficient than transmitting two private messages, because in the latter case the transmissions of the two private messages intefere with each other. In prominent previous examples where feedback increases the BC's capacity— e.g., in Dueck's example [11]—the common message is built up of past noise symbols. It is not clear how the idea of constructing a common message using past noise symbols should be adapted to the SD-BC, which is the focus of this paper; one receiver of the SD-BC is deterministic, and hence it is not clear why information that is constructed

[1]Somewhat similar to P-MSI is decoder cooperation on the BC, which allows the decoders to exchange information via finite-capacity links. The capacity region of the SD-BC with one-sided cooperation via a link from the deterministic to the stochastic receiver is established in [7]. In fact, as it is shown in [7], this network is operationally equivalent to a class of relay-broadcast channels whose capacity region is established in [8]. The physically-degraded BC with parallel conferencing and the BC with conferencing and degraded message sets are studied in [9].

only from previous noise symbols should be useful to this receiver.

In the current paper we show that—notwithstanding the above observations—feedback can increase the capacity region of the SD-BC (Theorem 17). To establish this result, we use the feedback to create an improved situation where the stochastic receiver has P-MSI. (As mentioned before, P-MSI at the stochastic receiver can increase the SD-BC's capacity.) More precisely, we use the idea that the encoder can create from the feedback a new message that is useful to the deterministic receiver and can be created (and hence is known) at the stochastic receiver. A similar idea was previously used by Wu and Wigger to construct a coding scheme for the general BC with rate-limited feedback [16], though their work does not make the connection to the BC with P-MSI explicit. They use the scheme to show that feedback can increase the capacity of a large class of stochastically- (but not physically-) degraded BCs as well as the capacity of a class of BCs that consist of a binary symmetric channel and a binary erasure channel [16]. The argument is particularly intuitive for the class of stochastically-degraded BCs that satisfy that one receiver is stronger than the other: it is shown in [16] that for any BC in this class the encoder can create a new message that is useful to the stronger receiver and can be created at the weaker receiver, and that the encoder can send this message without reducing the rates at which the fresh message-information is transmitted.

Unlike the above class of stochastically-degraded BCs, on the SD-BC there is a tradeoff between the rates at which fresh message-information and the message that the encoder constructs from the feedback are sent. Hence—even with the results of [16] at hand—showing that feedback can increase the capacity region of the SD-BC is nontrivial. We show by means of an example that—with a judicious choice of the rates at which the fresh and the feedback information are sent—we can increase the overall rates at which the messages are sent to the receivers (Example 18). From this we conclude that feedback can increase the capacity region of the SD-BC.

As already mentioned, in [16] the connection between the coding idea and the BC with P-MSI is not made explicit. We make the connection explicit, and this allows us to readily extend our feedback coding scheme for the SD-BC to the case where the receivers have P-MSI. Using this extension of the feedback code, we show that if the deterministic receiver has P-MSI, then feedback can increase the sum-rate capacity of the SD-BC (Theorem 17). For the case where the stochastic receiver has F-MSI, we show that feedback cannot increase capacity, irrespective of whether or not the deterministic receiver has P-MSI (Theorem 19).

The rest of this paper is structured as follows. We conclude this section by introducing some notation. Section II describes the channel model. Section III contains the results for the SD-BC with P-MSI, and Section IV studies the effect of feedback on the SD-BC with and without P-MSI.

A. Notation and Preliminaries

We use calligraphic letters to denote finite sets and $|\cdot|$ for their cardinality, e.g., $\mathcal{X}$ and $|\mathcal{X}|$. Random variables are denoted by upper-case letters and their realizations by lower-case letters, e.g., X and x. By X_i^j and x_i^j we denote the tuples $(X_i, \ldots, X_j)$ and $(x_i, \ldots, x_j)$, where $j > i$; and we drop the subscript $i = 1$, e.g., we write X^n instead of X_1^n. Sequences are in bold lower- or upper-case letters depending on whether they are deterministic or random, e.g., $\mathbf{x}$ denotes an n-length codeword.

By $Q \sim \mathrm{Unif}[1 : n]$ we indicate that the random variable Q is uniformly drawn from the set $\{1, \ldots, n\}$, and by $S \sim \mathrm{Ber}(p)$, where $p \in [0, 1]$, we indicate that S is a Bernoulli-p random variable. We denote the binary entropy function by $h_{\mathrm{b}}(\cdot)$ and its inverse on $[0, 1/2]$ by $h_{\mathrm{b}}^{-1}(\cdot)$.

A joint probability mass function (PMF), its marginal PMF, and its conditional PMF are all denoted by the same function $p(\cdot)$, with the exact meaning specified by the subscripts or arguments, e.g., $p_{X,Y}(0, 1)$ denotes the probability of the event $(X, Y) = (0, 1)$ and $p(x|y)$ the probability that $X = x$ given $Y = y$.

We denote the set of ϵ-typical length-n sequences defined in [17, Ch. 2] by $\mathcal{T}_\epsilon^{(n)}$. By $\delta(\epsilon)$ we denote any function of ϵ that converges to 0 as ϵ approaches 0; and $\{\epsilon_n\}$ can stand for any sequence of numbers that converges to 0 as n tends to infinity.

We shall use the following lemma, which is proved, e.g., in [18]:

Lemma 1 (Functional Representation Lemma): Given two random variables X and Y of finite support, there exist a chance variable S of finite support $\mathcal{S}$ that is independent of X and a function $g \colon \mathcal{X} \times \mathcal{S} \to \mathcal{Y}$ such that $Y = g(X, S)$.

II. CHANNEL MODEL

We consider the SD-BC of transition law

$$W(y, z|x) = \mathbb{1}_{\{y = f(x)\}} W(z|x),$$

where we assume that the channel-input alphabet $\mathcal{X}$ and the channel-output alphabets $\mathcal{Y}$ and $\mathcal{Z}$ are finite. Transmitting an n-tuple X^n, the encoder wants to convey the message-pairs $(M, M_{\mathcal{Y}})$ and $(M, M_{\mathcal{Z}})$ to the deterministic receiver $\mathcal{Y}$ and the stochastic receiver $\mathcal{Z}$, respectively, where M denotes the common message and $M_{\mathcal{Y}}$ and $M_{\mathcal{Z}}$ the private messages. We assume that M, $M_{\mathcal{Y}}$, and $M_{\mathcal{Z}}$ are independent, that M is uniformly drawn from a size-2^{nR} set, and that for each $\nu \in \{\mathcal{Y}, \mathcal{Z}\}$ message M_ν is uniformly drawn from a size-2^{nR_ν} set. We study the SD-BC with P-MSI, and we thus assume that each private message comprises two parts, i.e.,

$$M_\nu = \left(M_\nu^{(p)}, M_\nu^{(c)}\right), \quad \nu \in \{\mathcal{Y}, \mathcal{Z}\},$$

that Receiver $\mathcal{Y}$ knows $M_{\mathcal{Z}}^{(c)}$ and decodes $(M, M_{\mathcal{Y}})$ from $(Y^n, M_{\mathcal{Z}}^{(c)})$, and that Receiver $\mathcal{Z}$ knows $M_{\mathcal{Y}}^{(c)}$ and decodes $(M, M_{\mathcal{Z}})$ from $(Z^n, M_{\mathcal{Y}}^{(c)})$. For each $\nu \in \{\mathcal{Y}, \mathcal{Z}\}$ we assume that $M_\nu^{(p)}$ and $M_\nu^{(c)}$ are independent of each other and uniformly drawn from sets of size[2] $2^{nR_\nu^{(p)}}$ and $2^{nR_\nu^{(c)}}$, respectively,

---

[2]For simplicity, when we write 2^{nR}, for some $n, R \geq 0$, we implicitly assume that it is an integer value. It would be more precise to write $\lfloor 2^{nR} \rfloor$ instead. However, the ratio between the two expressions tends to 1 when $n \to \infty$, which is the regime of interest in this paper.
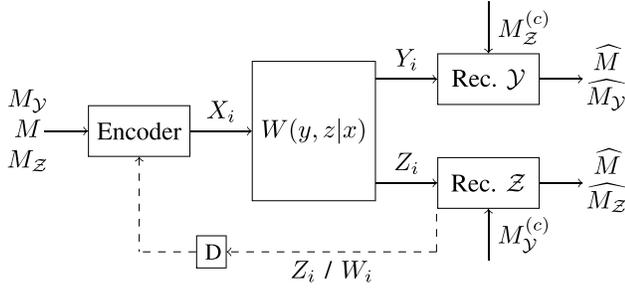
Fig. 1. SD-BC with P-MSI and feedback.

where

$$R_\mathcal{Y} = R_\mathcal{Y}^{(p)} + R_\mathcal{Y}^{(c)} \quad \text{and} \quad R_\mathcal{Z} = R_\mathcal{Z}^{(p)} + R_\mathcal{Z}^{(c)}.$$

Note the extreme cases:

- $(R_\mathcal{Y}^{(p)}, R_\mathcal{Y}^{(c)}) = (R_\mathcal{Y}, 0) \iff$ no MSI at Receiver $\mathcal{Z}$
- $(R_\mathcal{Y}^{(p)}, R_\mathcal{Y}^{(c)}) = (0, R_\mathcal{Y}) \iff$ F-MSI at Receiver $\mathcal{Z}$
- $(R_\mathcal{Z}^{(p)}, R_\mathcal{Z}^{(c)}) = (R_\mathcal{Z}, 0) \iff$ no MSI at Receiver $\mathcal{Y}$
- $(R_\mathcal{Z}^{(p)}, R_\mathcal{Z}^{(c)}) = (0, R_\mathcal{Z}) \iff$ F-MSI at Receiver $\mathcal{Y}$.

A rate-tuple $(R, R_\mathcal{Y}^{(p)}, R_\mathcal{Y}^{(c)}, R_\mathcal{Z}^{(p)}, R_\mathcal{Z}^{(c)})$ is achievable if there exists a sequence of encoders and decoders so that at each receiver the probability of a decoding error tends to zero as n tends to infinity. The capacity region is the closure of the set of all achievable rate-tuples.

We study the SD-BC with P-MSI in the absence and in the presence of feedback. In the absence of feedback, the encoder selects the channel-input sequence as a function of the triple $(M, M_\mathcal{Y}, M_\mathcal{Z})$, i.e., $X^n = \phi(M, M_\mathcal{Y}, M_\mathcal{Z})$. This setting corresponds to that of Figure 1 without the dashed links. We denote its capacity region by $\mathscr{C}_{\text{P-MSI}}$, and in the special case without MSI by $\mathscr{C}$.

When there is feedback, it is assumed to be one-sided from the stochastic receiver $\mathcal{Z}$ only. (Feedback from the deterministic receiver $\mathcal{Y}$ is useless, because the encoder can always compute Y^n from X^n.) We consider perfect and rate-limited feedback. Perfect feedback allows the encoder to form the Time-i input also as a function of Z^{i-1}, i.e.,

$$X_i = \phi_i(M, M_\mathcal{Y}, M_\mathcal{Z}, Z^{i-1}), \quad i \in [1:n].$$

Rate-limited feedback of rate R_{FB} allows Receiver $\mathcal{Z}$ to transmit after Transmission i a feedback signal $W_i = h(Z^i, M_\mathcal{Y}^{(c)}) \in \mathcal{W}_i$ to the encoder, and in turn the encoder can form the Time-i input also as a function of W^{i-1}, i.e.,

$$X_i = \phi_i(M, M_\mathcal{Y}, M_\mathcal{Z}, W^{i-1}), \quad i \in [1:n].$$

The rate-limitation implies that

$$\prod_{i=1}^{n} |\mathcal{W}_i| \leq 2^{nR_{\text{FB}}}. \tag{1}$$

The SD-BC with P-MSI and perfect feedback (rate-limited feedback) corresponds to the setting of Figure 1 when the dashed links transport the feedback signal Z_i (W_i). Note that perfect feedback is more powerful than rate-limited feedback: any rate-tuple that is achievable with rate-limited feedback

can also be achieved with perfect feedback. Rate-limited and perfect feedback are equally powerful when $R_{\text{FB}} \geq \log|\mathcal{Z}|$.

III. THE SD-BC WITH P-MSI

In this section, we assume that there is no feedback.

A. Capacity Region and Optimal Coding Scheme

Theorem 2 (Capacity With P-MSI): The capacity region $\mathscr{C}_{\text{P-MSI}}$ of the SD-BC with P-MSI is the set of rate-tuples $(R, R_\mathcal{Y}^{(p)}, R_\mathcal{Y}^{(c)}, R_\mathcal{Z}^{(p)}, R_\mathcal{Z}^{(c)})$ satisfying

$$R + R_\mathcal{Y} \leq H(Y) \tag{2a}$$
$$R + R_\mathcal{Z} \leq I(U; Z) \tag{2b}$$
$$R + R_\mathcal{Y} + R_\mathcal{Z}^{(p)} \leq I(V; Y) + H(Y|U) + I(U; Z|V) \tag{2c}$$
$$R + R_\mathcal{Y}^{(p)} + R_\mathcal{Z} \leq H(Y|U) + I(U; Z) \tag{2d}$$
$$2R + R_\mathcal{Y} + R_\mathcal{Z} \leq I(V; Y) + H(Y|U) + I(U; Z) \tag{2e}$$

for some PMF of the form

$$p(v, u, x, y, z) = p(v, u)\, p(x|u)\, W(y, z|x). \tag{3}$$

W.l.g., one can restrict X to be a function of (Y, U).

Proof: See Appendix A. □

In the following we sketch and discuss the proof of the direct part. The capacity-achieving code is described rigorously in Appendix A. Use Marton's code construction (see [17, Sec. 8.4]) to encode the "common message-tuple" $(M, M_\mathcal{Y}^{(c)}, M_\mathcal{Z}^{(c)})$ into a cloud-center V^n and the private messages $M_\mathcal{Y}^{(p)}$ and $M_\mathcal{Z}^{(p)}$ into satellites Y^n and U^n, respectively. Receiver $\mathcal{Y}$ decodes $(M, M_\mathcal{Y}^{(c)}, M_\mathcal{Z}^{(c)})$ and $M_\mathcal{Y}^{(p)}$ jointly, while taking into account its knowledge of $M_\mathcal{Z}^{(c)}$; and likewise Receiver $\mathcal{Z}$ decodes $(M, M_\mathcal{Y}^{(c)}, M_\mathcal{Z}^{(c)})$ and $M_\mathcal{Z}^{(p)}$ jointly, while taking into account its knowledge of $M_\mathcal{Y}^{(c)}$.

The tentative code can achieve all rate-tuples $(R, R_\mathcal{Y}^{(p)}, R_\mathcal{Y}^{(c)}, R_\mathcal{Z}^{(p)}, R_\mathcal{Z}^{(c)})$ that for some PMF of the form (3) satisfy (2) and

$$R_\mathcal{Y}^{(p)} \leq H(Y|V) \tag{4a}$$
$$R_\mathcal{Z}^{(p)} \leq I(U; Z|V) \tag{4b}$$
$$R_\mathcal{Y}^{(p)} + R_\mathcal{Z}^{(p)} \leq H(Y|U) + I(U; Z|V). \tag{4c}$$

Note that this achievable region differs from the capacity region $\mathscr{C}_{\text{P-MSI}}$ of the SD-BC with P-MSI in that the rates $R_\mathcal{Y}^{(p)}$ and $R_\mathcal{Z}^{(p)}$ must also satisfy (4). As we show in Appendix B, the region is—in general—strictly contained in $\mathscr{C}_{\text{P-MSI}}$.[3]

To get rid of the constraints (4) and hence achieve the entire capacity region $\mathscr{C}_{\text{P-MSI}}$, a fix is needed: the encoder must be able to convey more information about $M_\mathcal{Y}^{(p)}$ and $M_\mathcal{Z}^{(p)}$ by allowing the cloud-center V^n to depend not only on the triple $(M, M_\mathcal{Y}^{(c)}, M_\mathcal{Z}^{(c)})$ but also on $M_\mathcal{Y}^{(p)}$ and $M_\mathcal{Z}^{(p)}$. To this end we use the following code construction, which

---

[3] To show this, we shall use Corollary 5 ahead.
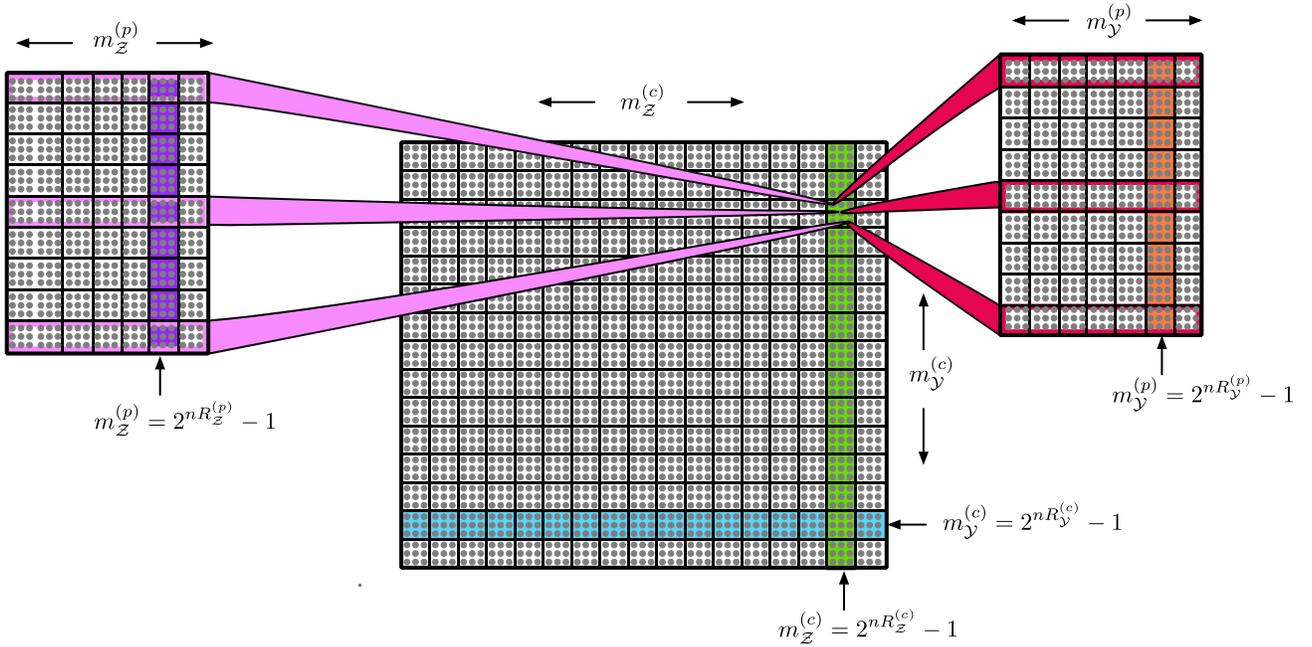
Fig. 2. Code construction without a common message.

is depicted in Figure III-A for the setting without a common message. (In Figure III-A each dot represents an n-length codeword.) Fix some PMF $p(v, u, x, y, z)$ of the form (3). For each triple $(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)})$ we generate a bin containing $2^{n\tilde{R}^{(c)}}$ n-tuples $\mathbf{v}$, which are drawn independently of each other and each from the PMF $\prod_{i=1}^{n} p(v_i)$. (In Figure III-A the light-blue row represents all bins and codewords that are associated with $m_{\mathcal{Y}}^{(c)} = 2^{nR_{\mathcal{Y}}^{(c)}} - 1$, and the light-green column represents all bins and codewords that are associated with $m_{\mathcal{Z}}^{(c)} = 2^{nR_{\mathcal{Z}}^{(c)}} - 1$.) For each cloud-center-bin we generate two satellite codebooks: one to encode $M_{\mathcal{Y}}^{(p)}$ and one to encode $M_{\mathcal{Z}}^{(p)}$. (Figure III-A depicts the satellite codebooks corresponding to the pair $(m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)}) = (3, 2^{nR_{\mathcal{Z}}^{(c)}} - 1)$: that for $M_{\mathcal{Y}}^{(p)}$ on the right, and that for $M_{\mathcal{Z}}^{(p)}$ on the left.) For each $m_{\mathcal{Y}}^{(p)}$ the first $2^{n(\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)})}$ $\mathbf{y}$-codewords in the satellite codebook corresponding to any pair $(m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)})$ are superpositioned on the first codeword in the corresponding cloud-center-bin; the following $2^{n(\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)})}$ codewords in each satellite codebook are superpositioned on the second codeword in the corresponding cloud-center-bin; and so on. That is, the first $2^{n(\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)})}$ $\mathbf{y}$-codewords are drawn according to the conditional PMF $\prod_{i=1}^{n} p(y_i | v_i^{(1)})$, where $v_i^{(1)}$ denotes the i-th component of the first codeword in the corresponding cloud-center-bin; the following $2^{n(\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)})}$ $\mathbf{y}$-codewords are drawn according to the conditional PMF $\prod_{i=1}^{n} p(y_i | v_i^{(2)})$, where $v_i^{(2)}$ denotes the i-th component of the second codeword in the corresponding cloud-center-bin; and so on. The $\mathbf{u}$-codewords in the satellite codebooks for $M_{\mathcal{Z}}^{(p)}$ are drawn similarly. (In Figure III-A the uppermost row on the right, which is framed in red, and the uppermost row on the left, which is
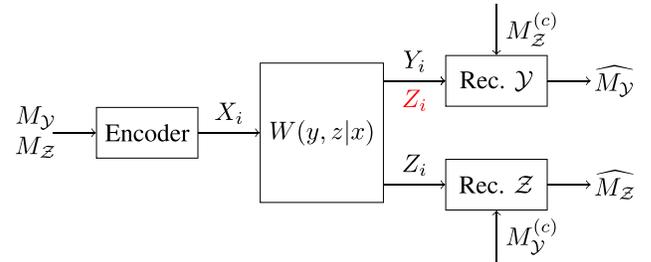
Fig. 3. Enhanced BC with P-MSI.

framed in lilac, correspond to the first codeword in the $(3, 2^{nR_{\mathcal{Z}}^{(c)}} - 1)$-cloud-center-bin.) The codewords in each satellite codebook are partitioned into as many different bins as there are possible realizations of the message $M_{\mathcal{Y}}^{(p)}$ or $M_{\mathcal{Z}}^{(p)}$, respectively, and each such bin is associated with a different realization $m_{\mathcal{Y}}^{(p)}$ or $m_{\mathcal{Z}}^{(p)}$, respectively. (In Figure 3 the orange column represents the bin that is associated with $m_{\mathcal{Y}}^{(p)} = 2^{nR_{\mathcal{Y}}^{(p)}} - 1$, and the purple column represents the bin that is associated with $m_{\mathcal{Z}}^{(p)} = 2^{nR_{\mathcal{Z}}^{(p)}} - 1$. Note that each bin comprises multiple subbins: one for each $(3, 2^{nR_{\mathcal{Z}}^{(c)}} - 1)$-cloud-center-codeword, where $(m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)}) = (3, 2^{nR_{\mathcal{Z}}^{(c)}} - 1)$ is the "common message-pair" to which the depicted satellite codebooks correspond.)

To transmit the message-tuple $(m, m_{\mathcal{Y}}^{(p)}, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(p)}, m_{\mathcal{Z}}^{(c)})$, the encoder first looks for n-tuples $\mathbf{v}$, $\mathbf{y}$, and $\mathbf{u}$ in the bins corresponding to $(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)})$, $m_{\mathcal{Y}}^{(p)}$, and $m_{\mathcal{Z}}^{(p)}$, respectively, satisfying that $(\mathbf{v}, \mathbf{y}, \mathbf{u})$ are jointly typical. It then generates the channel input X^n from the product distribution $\prod_{i=1}^{n} p(x_i | y_i, u_i)$. Receiver $\mathcal{Y}$ decodes $(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)})$ and $M_{\mathcal{Y}}^{(p)}$ jointly, while restricting attention to the column

TABLE I
SPECIAL CASES OF THEOREM 2.

		Receiver $\mathcal{Y}$		
		no MSI	P-MSI	F-MSI
Receiver $\mathcal{Z}$	no MSI	Corollary 3	Corollary 3	Corollary 3
	P-MSI	–	–	Corollary 4
	F-MSI	–	Corollary 5	Corollary 6

of the cloud-center that corresponds to the message $M_{\mathcal{Z}}^{(c)}$, which Receiver $\mathcal{Y}$ knows. Likewise, Receiver $\mathcal{Z}$ decodes $(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)})$ and $M_{\mathcal{Z}}^{(p)}$ jointly, while restricting attention to the row of the cloud-center that corresponds to the message $M_{\mathcal{Y}}^{(c)}$, which Receiver $\mathcal{Z}$ knows.[4]

As we explain in Appendix A, the effect of binning the cloud-center is the same as that of rate-splitting (see Remark 22 ahead). If we were to use rate-splitting instead of binning, then we would not bin the cloud-center, but instead we would divide the messages $M_{\mathcal{Y}}^{(p)}$ and $M_{\mathcal{Z}}^{(p)}$ into two parts each, i.e.,

$$M_{\nu}^{(p)} = (M_{\nu,s}^{(p)}, M_{\nu,c}^{(p)}), \quad \nu \in \{\mathcal{Y}, \mathcal{Z}\}.$$

Of these parts we would associate only $M_{\mathcal{Y},s}^{(p)}$ and $M_{\mathcal{Z},s}^{(p)}$ with the satellites Y^n and U^n, respectively, whereas we would encode $M_{\mathcal{Y},c}^{(p)}$ and $M_{\mathcal{Z},c}^{(p)}$ in the cloud-center. The benefit of binning the cloud-center is that it necessitates only one auxiliary rate: the rate at which the cloud-center is binned. In contrast, rate-splitting necessitates two auxiliary rates: the rates $R_{\mathcal{Y},c}^{(p)}$ and $R_{\mathcal{Z},c}^{(p)}$ associated with $M_{\mathcal{Y},c}^{(p)}$ and $M_{\mathcal{Z},c}^{(p)}$, respectively.

We next specialize Theorem 2 to cases where one or both of the receivers have no MSI or F-MSI, Table I illustrates to which.[5] These corollaries help understand when P-MSI increases capacity (see Subsection III-B).

Corollary 3 (No MSI at $\mathcal{Z}$): For each $R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$, the capacity region of the SD-BC without a common message and without MSI at Receiver $\mathcal{Z}$ ($R = 0$ and $R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$) is the set of rate-tuples $(0, R_{\mathcal{Y}}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying

$$R_{\mathcal{Y}} \leq H(Y) \tag{5a}$$

$$R_{\mathcal{Z}} \leq I(U; Z) \tag{5b}$$

$$R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq H(Y|U) + I(U; Z) \tag{5c}$$

for some PMF of the form

$$p(u, x, y, z) = p(u, x) W(y, z|x). \tag{6}$$

W.l.g., one can restrict X to be a function of (Y, U).

Proof: Let V be deterministic. In this case, and because $R = 0$ and $R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$, Constraints (2c) and (2e) are redundant in view of Constraint (2d). Hence, (5) is an inner bound on the capacity region. That (5) is also an outer bound follows from (2a), (2b), and (2d). □

---

[4]It is well-known that, without a cloud-center, this scheme achieves the capacity region of the SD-BC without a common message and without MSI (see, e.g., [17, Secs. 8.3.1–8.3.2]).

[5]Corollary 3 ahead is only for the setting without a common message ($R = 0$).

Corollary 4 (F-MSI at $\mathcal{Y}$): For each $R_{\mathcal{Y}}^{(p)} \in [0, R_{\mathcal{Y}}]$, the capacity region of the SD-BC with F-MSI at Receiver $\mathcal{Y}$ ($R_{\mathcal{Z}}^{(p)} = 0$) is the set of rate-tuples $(R, R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, 0, R_{\mathcal{Z}})$ satisfying

$$R + R_{\mathcal{Y}} \leq H(Y) \tag{7a}$$

$$R + R_{\mathcal{Z}} \leq I(U; Z) \tag{7b}$$

$$R + R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}} \leq H(Y|U) + I(U; Z) \tag{7c}$$

for some PMF of the form

$$p(u, x, y, z) = p(u, x) W(y, z|x). \tag{8}$$

Proof: For $V = U$ the constraints in (2) and (7) are equivalent, and hence (7) is an inner bound on the capacity region. That (7) is also an outer bound follows from (2a), (2b), and (2d). □

Corollary 5 (F-MSI at $\mathcal{Z}$): For each $R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$, the capacity region of the SD-BC with F-MSI at Receiver $\mathcal{Z}$ ($R_{\mathcal{Y}}^{(p)} = 0$) is the set of rate-tuples $(R, 0, R_{\mathcal{Y}}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying

$$R + R_{\mathcal{Y}} \leq H(Y) \tag{9a}$$

$$R + R_{\mathcal{Z}} \leq I(X; Z) \tag{9b}$$

$$R + R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} \leq I(X; Y, Z) \tag{9c}$$

for some PMF of the form

$$p(x, y, z) = p(x) W(y, z|x). \tag{10}$$

Proof: See Appendix C. □

From Corollary 5 we see that if both receivers have F-MSI, i.e., when

$$R_{\mathcal{Y}}^{(p)} = 0 \quad \text{and} \quad R_{\mathcal{Z}}^{(p)} = 0,$$

then (9c) is redundant. Consequently, we recover:

Corollary 6 (F-MSI at $\mathcal{Y}$ and $\mathcal{Z}$ From [4, Th. 1]):[6] The capacity region of the SD-BC with F-MSI at both receivers ($R_{\mathcal{Y}}^{(p)} = 0$ and $R_{\mathcal{Z}}^{(p)} = 0$) is the set of rate-tuples $(R, R_{\mathcal{Y}}, R_{\mathcal{Z}})$ satisfying

$$R + R_{\mathcal{Y}} \leq H(Y) \tag{11a}$$

$$R + R_{\mathcal{Z}} \leq I(X; Z) \tag{11b}$$

for some PMF of the form

$$p(x, y, z) = p(x) W(y, z|x). \tag{12}$$

B. How P-MSI Affects Capacity

In this section we study how P-MSI at the receivers affects the capacity region of the SD-BC.

Remark 7 (P-MSI at $\mathcal{Y}$): P-MSI at the deterministic receiver $\mathcal{Y}$ can increase capacity if, and only if, the stochastic receiver $\mathcal{Z}$ has P-MSI ($R_{\mathcal{Y}}^{(p)} < R_{\mathcal{Y}}$) or a common message is transmitted ($R > 0$).

In particular, the "only if"-direction implies:

$$(0, R_{\mathcal{Y}}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}} \iff (0, R_{\mathcal{Y}}, R_{\mathcal{Z}}) \in \mathscr{C}, \tag{13}$$

---

[6]For the setting without a common message, the capacity region of the general BC with F-MSI at both receivers was established in [4, Th. 1]. This result readily extends to the setting with a common message.

where $\mathscr{C}$ denotes the capacity region of the SD-BC without MSI.

Proof: The "only-if" direction follows from Corollary 3. The "if-direction" follows from Examples 9 and 10 ahead. More specifically, Example 9 shows that F-MSI at Receiver $\mathcal{Y}$ can increase capacity if Receiver $\mathcal{Z}$ already has F-MSI; and Example 10 shows that F-MSI at Receiver $\mathcal{Y}$ can increase capacity if a common message is transmitted.[7] $\qquad\square$

Remark 8 (P-MSI at $\mathcal{Z}$): P-MSI at the stochastic receiver $\mathcal{Z}$ can increase the capacity region of the SD-BC; and this holds irrespective of whether or not the deterministic receiver has P-MSI or the encoder transmits a common message.

Proof: Assume no common message ($R = 0$). By Corollary 3 the capacity region without MSI at Receiver $\mathcal{Z}$ does not depend on whether or not Receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$), and hence we obtain from Example 9 ahead that F-MSI at Receiver $\mathcal{Z}$ can increase capacity, irrespective of whether or not Receiver $\mathcal{Y}$ has P-MSI.[8] Continuity considerations imply that the statement also holds with a common message, i.e., for some $R > 0$. $\qquad\square$

Example 9 (P-MSI Without a Common Message): Consider the SD-BC with binary input X and binary outputs

$$Y = X \quad \text{and} \quad Z = X \oplus S, \qquad (14)$$

where $S \sim \text{Ber}(p)$ is independent of X, and where $p \in (0, 1/2)$. Assume that the transmitter conveys only private messages ($R = 0$).

Using the capacity results of Section III-A, we can characterize the capacity region of the SC-BC (14) for the cases where no receiver has MSI and where the stochastic receiver $\mathcal{Z}$ has F-MSI:

- *Assume no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$). As shown in [17, Sec. 5.4.2], the capacity region $\mathscr{C}$ without MSI is the set of all rate-pairs $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$ that for some $\alpha \in [0, 1/2]$ satisfy*

$$R_{\mathcal{Y}} \leq h_{\text{b}}(\alpha) \qquad (15a)$$
$$R_{\mathcal{Z}} \leq 1 - h_{\text{b}}(\alpha + p - 2\alpha p). \qquad (15b)$$

In particular, (15) implies that without MSI the sum-rate is at most 1, and that it is strictly smaller than 1 whenever $R_{\mathcal{Z}} > 0$.

- *Assume F-MSI at $\mathcal{Z}$ ($R_{\mathcal{Y}}^{(p)} = 0$). By Corollary 5 (with $X \sim \text{Ber}(1/2)$) the capacity region $\mathscr{C}_{\text{P-MSI}}$ with F-MSI at Receiver $\mathcal{Z}$ is the set of rate-tuples $(0, R_{\mathcal{Y}}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ that satisfy*

$$R_{\mathcal{Z}} \leq 1 - h_{\text{b}}(p) \qquad (16a)$$
$$R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} \leq 1. \qquad (16b)$$

For the case where Receiver $\mathcal{Y}$ has no MSI and Receiver $\mathcal{Z}$ has F-MSI ($R_{\mathcal{Y}}^{(p)} = 0$ and $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$), Constraints (16) imply that the sum-rate is at most 1 but can also be 1 when $R_{\mathcal{Z}} > 0$.

For the case where both receivers have F-MSI ($R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Y}}^{(p)} = 0$), Constraints (16) imply that the sum-rate can exceed 1.

From the above observations we see that the capacity of the studied SD-BC (14) satisfies the following two:

1) *The capacity region without MSI is strictly contained in the capacity region without MSI at Receiver $\mathcal{Y}$ and with F-MSI at Receiver $\mathcal{Z}$.*

2) *The capacity region without MSI at Receiver $\mathcal{Y}$ and with F-MSI at Receiver $\mathcal{Z}$ is strictly contained in the capacity region with F-MSI at both receivers.*

Example 10 (P-MSI With a Common Message): Consider the SD-BC with input $X = (X_1, X_2)$, where X_1 and X_2 are binary, and with outputs

$$Y = X_1 + X_2 \quad \text{and} \quad Z = \begin{cases} X_2 & S = 0, \\ ? & S = 1, \end{cases} \qquad (17)$$

where $S \sim \text{Ber}(p)$ is independent of X, and where $p \in (0, 1)$.

In Appendix D we prove the following facts on the maximum sum-rate that is achievable on the SD-BC (17):

- *Assume F-MSI at $\mathcal{Y}$ and no MSI at $\mathcal{Z}$ ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} = 0$). Denote the set of PMFs $p(u, x, y, z)$ satisfying (8) by $\mathcal{P}_u$. The maximum achievable sum-rate $R + R_{\mathcal{Y}} + R_{\mathcal{Z}}$ is*

$$\max_{p(u,x,y,z)\in\mathcal{P}_u} \{H(Y|U) + I(U; Z)\}$$
$$= 1 - p + \left[p\, h_{\text{b}}\left(\frac{1}{1 + 2^{1/p}}\right) + \frac{2^{1/p}}{1 + 2^{1/p}} \right]. \quad (18)$$

Let $\mathcal{P}_u^\star$ denote the set of PMFs $p(u, x, y, z) \in \mathcal{P}_u$ that maximize the LHS of (18). W.r.t. every PMF $p(u, x, y, z) \in \mathcal{P}_u^\star$

$$I(U; Y) < \min\{H(Y), I(U; Z)\}. \qquad (19)$$

Moreover, the largest common-message rate $R_{\text{F-MSI@}\mathcal{Y}}^\star$ for which the maximum sum-rate is achievable is

$$R_{\text{F-MSI@}\mathcal{Y}}^\star = \max_{p(u,x,y,z)\in\mathcal{P}_u^\star} \min\{H(Y), I(U; Z)\}. \quad (20)$$

(Note that—by continuity and because the set $\mathcal{P}_u^\star$ is compact—the maxima in (18) and (20) are attained.)

- *Assume no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$). The maximum achievable sum-rate is again (18), and it can be achieved only if $p(u, x, y, z) \in \mathcal{P}_u$. Let $\mathcal{P}_{uv}^\star$ denote the set of PMFs $p(u, v, x, y, z)$ that for some $p(u, x, y, z) \in \mathcal{P}_u^\star$ are of the form*

$$p(u, v, x, y, z) = p(v|u)\, p(u, x, y, z).$$

The largest common-message rate $R_{\text{no-MSI}}^\star$ for which the maximum sum-rate is achievable satisfies

$$R_{\text{no-MSI}}^\star \leq \max_{p(u,v,x,y,z)\in\mathcal{P}_{uv}^\star} I(V; Y). \qquad (21)$$

---

[7]Continuity considerations imply that it is not necessary to assume F-MSI ($R_{\mathcal{Y}}^{(p)} = 0$ or $R_{\mathcal{Z}}^{(p)} = 0$), but that the statements also hold for P-MSI of the form $R_{\mathcal{Y}}^{(p)} \in (0, R_{\mathcal{Y}})$ or $R_{\mathcal{Z}}^{(p)} \in (0, R_{\mathcal{Z}})$.

[8]Continuity considerations imply that it is not necessary to assume F-MSI ($R_{\mathcal{Z}}^{(p)} = 0$), i.e., that the statement also holds for P-MSI of the form $R_{\mathcal{Y}}^{(p)} \in (0, R_{\mathcal{Y}})$.

We can use the above observations to compare $R^\star_{\text{no-MSI}}$ and $R^\star_{\text{F-MSI@}\mathcal{Y}}$:

$$
\begin{aligned}
R^\star_{\text{no-MSI}} &\overset{(a)}{\leq} \max_{p(u,v,x,y,z)\in\mathcal{P}^\star_{uv}} I(V;Y) \\
&\overset{(b)}{\leq} \max_{p(u,x,y,z)\in\mathcal{P}^\star_u} I(U;Y) \\
&\overset{(c)}{<} \max_{p(u,x,y,z)\in\mathcal{P}^\star_u} \min\{H(Y), I(U;Z)\} \\
&\overset{(d)}{=} R^\star_{\text{F-MSI@}\mathcal{Y}},
\end{aligned}
\tag{22}
$$

where (a) follows from (21); (b) holds by definition of $\mathcal{P}^\star_{uv}$; (c) follows from (19); and (d) follows from (20). The comparison reveals that for the SD-BC (17) the capacity region with F-MSI at Receiver $\mathcal{Y}$ and without MSI at Receiver $\mathcal{Z}$ strictly contains that without MSI.

C. Intuition on the Results

In this section we provide some intuition on the results of Section III-B. To keep the exposition simple we consider only the case without a common message.

When a transmitter sends two independent messages over a BC to two receivers, then the transmission to each of the receivers is interfered by the transmission to the other receiver. Since the transmitter knows the two messages, it knows the two transmissions acausally and can thus partially mitigate the interference experienced at each receiver (see Marton's scheme [1]).

Suppose now that a receiver has P-MSI. Such a receiver has partial knowledge of the transmission to the other receiver, and in general knowing interference at both the transmitter and the receiver is better (in terms of achievable rates) than knowing it only at the transmitter (cf. the state-dependent single-user channel with acausal state-information (SI) at the transmitter). Consequently, P-MSI at a receiver allows to better mitigate interference on the BC and hence to achieve larger rates. This provides some intuition for our finding that—in the setting without a common message—P-MSI at the SD-BC's stochastic receiver only can increase capacity.

In contrast, we have seen that—in the setting without a common message—P-MSI at the SD-BC's deterministic receiver only cannot increase capacity. Also for this result we can obtain some intuition from the single-user channel whose transmission is subject to interference. Indeed, if the single-user channel's outputs can be computed from its inputs and the interference, then knowing the interference only at the transmitter is as beneficial as knowing it also at the receiver.[9]

Interference can be mitigated more efficiently if both receivers have P-MSI. For example, if both receivers have F-MSI ($R^{(p)}_{\mathcal{Y}} = 0$ and $R^{(p)}_{\mathcal{Z}} = 0$) and, moreover, $R^{(c)}_{\mathcal{Y}} = R^{(c)}_{\mathcal{Z}}$,

---

[9]To see this, consider a deterministic state-dependent single-user channel $W(y|x,s)$, whose output is a function of its input and the state. That the capacity of this channel with acausal SI at the encoder does not depend on whether or not the state is revealed to the receiver can be seen as follows. If the receiver does not observe the state, then the capacity is given by the Gelfand-Pinsker formula [19]

$$
C = \max_{p(u,x|s)} I(U;Y) - I(U;S),
$$

then the transmitter can send the "x-or" of the messages $M^{(c)}_{\mathcal{Y}}$ and $M^{(c)}_{\mathcal{Z}}$ as a "common message" to both receivers, and each receiver can recover its message by first recovering the common message and then subtracting the message for the other receiver, which it knows. This provides some intuition for our finding that on the SD-BC P-MSI at both receiver's is better (in terms of achievable rates) than P-MSI at the stochastic receiver only.

IV. FEEDBACK ON THE SD-BC

This section investigates how feedback can be used on the SD-BC. The feedback that we consider is perfect or rate-limited, and it is only from the stochastic receiver $\mathcal{Z}$. (Recall that feedback from the deterministic receiver $\mathcal{Y}$ is useless.) For simplicity of exposition, we assume that there is no common message ($R = 0$). But our ideas easily extend to the common-message setting.

A. Preliminaries: An Enhanced BC

Consider the *enhanced BC* with P-MSI of Figure 3, which is obtained from the SD-BC with P-MSI by revealing the stochastic outputs $\{Z_i\}$ also to the deterministic receiver $\mathcal{Y}$. It plays an important role in the feedback code that we present in the next section.

The capacity region of the enhanced BC is defined similarly as that of the SD-BC (see Section II). We denote it by $\mathscr{C}^{(\text{enh})}_{\text{P-MSI}}$, and in the special case without MSI by $\mathscr{C}^{(\text{enh})}$.

Proposition 11 (Enhanced BC With P-MSI): The capacity region $\mathscr{C}^{(\text{enh})}_{\text{P-MSI}}$ of the enhanced BC with P-MSI is the set of rate-tuples $(R^{(p)}_{\mathcal{Y}}, R^{(c)}_{\mathcal{Y}}, R^{(p)}_{\mathcal{Z}}, R^{(c)}_{\mathcal{Z}})$ satisfying

$$
R_{\mathcal{Z}} \leq I(U;Z) \tag{23a}
$$

$$
R_{\mathcal{Y}} + R^{(p)}_{\mathcal{Z}} \leq I(X;Y,Z) \tag{23b}
$$

$$
R^{(p)}_{\mathcal{Y}} + R_{\mathcal{Z}} \leq I(X;Y,Z|U) + I(U;Z) \tag{23c}
$$

for some PMF of the form

$$
p(u,x,y,z) = p(u,x)\, W(y,z|x). \tag{24}
$$

Proof: The result is an immediate consequence of [4, Th. 3], which characterizes the capacity region of the BC with P-MSI and degraded message sets. This holds because, as we argue next, the capacity region of the enhanced BC remains unchanged if the stronger receiver $\mathcal{Y}$ must decode also the pair $(M^{(p)}_{\mathcal{Z}}, M^{(c)}_{\mathcal{Z}})$. To see this, note that a coding scheme is reliable on the enhanced BC with P-MSI if, and only if, the following two conditions hold: 1) Receiver $\mathcal{Z}$ can decode $(M^{(p)}_{\mathcal{Z}}, M^{(c)}_{\mathcal{Z}})$ reliably from $M^{(c)}_{\mathcal{Y}}$ and the outputs Z^n; and 2) Receiver $\mathcal{Y}$ can decode $(M^{(p)}_{\mathcal{Y}}, M^{(c)}_{\mathcal{Y}})$ reliably from $M^{(c)}_{\mathcal{Z}}$ and the outputs (Y^n, Z^n). But these two conditions are

which for $U = Y(X,S)$ evaluates to $\max_{p(x|s)} H(Y|S)$. If the receiver observes the state, then the capacity is

$$
C = \max_{p(x|s)} I(X;Y|S),
$$

which is equivalent to $\max_{p(x|s)} H(Y|S)$, because $Y = Y(X,S)$. This, and the fact that the receiver can always ignore the state that it observes, prove the claim.

equivalent to the conditions that result when in 2) we require that Receiver $\mathcal{Y}$ can—in addition to $(M_{\mathcal{Y}}^{(p)}, M_{\mathcal{Y}}^{(c)})$—decode also the pair $(M_{\mathcal{Z}}^{(p)}, M_{\mathcal{Z}}^{(c)})$ reliably. $\qquad\square$

Corollary 12 (Enhanced BC Without MSI): The capacity region $\mathscr{C}^{(\text{enh})}$ of the enhanced BC without MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$) is the set of rate-tuples $(R_{\mathcal{Y}}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying

$$R_{\mathcal{Y}} \leq I(X; Y, Z|U) \tag{25a}$$
$$R_{\mathcal{Z}} \leq I(U; Z) \tag{25b}$$

for some PMF of the form (24).

From Proposition 11 and Corollary 12 it follows that on the enhanced BC P-MSI at Receiver $\mathcal{Y}$ only cannot increase capacity:

Corollary 13 (Enhanced BC Without MSI at $\mathcal{Z}$): If Receiver $\mathcal{Z}$ has no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$), then, irrespective of $R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$,

$$(R_{\mathcal{Y}}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}^{(\text{enh})} \iff (R_{\mathcal{Y}}, R_{\mathcal{Z}}) \in \mathscr{C}^{(\text{enh})}.$$

B. Coding Scheme With Rate-Limited Feedback

Assume that the feedback is rate-limited (see Figure 1 when the dashed links transport the feedback signal W_i).[10] We first describe our feedback scheme for the SD-BC without MSI at the stochastic receiver $\mathcal{Z}$. Later, we shall generalize it to the case with P-MSI at both receivers.

The scheme has two phases: for some fixed $\alpha \in (0, 1)$, Phase 1 comprises the first αn channel uses $[1 : \alpha n]$, and Phase 2 comprises the remaining $(1 - \alpha)n$ channel uses $[\alpha n + 1 : n]$. We next describe Phase 1 and Phase 2, beginning with Phase 1.

Phase 1: In Phase 1 the transmitter codes for the enhanced BC with P-MSI at the deterministic receiver $\mathcal{Y}$ at rates

$$\left(\tilde{R}_{\mathcal{Y}}^{(p)}, 0, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}\right) \in \mathscr{C}_{\text{P-MSI}}^{(\text{enh})}. \tag{26a}$$

At the end of the phase the stochastic receiver $\mathcal{Z}$ decodes its intended rate-$(\tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)})$ messages. Moreover, it compresses its Phase-1 channel-outputs and uses the rate-limited feedback-link to send the compression index that it obtains to the encoder.[11] The deterministic receiver $\mathcal{Y}$ performs no action in Phase 1.

Phase 2: In Phase 2 the encoder transmits the compression index, which Receiver $\mathcal{Z}$ sent over the feedback link in Phase 1, along with fresh message-information. It transmits

---

[10]Recall from Section II that perfect feedback is equivalent to rate-limited feedback of rate $R_{\text{FB}} \geq \log |\mathcal{Z}|$.

[11]Receiver $\mathcal{Z}$ uses the feedback link only between the two phases, i.e., only at Time αn, and for every $i \neq \alpha n$ the feedback signal W_i thus takes value in a size-1 set $\mathcal{W}_i$. This allows the Receiver $\mathcal{Z}$ to choose the Time-αn feedback $W_{\alpha n}$ from a size-$2^{n R_{\text{FB}}}$ set $\mathcal{W}_{\alpha n}$ while guaranteeing that the rate-limitation (1) be met. By using B instances of the two phases, and by starting the transmission with sufficiently many instances of Phase 1, one could use the feedback link more evenly and (for $B \to \infty$) guarantee that each alphabet $\mathcal{W}_i$, $i \in [1 : n]$ be of size at most $2^{R_{\text{FB}}}$.

the compression index at rate $R_{\mathcal{Y}}^{(c)}$ and the fresh message-information at rates $(R_{\mathcal{Y}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$, where

$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}. \tag{26b}$$

Note that Receiver $\mathcal{Z}$ knows the rate-$R_{\mathcal{Y}}^{(c)}$ compression-index and can thus use it as P-MSI. At the end of the phase Receiver $\mathcal{Z}$ decodes its intended rate-$(R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ messages; and Receiver $\mathcal{Y}$ performs the following actions:

1) Based on its Phase-2 channel-outputs $Y_{\alpha n+1}^n$, it decodes the rate-$R_{\mathcal{Y}}^{(p)}$ message and the rate-$R_{\mathcal{Y}}^{(c)}$ compression-index that were sent to it in Phase 2.
2) Based on its estimate of the compression index and its Phase-1 channel-outputs $Y^{\alpha n}$, it decodes Receiver $\mathcal{Z}$'s Phase-1 channel-outputs $Z^{\alpha n}$.
3) Based on its Phase-1 channel-outputs $Y^{\alpha n}$ and its estimate of Receiver $\mathcal{Z}$'s Phase-1 channel-outputs $Z^{\alpha n}$, it decodes the rate-$\tilde{R}_{\mathcal{Y}}^{(p)}$ message that was sent to it in Phase 1.

By (26a) and (26b) we can guarantee that the probability of a decoding error tend to 0 as n tends to infinity whenever Receiver $\mathcal{Y}$ can recover Receiver $\mathcal{Z}$'s Phase-1 channel-outputs $Z^{\alpha n}$. As we argue next, we can guarantee that—with probability tending to 1 as n tends to infinity—the latter hold whenever the feedback-rate R_{FB} and the rate $R_{\mathcal{Y}}^{(c)}$ at which the compression index is sent in Phase 2 satisfy

$$\alpha H(Z|Y) < (1 - \alpha) R_{\mathcal{Y}}^{(c)} \tag{26c}$$
$$(1 - \alpha) R_{\mathcal{Y}}^{(c)} \leq R_{\text{FB}}, \tag{26d}$$

where the conditional entropy $H(Z|Y)$ is computed w.r.t. some PMF $p(u, x, y, z)$ of the form (24) for which the rate-tuple $(\tilde{R}_{\mathcal{Y}}^{(p)}, 0, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)})$ satisfies (23). Indeed, Condition (26c) guarantees that—with probability tending to 1 as n tends to infinity—Receiver $\mathcal{Y}$ can decode Receiver $\mathcal{Z}$'s Phase-1 channel-outputs $Z^{\alpha n}$ from its own Phase-1 channel-outputs and the compression index that is sent to it in Phase 2, and Condition (26d) guarantees that the compression index can be sent over the feedback link. To see this, recall that the compression index is sent in Phase 2, and that the rate $R_{\mathcal{Y}}^{(c)}$ at which it is sent is thus computed w.r.t. the $(1 - \alpha)n$ channel uses that Phase 2 comprises. Moreover, we can guarantee that—with probability tending to 1 as n tends to infinity—Receiver $\mathcal{Y}$ can decode Receiver $\mathcal{Z}$'s Phase-1 channel-outputs $Z^{\alpha n}$ based on the compression index and its own Phase-1 channel-outputs Y^n whenever the rate of the compression index—computed w.r.t. the first αn channel uses—exceeds $H(Z|Y)$.

From the above we conclude that our feedback code achieves any rate-tuple of the form

$$\alpha(\tilde{R}_{\mathcal{Y}}^{(p)}, 0, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}) + (1 - \alpha)(R_{\mathcal{Y}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \tag{27}$$

for which Conditions (26) hold. A more detailed analysis of the scheme can be found in Appendix E.

We can easily adapt the above feedback code to allow P-MSI also at the stochastic receiver $\mathcal{Z}$, i.e., $R_{\mathcal{Y}}^{(p)} \in [0, R_{\mathcal{Y}}]$

and $R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$. To this end we modify the code as follows:

- In Phase 1 the encoder codes for the enhanced BC with P-MSI at *both* receivers.
- In Phase 2 the compression index that is sent to the deterministic receiver constitutes *additional* P-MSI at the stochastic receiver, i.e., P-MSI beyond the one that the stochastic receiver already had in the beginning.

In Phase 1 of the adapted code the encoder thus codes at rates

$$(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}^{(\text{enh})}; \qquad (28a)$$

and in Phase 2 the encoder sends fresh message-information at rates $(R_{\mathcal{Y}}^{(p)}, \hat{R}_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$, where

$$\hat{R}_{\mathcal{Y}}^{(c)} \leq R_{\mathcal{Y}}^{(c)} - \frac{\alpha}{1-\alpha} H(Z|Y), \qquad (28b)$$

and

$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}. \qquad (28c)$$

The adapted feedback code achieves any rate-tuple of the form

$$\alpha(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Y}}^{(c)}, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}) + (1-\alpha)(R_{\mathcal{Y}}^{(p)}, \hat{R}_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$$

for which Conditions (26d) and (28) hold.

The following proposition summarizes which rate-tuples our feedback scheme can achieve:

Proposition 14 (Performance of the Feedback Code): Fix any PMF $p(u, x, y, z)$ of the form (24), and let $(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Y}}^{(c)}, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)})$ be any rate-tuple that—for the fixed PMF $p(u, x, y, z)$—satisfies (23). In addition, pick any rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}$, any nonnegative number

$$\alpha \leq \min\left\{ \frac{R_{\mathcal{Y}}^{(c)}}{R_{\mathcal{Y}}^{(c)} + H(Z|Y)}, \frac{R_{\text{FB}}}{H(Z|Y)} \right\}, \qquad (29)$$

where the conditional entropy $H(Z|Y)$ is computed w.r.t. the fixed PMF $p(u, x, y, z)$, and any nonnegative rate

$$\hat{R}_{\mathcal{Y}}^{(c)} \leq R_{\mathcal{Y}}^{(c)} - \frac{\alpha}{1-\alpha} H(Z|Y). \qquad (30)$$

The capacity region of the SD-BC with P-MSI and rate-limited feedback of rate R_{FB} from the stochastic receiver $\mathcal{Z}$ contains the rate-tuple

$$\alpha(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Y}}^{(c)}, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}) + (1-\alpha)(R_{\mathcal{Y}}^{(p)}, \hat{R}_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}).$$

Proof: See Appendix E. □

Our feedback code can be generalized along the following guidelines:

- To obtain a feedback code for the general BC, in Phase 2 one can replace the capacity-achieving code for the SD-BC with P-MSI by a "good" code for the general BC with P-MSI, e.g., by the code of [6].
- Instead of recovering $Z^{\alpha n}$ losslessly, after Phase 2 Receiver $\mathcal{Y}$ could recover a lossy version of these outputs.[12] To allow for this generalization, the enhanced BC would have to be adapted so that Receiver $\mathcal{Y}$ does not

observe Receiver $\mathcal{Z}$'s output but only a lossy version of it. In general, the capacity of such an enhanced BC (with P-MSI) is unknown; and hence the encoder would have to use a "good" rather than a capacity-achieving code for it, e.g., the code of [6] for the general BC with P-MSI.

- Instead of partitioning the transmission into two phases, one could use a block-Markov framework as in [16].

In the absence of MSI, extending our feedback code along the above guidelines results in the feedback code of [16], which—in the absence of MSI—is thus more general than our code.[13]

Using Proposition 14, we next identify sufficient conditions for feedback to increase the capacity of the SD-BC. For simplicity we assume that the stochastic receiver $\mathcal{Z}$ has no MSI, i.e., that $R_{\mathcal{Y}}^{(p)} = 0$. Proposition 15 treats the case where Receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$), and Proposition 16 treats the case without MSI ($R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Y}}$).

Proposition 15 (Sufficient Conditions With P-MSI at $\mathcal{Y}$): Consider an SD-BC with P-MSI only at the deterministic receiver $\mathcal{Y}$ ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}})$). If there exists a rate-triple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying

$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)} + R_{\mathcal{Z}}^{(c)}) \in (\partial\mathscr{C} \cap (\mathscr{C}_{\text{enh}} \setminus \partial\mathscr{C}_{\text{enh}})), \qquad (31)$$

and if, for some PMF $p(u, x, y, z)$ of the form (6), Conditions (5) and

$$R_{\mathcal{Y}} < H(Y) \quad \text{and} \quad 0 < I(U; Y) \qquad (32)$$

hold, then, irrespective of $R_{\text{FB}} > 0$, $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ is in the interior of the feedback capacity region, and feedback thus increases the capacity region.

Proof: See Appendix F. □

Proposition 16 (Sufficient Conditions Without MSI): Consider an SD-BC without MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$). If there exists a rate-pair $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$ satisfying

$$(R_{\mathcal{Y}}, R_{\mathcal{Z}}) \in (\partial\mathscr{C} \cap (\mathscr{C}_{\text{enh}} \setminus \partial\mathscr{C}_{\text{enh}})), \qquad (33)$$

and if, for some PMF $p(v, u, x, y, z)$ of the form (3), Conditions (5) and

$$R_{\mathcal{Y}} < H(Y) \quad \text{and} \quad 0 < I(V; Y) - I(V; Z) \qquad (34)$$

hold, then, irrespective of $R_{\text{FB}} > 0$, $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$ is in the interior of the feedback capacity region, and feedback thus increases the capacity region.

Proof: See Appendix F. □

Proposition 16 is used in the analysis of Example 18 ahead (see Appendix G), which proves that feedback can increase the capacity of the SD-BC without MSI. In this analysis one chooses $U = V$, which turns Marton coding into the simpler superposition coding with no satellite for Receiver $\mathcal{Z}$. There are other SD-BCs for which one has to apply Proposition 16 with full Marton coding, i.e., with $U \neq V$, in order to show that feedback increases their capacity (see Remark 25 in Appendix G).

---

[12]In particular, this generalization would allow us to extend the feedback code to continuous output alphabets $\mathcal{Z}$.

[13]In particular, in the absence of MSI the rate region that is achievable with our feedback code is contained in the rate region that is achievable with the feedback code of [16].

It is perhaps surprising that Proposition 16 can be used to show that feedback can increase the capacity of the SD-BC without MSI: The boundary points of the no-feedback capacity region of the SD-BC without MSI are known to be achievable using only satellites but no cloud-center, i.e., with $V = \emptyset$. This notwithstanding, Proposition 16 relies on the assumption that one can achieve boundary points using full Marton coding with a cloud-center, i.e., with $V \neq \emptyset$ (see (34)).

C. How Feedback Affects Capacity

We use the sufficient conditions of Propositions 15 and 16 to show that feedback can increase the capacity of the SD-BC without P-MSI and the sum-rate capacity of the SD-BC with P-MSI at the deterministic receiver $\mathcal{Y}$. (The assumption that the stochastic receiver $\mathcal{Z}$ has no MSI is made for simplicity.) Key to the proof are the following two observations:

1) P-MSI to the stochastic receiver can increase the capacity of the SD-BC.
2) The capacity of the enhanced BC (with P-MSI) is typically larger than that of the SD-BC (with P-MSI).

Theorem 17 (Feedback Can Help): If the stochastic receiver $\mathcal{Z}$ has no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$), then, irrespective of whether or not the deterministic receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$), rate-limited feedback of any positive rate can increase the capacity region of the SD-BC. In particular, it can increase the sum-rate capacity if Receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$).

The theorem follows from the following example, which we analyze in Appendix G:

Example 18 (Feedback Helps): Consider the SD-BC whose input is $X = (X_1, X_2)$, where X_1 and X_2 are binary, and whose outputs are $Y = X_1 + X_2$ and

$$Z = \begin{cases} X_2 & S = 0, \\ ? & S = 1, \end{cases}$$

where $S \sim \mathrm{Ber}(p)$, $p \in (0, 1)$ is independent of X. If the deterministic receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$) and the stochastic receiver $\mathcal{Z}$ has no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$), then rate-limited feedback of any positive rate increases the sum-rate capacity. If no receiver has MSI ($R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$ and $R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$) and $p > 1/2$, then such feedback increases the capacity region.

As our next result shows, the setting where the stochastic receiver $\mathcal{Z}$ has F-MSI is different:

Theorem 19 (With F-MSI at $\mathcal{Z}$, Feedback is Useless): If the stochastic receiver $\mathcal{Z}$ has F-MSI ($R_{\mathcal{Y}}^{(p)} = 0$), then, irrespective of whether or not the deterministic receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$), even perfect feedback cannot increase the capacity region of the SD-BC.

Proof: See Appendix H. □

From Theorems 17 and 19 it follows that whether feedback can increase the capacity of the SD-BC with P-MSI depends on the P-MSI. Of course, it also depends on the SD-BC. The following is an example of an SD-BC—which does not consists of two noninterfering single-user channels and is neither deterministic nor physically-degraded—whose capacity

without MSI at the stochastic receiver $\mathcal{Z}$ cannot be increased by feedback, irrespective of whether or not the deterministic receiver $\mathcal{Y}$ has P-MSI. The example is analyzed in Appendix I.

Example 20 (Also Without MSI at $\mathcal{Z}$, Feedback Can be Useless): Consider the SD-BC with input $X \in \mathcal{X}$ and outputs

$$Y = f(X) \quad and \quad Z = \begin{cases} X & if\ S = 0, \\ ? & if\ S = 1, \end{cases} \tag{35}$$

where $S \sim \mathrm{Ber}(p)$ is independent of X, where $p \in (0, 1)$, and where $? \notin \mathcal{X}$. Let $\bar{p} = 1 - p$ and

$$\mathcal{X}_y = \{x \in \mathcal{X}\colon f(x) = y\}, \quad y \in \mathcal{Y}.$$

For the case where the stochastic receiver $\mathcal{Z}$ has no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$), we show the following statement in Appendix I. Irrespective of $R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$, the capacity region is the set of rate-tuples $(R_{\mathcal{Y}}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying

$$R_{\mathcal{Y}} \leq H(Y) \tag{36a}$$

$$R_{\mathcal{Z}} \leq \bar{p} I(U; Y) + \bar{p} \sum_{y \in \mathcal{Y}} p(y) \log |\mathcal{X}_y| \tag{36b}$$

$$R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq H(Y) - p I(U; Y) + \bar{p} \sum_{y \in \mathcal{Y}} p(y) \log |\mathcal{X}_y| \tag{36c}$$

for some PMF $p(u, x, y, z)$ of the form (6) under which X is a function of (Y, U). Further, we show that this holds also in the presence of feedback. These observations imply that feedback cannot increase the capacity region of the considered SD-BC without MSI at Receiver $\mathcal{Z}$.

D. Strictly Causal SI on the State-Dependent SD-BC

In this section we consider the state-dependent SD-BC, whose transition law is governed by an IID state-sequence of finite support $\mathcal{S}$, i.e.,

$$W(y, z|x, s) = \mathbb{1}_{\{y = f(x)\}} W(z|x, s), \quad s \in \mathcal{S}.$$

We study how strictly-causal SI at the encoder affects the capacity of this channel. Because the setting with strictly-causal SI at the encoder is related to that with feedback, for the state-dependent SD-BC with strictly-causal SI at the encoder we can easily obtain results that are equivalent to those of Theorems 17 and 19 for the SD-BC with feedback.

Theorem 21 (Strictly-Causal SI): Consider the state-dependent SD-BC

$$W(y, z|x, s) = \mathbb{1}_{\{y = f(x)\}} W(z|x, s), \quad s \in \mathcal{S}.$$

If the stochastic receiver $\mathcal{Z}$ has no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$), then, irrespective of whether or not the deterministic receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$), strictly-causal SI can increase the capacity region of the SD-BC. In particular, it can increase the sum-rate capacity if Receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$). If Receiver $\mathcal{Z}$ has F-MSI ($R_{\mathcal{Y}}^{(p)} = 0$), then, irrespective of whether or not Receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$), strictly-causal SI cannot increase capacity.

Proof: The Functional Representation lemma (Lemma 1) allows us to view the state-less SD-BC $W(y, z|x)$ as a state-dependent SD-BC $W(y, z|x, s)$ whose stochastic output Z can

be computed from its input and state. If the state is revealed strictly-causally to the encoder, then the encoder can compute Receiver $\mathcal{Z}$'s past channel outputs from the past states and channel inputs, and hence feedback cannot be better than strictly-causal SI. Consequently, the first two claims of the theorem follow as a corollary to Theorem 17.

The last claim of the theorem (i.e., that if Receiver $\mathcal{Z}$ has F-MSI ($R_{\mathcal{Y}}^{(p)} = 0$), then, irrespective of whether or not Receiver $\mathcal{Y}$ has P-MSI ($R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$), strictly-causal SI cannot increase capacity) can be established along the line of argument in the proof of Theorem 19. □

APPENDIX A
PROOF OF THEOREM 2

A. Direct Part

1) Codebook Generation: Fix positive real numbers $\tilde{\epsilon} > \epsilon > 0$, a PMF $p(v, u, x) = p(v, u)\, p(x|u)$, and rates $\tilde{R}^{(c)}, \tilde{R}_{\mathcal{Y}}, \tilde{R}_{\mathcal{Z}} > 0$ for which $\tilde{R}^{(c)} \leq \min\{\tilde{R}_{\mathcal{Y}}, \tilde{R}_{\mathcal{Z}}\}$. For each $(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)}) \in \mathcal{M} \times \mathcal{M}_{\mathcal{Y}}^{(c)} \times \mathcal{M}_{\mathcal{Z}}^{(c)}$ draw $2^{n\tilde{R}^{(c)}}$ n-tuples $\mathbf{v}$ from the product distribution $\prod_{i=1}^{n} p(v_i)$, label them by $k \in [1 : 2^{n\tilde{R}^{(c)}}]$, where $\mathbf{v}(k)$ denotes the n-tuple labelled by k, and place them in a bin $\mathscr{B}(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)})$. For each n-tuple $\mathbf{v}$ draw $2^{n(\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)})}$ n-tuples $\mathbf{y} \in \mathcal{Y}^n$ independently of each other and each from the product distribution $\prod_{i=1}^{n} p(y_i|v_i)$, and label them by $\ell \in [1 : 2^{n(\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)})}]$, where $\mathbf{y}(k, \ell)$ denotes the ℓ-th n-tuple corresponding to $\mathbf{v}(k)$. Randomly allocate the $2^{n\tilde{R}_{\mathcal{Y}}}$ n-tuples to $2^{nR_{\mathcal{Y}}^{(p)}}$ bins $\mathscr{B}_{\mathcal{Y}}(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)}, m_{\mathcal{Y}}^{(p)})$, where $m_{\mathcal{Y}}^{(p)} \in \mathcal{M}_{\mathcal{Y}}^{(p)}$. Similarly, for each n-tuple $\mathbf{v}$ draw $2^{n(\tilde{R}_{\mathcal{Z}} - \tilde{R}^{(c)})}$ n-tuples $\mathbf{u} \in \mathcal{U}^n$ independently of each other and each from the product distribution $\prod_{i=1}^{n} p(u_i|v_i)$, and label them by $j \in [1 : 2^{n(\tilde{R}_{\mathcal{Z}} - \tilde{R}^{(c)})}]$, where $\mathbf{u}(k, j)$ denotes the j-th n-tuple $\mathbf{u}$ corresponding to $\mathbf{v}(k)$. Randomly allocate the $2^{n\tilde{R}_{\mathcal{Z}}}$ n-tuples to $2^{nR_{\mathcal{Z}}^{(p)}}$ bins $\mathscr{B}_{\mathcal{Z}}(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)}, m_{\mathcal{Z}}^{(p)})$, where $m_{\mathcal{Z}}^{(p)} \in \mathcal{M}_{\mathcal{Z}}^{(p)}$.

2) Encoding: If there are n-tuples $\mathbf{v} \in \mathscr{B}(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)})$, $\mathbf{y} \in \mathscr{B}_{\mathcal{Y}}(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, M_{\mathcal{Y}}^{(p)})$, and $\mathbf{u} \in \mathscr{B}_{\mathcal{Z}}(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, M_{\mathcal{Z}}^{(p)})$ for which $(\mathbf{v}, \mathbf{y}, \mathbf{u}) \in \mathcal{T}_{\epsilon}^{(n)}(V, Y, U)$, then the encoder draws the length-n channel-input-sequence X^n from the product distribution $\prod_{i=1}^{n} p(x_i|y_i, u_i)$. Otherwise the encoding is unsuccessful.

3) Decoding: Upon observing $(Y^n, M_{\mathcal{Z}}^{(c)})$, Receiver $\mathcal{Y}$ decodes $(m, m_{\mathcal{Y}})$ if it is the unique element of $\mathcal{M} \times \mathcal{M}_{\mathcal{Y}}$ for which $Y^n \in \mathscr{B}_{\mathcal{Y}}(m, m_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, m_{\mathcal{Y}}^{(p)}) \cap \mathcal{T}_{\tilde{\epsilon}}^{(n)}(Y)$. Otherwise it declares an error. Upon observing $(Z^n, M_{\mathcal{Y}}^{(c)})$, Receiver $\mathcal{Z}$ decodes $(m, m_{\mathcal{Z}})$ if it is the unique element of $\mathcal{M}_{\mathcal{Z}}$ for which Bin $\mathscr{B}_{\mathcal{Z}}(m, M_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)}, m_{\mathcal{Z}}^{(p)})$ contains an n-tuple $\mathbf{u}$ satisfying $(\mathbf{u}, Z^n) \in \mathcal{T}_{\tilde{\epsilon}}^{(n)}(U, Z)$. Otherwise it declares an error.

4) Analysis of the Error Probability: We first analyze the probability that the encoding is unsuccessful. By symmetry, this probability does not depend on the messages' realizations, and we thus assume w.l.g. that $M = M_{\mathcal{Y}}^{(p)} = M_{\mathcal{Y}}^{(c)} = M_{\mathcal{Z}}^{(p)} = M_{\mathcal{Z}}^{(c)} = 1$. For ease of notation, let $\mathscr{B} = \mathscr{B}(1, 1, 1)$, $\mathscr{B}_{\mathcal{Y}} = \mathscr{B}_{\mathcal{Y}}(1, 1, 1, 1)$, and $\mathscr{B}_{\mathcal{Z}} = \mathscr{B}_{\mathcal{Z}}(1, 1, 1, 1)$. Note that

the encoder can choose one out of

$$N = \sum_{k, \ell, j} \mathbb{1}_{(\mathbf{v}(k), \mathbf{y}(k, \ell), \mathbf{u}(k, j)) \in \mathcal{T}_{\epsilon}^{(n)}} \mathbb{1}_{\mathbf{y}(k, \ell) \in \mathscr{B}_{\mathcal{Y}}} \mathbb{1}_{\mathbf{u}(k, j) \in \mathscr{B}_{\mathcal{Z}}}$$

triples $(\mathbf{v}, \mathbf{y}, \mathbf{u})$, where $k \in [1 : 2^{n\tilde{R}^{(c)}}]$, $\ell \in [1 : 2^{n(\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)})}]$, and $j \in [1 : 2^{n(\tilde{R}_{\mathcal{Z}} - \tilde{R}^{(c)})}]$. The encoding is unsuccessful if $N = 0$. As in the proof of the mutual covering lemma [17, Lemma 8.1], we obtain from Chebyshev's inequality that

$$\mathbb{P}[N = 0] \leq \mathbb{P}\left[\left(N - \mathbb{E}[N]\right)^2 \geq \mathbb{E}[N]^2\right] \leq \frac{\mathrm{Var}(N)}{\mathbb{E}[N]^2}. \quad (37)$$

To conclude that—on average over the realization of the code—the encoding is with high probability successful, it thus suffices to show that

$$\frac{\mathrm{Var}(N)}{\mathbb{E}[N]^2} \to 0 \ (n \to \infty).$$

Define

$$E^{k, \ell, j} = \mathbb{1}_{(\mathbf{v}(k), \mathbf{y}(k, \ell), \mathbf{u}(k, j)) \in \mathcal{T}_{\epsilon}^{(n)}} \mathbb{1}_{\mathbf{y}(k, \ell) \in \mathscr{B}_{\mathcal{Y}}} \mathbb{1}_{\mathbf{u}(k, j) \in \mathscr{B}_{\mathcal{Z}}} \quad (38a)$$

$$T^{k, \ell, j} = \mathbb{1}_{(\mathbf{v}(k), \mathbf{y}(k, \ell), \mathbf{u}(k, j)) \in \mathcal{T}_{\epsilon}^{(n)}}, \quad (38b)$$

and note that

$$\mathbb{E}[N^2] = \sum_{k, \ell, j} \sum_{k', \ell', j'} \mathbb{E}\left[E^{k, \ell, j} E^{k', \ell', j'}\right] \quad (39)$$

$$= \sum_{k, \ell, j} \sum_{k' \neq k, \ell', j'} \mathbb{E}\left[E^{k, \ell, j} E^{k', \ell', j'}\right]$$

$$+ \sum_{k, \ell, j} \sum_{\ell' \neq \ell, j' \neq j} \mathbb{E}\left[E^{k, \ell, j} E^{k, \ell', j'}\right]$$

$$+ \sum_{k, \ell, j} \sum_{j' \neq j} \mathbb{E}\left[E^{k, \ell, j} E^{k, \ell, j'}\right]$$

$$+ \sum_{k, \ell, j} \sum_{\ell' \neq \ell} \mathbb{E}\left[E^{k, \ell, j} E^{k, \ell', j}\right]$$

$$+ \sum_{k, \ell, j} \mathbb{E}\left[E^{k, \ell, j}\right], \quad (40)$$

where we used that $E^{k, \ell, j} E^{k, \ell, j} = E^{k, \ell, j}$. If $k \neq k'$, then $E^{k, \ell, j}$ and $E^{k', \ell', j'}$ are independent and

$$\sum_{k, \ell, j} \sum_{k' \neq k, \ell', j'} \mathbb{E}\left[E^{k, \ell, j} E^{k', \ell', j'}\right]$$

$$= \sum_{k, \ell, j} \sum_{k' \neq k, \ell', j'} \mathbb{E}\left[E^{k, \ell, j}\right] \mathbb{E}\left[E^{k', \ell', j'}\right] \quad (41)$$

$$\leq \left(\sum_{k, \ell, j} \mathbb{E}\left[E^{k, \ell, j}\right]\right)^2 \quad (42)$$

$$= \mathbb{E}[N]^2. \quad (43)$$

If $k = k'$, $\ell \neq \ell'$, and $j \neq j'$, then

$$\sum_{k,\ell,j} \sum_{\ell' \neq \ell, j' \neq j} \mathbb{E}\left[E^{k,\ell,j} E^{k,\ell',j'}\right]$$

$$\overset{(a)}{\leq} 2^{-2n(R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)})} \sum_{k,\ell,j} \sum_{\ell' \neq \ell, j' \neq j} \mathbb{E}\left[T^{k,\ell,j} T^{k,\ell',j'}\right] \quad (44)$$

$$\overset{(b)}{\leq} 2^{-2n(R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)})} \sum_{k,\ell,j} \sum_{\ell' \neq \ell, j' \neq j} 2^{-n(2I(Y;U|V) - \delta(\epsilon))} \quad (45)$$

$$= 2^{n(2\tilde{R}_{\mathcal{Y}} - 2R_{\mathcal{Y}}^{(p)} + 2\tilde{R}_{\mathcal{Z}} - 2R_{\mathcal{Z}}^{(p)} - 3\tilde{R}^{(c)} - 2I(Y;U|V) + \delta(\epsilon))}, \quad (46)$$

where (a) holds because $\mathbb{1}_{\mathbf{y}(k,\ell) \in \mathscr{B}_{\mathcal{Y}}}$, $\mathbb{1}_{\mathbf{u}(k,j) \in \mathscr{B}_{\mathcal{Z}}}$, $\mathbb{1}_{\mathbf{y}(k,\ell') \in \mathscr{B}_{\mathcal{Y}}}$, and $\mathbb{1}_{\mathbf{u}(k,j') \in \mathscr{B}_{\mathcal{Z}}}$ are independent of each other and of $T^{k,\ell,j} T^{k,\ell',j'}$, and because

$$\mathbb{E}\left[\mathbb{1}_{\mathbf{y}(k,\ell) \in \mathscr{B}_{\mathcal{Y}}}\right] = 2^{-nR_{\mathcal{Y}}^{(p)}}$$
$$\mathbb{E}\left[\mathbb{1}_{\mathbf{u}(k,j) \in \mathscr{B}_{\mathcal{Z}}}\right] = 2^{-nR_{\mathcal{Z}}^{(p)}};$$

and where (b) holds because of the properties of typical sequences. (Recall that $\delta(\cdot)$ denotes any function of ϵ that converges to 0 as ϵ approaches 0.) If $k = k'$, $\ell = \ell'$, and $j \neq j'$, then

$$\sum_{k,\ell,j} \sum_{j' \neq j} \mathbb{E}\left[E^{k,\ell,j} E^{k,\ell,j'}\right]$$

$$\overset{(a)}{\leq} 2^{-n(R_{\mathcal{Y}}^{(p)} + 2R_{\mathcal{Z}}^{(p)})} \sum_{k,\ell,j} \sum_{j' \neq j} \mathbb{E}\left[T^{k,\ell,j} T^{k,\ell,j'}\right] \quad (47)$$

$$\overset{(b)}{\leq} 2^{-n(R_{\mathcal{Y}}^{(p)} + 2R_{\mathcal{Z}}^{(p)})} \sum_{k,\ell,j} \sum_{j' \neq j} 2^{-n(2I(Y;U|V) - \delta(\epsilon))} \quad (48)$$

$$= 2^{n(\tilde{R}_{\mathcal{Y}} - R_{\mathcal{Y}}^{(p)} + 2\tilde{R}_{\mathcal{Z}} - 2R_{\mathcal{Z}}^{(p)} - 2\tilde{R}^{(c)} - 2I(Y;U|V) + \delta(\epsilon))}, \quad (49)$$

where (a) holds because $\mathbb{1}_{\mathbf{y}(k,\ell) \in \mathscr{B}_{\mathcal{Y}}} \mathbb{1}_{\mathbf{y}(k,\ell) \in \mathscr{B}_{\mathcal{Y}}} = \mathbb{1}_{\mathbf{y}(k,\ell) \in \mathscr{B}_{\mathcal{Y}}}$ and because $\mathbb{1}_{\mathbf{y}(k,\ell) \in \mathscr{B}_{\mathcal{Y}}}$, $\mathbb{1}_{\mathbf{u}(k,j) \in \mathscr{B}_{\mathcal{Z}}}$, and $\mathbb{1}_{\mathbf{u}(k,j') \in \mathscr{B}_{\mathcal{Z}}}$ are independent of each other and of $T^{k,\ell,j} T^{k,\ell,j'}$; and where (b) follows from the properties of typical sequences. Similarly, if $k = k'$, $j = j'$, and $\ell \neq \ell'$, then

$$\sum_{k,\ell,j} \sum_{\ell' \neq \ell} \mathbb{E}\left[E^{k,\ell,j} E^{k,\ell',j}\right]$$

$$= 2^{n(2\tilde{R}_{\mathcal{Y}} - 2R_{\mathcal{Y}}^{(p)} + \tilde{R}_{\mathcal{Z}} - R_{\mathcal{Z}}^{(p)} - 2\tilde{R}^{(c)} - 2I(Y;U|V) + \delta(\epsilon))}. \quad (50)$$

Finally, if $k = k'$, $j = j'$, and $\ell = \ell'$, then

$$\sum_{k,\ell,j} \mathbb{E}\left[E^{k,\ell,j}\right] = \mathbb{E}[N]. \quad (51)$$

Using that $\text{Var}(N) = \mathbb{E}\left[N^2\right] - \mathbb{E}[N]^2$, we obtain from (37), (40), (43), (46), (49), (50), and (51) that

$$\mathbb{P}[N = 0] \to 0 \quad (n \to \infty)$$

holds whenever the RHS of each of the equations (46), (49), (50), and (51) is—asymptotically—negligibly-small compared

to $\mathbb{E}[N]^2$. Note that

$$\mathbb{E}[N] = \sum_{k,\ell,j} \mathbb{E}\left[E^{k,\ell,j}\right] \quad (52)$$

$$\overset{(a)}{=} 2^{-n(R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)})} \sum_{k,\ell,j} \mathbb{E}\left[T^{k,\ell,j}\right] \quad (53)$$

$$\overset{(b)}{\geq} 2^{-n(R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)})} \sum_{k,\ell,j} 2^{-n(I(Y;U|V) + \delta(\epsilon))} \quad (54)$$

$$= 2^{n(\tilde{R}_{\mathcal{Y}} - R_{\mathcal{Y}}^{(p)} + \tilde{R}_{\mathcal{Z}} - R_{\mathcal{Z}}^{(p)} - \tilde{R}^{(c)} - I(Y;U|V) - \delta(\epsilon))}, \quad (55)$$

where (a) is true because $\mathbb{1}_{\mathbf{y}(k,\ell) \in \mathscr{B}_{\mathcal{Y}}}$ and $\mathbb{1}_{\mathbf{u}(k,j) \in \mathscr{B}_{\mathcal{Z}}}$ are independent of each other and of $T^{k,\ell,j}$; and where (b) holds for all sufficiently-large n by the properties of typical sequences. This proves that—on average over the realization of the code—the probability that the encoding is unsuccessful converges to 0 as n tends to infinity whenever

$$\tilde{R}^{(c)} > 3\delta(\epsilon) \quad (56a)$$
$$\tilde{R}_{\mathcal{Y}} - R_{\mathcal{Y}}^{(p)} > 3\delta(\epsilon) \quad (56b)$$
$$\tilde{R}_{\mathcal{Z}} - R_{\mathcal{Z}}^{(p)} > 3\delta(\epsilon) \quad (56c)$$
$$\tilde{R}_{\mathcal{Y}} - R_{\mathcal{Y}}^{(p)} + \tilde{R}_{\mathcal{Z}} - R_{\mathcal{Z}}^{(p)} - \tilde{R}^{(c)} > I(Y;U|V) + \delta(\epsilon). \quad (56d)$$

Suppose now that the encoding was successful. Under this assumption, we next analyze the probability that the decoding is unsuccessful. Because Receiver $\mathcal{Y}$ observes the n-tuple $\mathbf{y}$ that the encoder selected, and because $\mathcal{T}_{\epsilon}^{(n)}(Y) \subseteq \mathcal{T}_{\tilde{\epsilon}}^{(n)}(Y)$, it is clear that $Y^n \in \mathscr{B}_{\mathcal{Y}}(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, M_{\mathcal{Y}}^{(p)}) \cap \mathcal{T}_{\tilde{\epsilon}}^{(n)}(Y)$. Hence, the pair $(M, M_{\mathcal{Y}})$ satisfies the decoding requirements. Moreover, for all sufficiently-large n it holds with high probability that $(U^n, Z^n) \in \mathcal{T}_{\tilde{\epsilon}}^{(n)}(P_{U,Z})$, and hence that the pair $(M, M_{\mathcal{Z}})$ satisfies the decoding requirements. From this we conclude that we need to worry only about the event that also other message-pairs $(m, m_{\mathcal{Y}}) \in \mathcal{M} \times \mathcal{M}_{\mathcal{Y}} \setminus \{(M, M_{\mathcal{Y}})\}$ or $(m, m_{\mathcal{Z}}) \in \mathcal{M} \times \mathcal{M}_{\mathcal{Z}} \setminus \{(M, M_{\mathcal{Z}})\}$ satisfy the decoding requirements. For Receiver $\mathcal{Y}$ this can happen if an n-tuple $\mathbf{y}$ that corresponds to the same length-n sequence $\mathbf{v}$ as Y^n satisfies the decoding requirements, or if an n-tuple $\mathbf{y}$ that corresponds to a different length-n sequence $\mathbf{v}$ than Y^n satisfies the decoding requirements; and similarly for Receiver $\mathcal{Z}$. Hence, by the properties of typical sequences, the probability that the decoding is unsuccessful converges to 0 as n tends to infinity if

$$\tilde{R}_{\mathcal{Y}} - \tilde{R}^{(c)} < H(Y|V) - \delta(\tilde{\epsilon}) \quad (57a)$$
$$\tilde{R}_{\mathcal{Y}} + R + R_{\mathcal{Y}}^{(c)} < H(Y) - \delta(\tilde{\epsilon}) \quad (57b)$$
$$\tilde{R}_{\mathcal{Z}} - \tilde{R}^{(c)} < I(U;Z|V) - \delta(\tilde{\epsilon}) \quad (57c)$$
$$\tilde{R}_{\mathcal{Z}} + R + R_{\mathcal{Z}}^{(c)} < I(U;Z) - \delta(\tilde{\epsilon}). \quad (57d)$$

Now let ϵ and $\tilde{\epsilon}$ tend to 0. Then, also $\delta(\epsilon)$ and $\delta(\tilde{\epsilon})$ tend to 0, and hence we conclude that a rate-tuple $(R, R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ is achievable if there exist rates

$\tilde{R}^{(c)}$, $\tilde{R}_{\mathcal{Y}}$, and $\tilde{R}_{\mathcal{Z}}$ for which

$$-\tilde{R}^{(c)} < 0 \tag{58a}$$

$$R_{\mathcal{Y}}^{(p)} - \tilde{R}_{\mathcal{Y}} < 0 \tag{58b}$$

$$R_{\mathcal{Z}}^{(p)} - \tilde{R}_{\mathcal{Z}} < 0 \tag{58c}$$

$$R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)} + \tilde{R}^{(c)} - \tilde{R}_{\mathcal{Y}} - \tilde{R}_{\mathcal{Z}} < -I(Y;U|V) \tag{58d}$$

$$\tilde{R}^{(c)} - \tilde{R}_{\mathcal{Y}} < 0 \tag{58e}$$

$$-\tilde{R}^{(c)} + \tilde{R}_{\mathcal{Y}} < H(Y|V) \tag{58f}$$

$$R + R_{\mathcal{Y}}^{(c)} + \tilde{R}_{\mathcal{Y}} < H(Y) \tag{58g}$$

$$\tilde{R}^{(c)} - \tilde{R}_{\mathcal{Z}} < 0 \tag{58h}$$

$$-\tilde{R}^{(c)} + \tilde{R}_{\mathcal{Z}} < I(U;Z|V) \tag{58i}$$

$$R + R_{\mathcal{Z}}^{(c)} + \tilde{R}_{\mathcal{Z}} < I(U;Z). \tag{58j}$$

A Fourier-Motzkin elimination reveals that the above inequalities hold if the rate-tuple $(R, R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ lies in the interior of $\mathscr{C}_{\text{P-MSI}}$.

At first sight, it is perhaps surprising that we bin the cloud-center. As the following remark shows, this is simply an equivalent alternative to rate-splitting. The benefit is that binning the cloud-center requires only one auxiliary rate, namely $\tilde{R}^{(c)}$, whereas rate-splitting requires two auxiliary rates, namely one for each rate $R_{\mathcal{Y}}^{(p)}$ and $R_{\mathcal{Z}}^{(p)}$.

Remark 22: The effect of binning the cloud-center is that of rate-splitting. More precisely, instead of generating $2^{n\tilde{R}^{(c)}}$ cloud-centers $\mathbf{v}$ for each triple $(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)})$, we could execute the following three steps: 1) divide the messages $M_{\mathcal{Y}}^{(p)}$ and $M_{\mathcal{Z}}^{(p)}$ into two parts, i.e., $M_{\mathcal{Y}}^{(p)} = (M_{\mathcal{Y},s}^{(p)}, M_{\mathcal{Y},c}^{(p)})$ and $M_{\mathcal{Z}}^{(p)} = (M_{\mathcal{Z},s}^{(p)}, M_{\mathcal{Z},c}^{(p)})$; 2) generate for each tuple $(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)}, m_{\mathcal{Y},c}^{(p)}, m_{\mathcal{Z},c}^{(p)})$ a cloud-center $\mathbf{v}$; and 3) allocate the associated satellites $\mathbf{y}$ and $\mathbf{u}$ to $2^{nR_{\mathcal{Y},s}^{(p)}}$ and $2^{nR_{\mathcal{Z},s}^{(p)}}$ instead of $2^{nR_{\mathcal{Y}}^{(p)}}$ and $2^{nR_{\mathcal{Z}}^{(p)}}$ bins, respectively.

To see that binning the cloud-center is tantamount to rate-splitting, note that with rate-splitting we generate $2^{n(R_{\mathcal{Y},c}^{(p)} + R_{\mathcal{Z},c}^{(p)})}$ cloud-centers $\mathbf{v}$ per triple $(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)})$, namely one for each pair $(m_{\mathcal{Y},c}^{(p)}, m_{\mathcal{Z},c}^{(p)})$. Therefore, we can identify $R_{\mathcal{Y},c}^{(p)} + R_{\mathcal{Z},c}^{(p)}$ in the rate-splitting code with $\tilde{R}^{(c)}$ in the code where we also bin the cloud-center. Moreover, associating every cloud-center $\mathbf{v}$ that corresponds to some triple $(m, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(c)})$ with a different pair $(m_{\mathcal{Y},c}^{(p)}, m_{\mathcal{Z},c}^{(p)})$ and allocating the satellites $\mathbf{y}$ and $\mathbf{u}$ to $2^{nR_{\mathcal{Y},s}^{(p)}}$ and $2^{nR_{\mathcal{Z},s}^{(p)}}$ bins, respectively, is tantamount to not associating the cloud-center with anything while allocating the satellites $\mathbf{y}$ and $\mathbf{u}$ to $2^{nR_{\mathcal{Y}}^{(p)}}$ and $2^{nR_{\mathcal{Z}}^{(p)}}$ bins, respectively. From these observations it follows that the only difference between binning the cloud-center and rate-splitting is that for the latter $\tilde{R}^{(c)} = R_{\mathcal{Y},c}^{(p)} + R_{\mathcal{Z},c}^{(p)}$ must satisfy the upper bound $\tilde{R}^{(c)} \leq R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)}$. This upper bound is not, however, restrictive: if $\tilde{R}^{(c)} > R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)}$, then we have more cloud-centers $\mathbf{v}$ than message-tuples $(m, m_{\mathcal{Y}}^{(p)}, m_{\mathcal{Y}}^{(c)}, m_{\mathcal{Z}}^{(p)}, m_{\mathcal{Z}}^{(c)})$; and this cannot be better than having for each message-tuple a different cloud-center.

B. Converse

Let $Q \sim \text{Unif}[1:n]$ be independent of $(M, M_{\mathcal{Y}}, M_{\mathcal{Z}})$, and introduce

$$V_Q = (M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Z^{Q-1}), \tag{59a}$$

$$U_Q = (M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}), \tag{59b}$$

$$V = (V_Q, Q), \quad U = (U_Q, Q), \tag{59c}$$

$$X = X_Q, \quad Y = Y_Q, \quad Z = Z_Q. \tag{59d}$$

Note that V, U, X, and (Y, Z) form a Markov chain in that order, i.e., that their PMF is of the form (3).

The rate of the message-pair $(M, M_{\mathcal{Y}})$ intended to Receiver $\mathcal{Y}$ satisfies

$$R + R_{\mathcal{Y}} - \epsilon_n \overset{(a)}{\leq} \frac{1}{n} I(M, M_{\mathcal{Y}}; Y^n, M_{\mathcal{Z}}^{(c)}) \tag{60}$$

$$\overset{(b)}{=} I(M, M_{\mathcal{Y}}; Y_Q | M_{\mathcal{Z}}^{(c)}, Y^{Q-1}, Q) \tag{61}$$

$$\overset{(c)}{\leq} H(Y), \tag{62}$$

where (a) follows from Fano's inequality; (b) follows from the chain-rule and the independence of $(M, M_{\mathcal{Y}})$ and $M_{\mathcal{Z}}^{(c)}$; and (c) holds because conditional entropy is nonnegative and conditioning cannot increase entropy.

The rate of the message-pair $(M, M_{\mathcal{Z}})$ intended to Receiver $\mathcal{Z}$ satisfies

$$R + R_{\mathcal{Z}} - \epsilon_n \overset{(a)}{\leq} \frac{1}{n} I(M, M_{\mathcal{Z}}; Z^n, M_{\mathcal{Y}}^{(c)}) \tag{63}$$

$$\overset{(b)}{=} I(M, M_{\mathcal{Z}}; Z_Q | M_{\mathcal{Y}}^{(c)}, Z^{Q-1}, Q) \tag{64}$$

$$\overset{(c)}{\leq} I(U;Z), \tag{65}$$

where (a) follows from Fano's inequality; (b) holds because of the chain-rule and because $(M, M_{\mathcal{Z}})$ and $M_{\mathcal{Y}}^{(c)}$ are independent; and (c) is true because conditioning cannot increase entropy.

We prove the sum-rate constraints using Csiszár's sum-identity, which states that, for every tuple (A^n, B^n, T),

$$0 = \frac{1}{n} \sum_{i=1}^n \left[I(A_{i+1}^n; B_i | B^{i-1}, T) - I(B^{i-1}; A_i | A_{i+1}^n, T) \right]$$

$$= I(A_{Q+1}^n; B_Q | B^{Q-1}, T, Q) - I(B^{Q-1}; A_Q | A_{Q+1}^n, T, Q),$$

where $Q \sim \text{Unif}[1:n]$ is independent of (A^n, B^n, T). The first sum-rate constraint that we prove is that

$$R + R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}} - \epsilon_n$$

$$\overset{(a)}{\leq} \frac{1}{n} \Big[I(M_{\mathcal{Y}}^{(p)}; Y^n, M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}})$$

$$\quad + I(M, M_{\mathcal{Z}}; Z^n, M_{\mathcal{Y}}^{(c)}) \Big] \tag{66}$$

$$\overset{(b)}{=} I(M_{\mathcal{Y}}^{(p)}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Q)$$

$$\quad + I(M, M_{\mathcal{Z}}; Z_Q | M_{\mathcal{Y}}^{(c)}, Z^{Q-1}, Q) \tag{67}$$

$$\overset{(c)}{=} I(M_{\mathcal{Y}}^{(p)}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q)$$

$$\quad - I(Z^{Q-1}; Y_Q | M, M_{\mathcal{Y}}, M_{\mathcal{Z}}, Y_{Q+1}^n, Q)$$

$$\quad + I(Z^{Q-1}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Q)$$

$$+ I\big(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q; Z_Q\big)$$
$$- I\big(M_{\mathcal{Y}}^{(c)}, Z^{Q-1}, Q; Z_Q\big)$$
$$- I\big(Y_{Q+1}^n; Z_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Z^{Q-1}, Q\big) \qquad (68)$$

$$\overset{(d)}{\leq} I\big(M_{\mathcal{Y}}^{(p)}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q\big)$$
$$+ I\big(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q; Z_Q\big) \qquad (69)$$

$$\overset{(e)}{\leq} H\big(Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q\big)$$
$$+ I\big(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q; Z_Q\big) \qquad (70)$$

$$\overset{(f)}{=} H(Y|U) + I(U; Z), \qquad (71)$$

where (a) follows from Fano's inequality; (b) follows from the chain-rule and the independence of M, $M_{\mathcal{Y}}^{(p)}$, $M_{\mathcal{Y}}^{(c)}$, and $M_{\mathcal{Z}}$; (c) follows from the chain-rule; (d) holds because of Csiszár's sum-identity and because mutual information is nonnegative; (e) holds because conditional entropy is nonnegative; and (f) follows from (59). Similarly, we obtain the sum-rate constraint

$$R + R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} - \epsilon_n$$
$$\overset{(a)}{\leq} \frac{1}{n}\Big[I\big(M, M_{\mathcal{Y}}^{(c)}; Y^n, M_{\mathcal{Z}}^{(c)}\big)$$
$$+ I\big(M_{\mathcal{Y}}^{(p)}; Y^n, M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}\big)$$
$$+ I\big(M_{\mathcal{Z}}^{(p)}; Z^n, M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}\big)\Big] \qquad (72)$$

$$\overset{(b)}{=} I\big(M, M_{\mathcal{Y}}^{(c)}; Y_Q | M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Q\big)$$
$$+ I\big(M_{\mathcal{Y}}^{(p)}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Q\big)$$
$$+ I\big(M_{\mathcal{Z}}^{(p)}; Z_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Z^{Q-1}, Q\big) \qquad (73)$$

$$\overset{(c)}{=} I\big(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Z^{Q-1}, Q; Y_Q\big)$$
$$- I\big(M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Q; Y_Q\big)$$
$$- I\big(Z^{Q-1}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Q\big)$$
$$+ I\big(M_{\mathcal{Y}}^{(p)}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q\big)$$
$$- I\big(Z^{Q-1}; Y_Q | M, M_{\mathcal{Y}}, M_{\mathcal{Z}}, Y_{Q+1}^n, Q\big)$$
$$+ I\big(Z^{Q-1}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Q\big)$$
$$+ I\big(M_{\mathcal{Z}}^{(p)}; Z_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Z^{Q-1}, Q\big)$$
$$- I\big(Y_{Q+1}^n; Z_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Z^{Q-1}, Q\big)$$
$$+ I\big(Y_{Q+1}^n; Z_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Z^{Q-1}, Q\big) \qquad (74)$$

$$\overset{(d)}{\leq} I\big(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Z^{Q-1}, Q; Y_Q\big)$$
$$+ I\big(M_{\mathcal{Y}}^{(p)}; Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q\big)$$
$$+ I\big(M_{\mathcal{Z}}^{(p)}; Z_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Z^{Q-1}, Q\big) \qquad (75)$$

$$\overset{(e)}{\leq} I\big(M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Z^{Q-1}, Q; Y_Q\big)$$
$$+ H\big(Y_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}, Y_{Q+1}^n, Z^{Q-1}, Q\big)$$
$$+ I\big(M_{\mathcal{Z}}^{(p)}; Z_Q | M, M_{\mathcal{Y}}^{(c)}, M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Z^{Q-1}, Q\big) \qquad (76)$$

$$\overset{(f)}{=} I(V; Y) + H(Y|U) + I(U; Z|V), \qquad (77)$$

where (a) follows from Fano's inequality; (b) holds because of the chain-rule and because M, $M_{\mathcal{Y}}^{(p)}$, $M_{\mathcal{Y}}^{(c)}$, $M_{\mathcal{Z}}^{(p)}$, and

$M_{\mathcal{Z}}^{(c)}$ are independent; (c) follows from the chain-rule; (d) is obtained by using Csiszár's sum-identity twice and by using that mutual information is nonnegative; (e) holds because conditional entropy is nonnegative; and (f) follows from (59). Our last sum-rate constraint is that

$$2R + R_{\mathcal{Y}} + R_{\mathcal{Z}} - \epsilon_n$$
$$\overset{(a)}{\leq} \frac{1}{n} I\big(M, M_{\mathcal{Y}}^{(c)}; Y^n, M_{\mathcal{Z}}^{(c)}\big) + H(Y|U) + I(U; Z) \qquad (78)$$
$$\overset{(b)}{\leq} I\big(M, M_{\mathcal{Y}}^{(c)}; Y_Q | M_{\mathcal{Z}}^{(c)}, Y_{Q+1}^n, Q\big) + H(Y|U) + I(U; Z)$$
$$\overset{(c)}{\leq} I(V; Y) + H(Y|U) + I(U; Z), \qquad (79)$$

where (a) follows from Fano's inequality and (71); (b) follows from the chain-rule and the independence of M, $M_{\mathcal{Y}}^{(c)}$, and $M_{\mathcal{Z}}^{(c)}$; and (c) holds because conditioning cannot increase entropy.

So far, we have shown that the set of rate-tuples satisfying (2) for some PMF of the form (3) is an outer bound on the capacity region of the SD-BC with P-MSI. To conclude, it remains to establish that we can w.l.g. restrict X to be a function of (Y, U). To this end we note that, by the Functional Representation lemma (Lemma 1), there exists some $\hat{U}$ that is of finite support and independent of (Y, U) for which X is a function of $(Y, U, \hat{U})$. Further, we note that the RHS of each constraint in (2) is either unaffected or increases if we replace U by the pair $(U, \hat{U})$. From this we conclude that we can w.l.g. restrict X to be a function of (Y, U).

APPENDIX B
BINNING THE CLOUD-CENTER (OR RATE-SPLITTING) IS NECESSARY TO ACHIEVE THE CAPACITY REGION OF THEOREM 2

Throughout this section, we consider the case where the encoder conveys only private messages ($R = 0$). We show that with a coding scheme as in Appendix A-A but where the cloud-center is not binned, we cannot—in general—achieve the capacity region $\mathscr{C}_{\text{P-MSI}}$ of the SD-BC with P-MSI. Recall from the proof-sketch after Theorem 2 that—without binning the cloud-center—we can achieve every rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ that satisfies (2) and (4) for some PMF $p(v, u, x, y, z)$ of the form (3). Moreover, we can show that w.l.g. we can restrict X to be a function of (Y, U) (this follows from the Functional Representation lemma (Lemma 1); a similar argument can be found in Appendix A-B). Note that the rate-region that can be achieved without binning the cloud-center is contained in the capacity region $\mathscr{C}_{\text{P-MSI}}$ of the SD-BC with P-MSI. As the following example shows, the containment can be strict:

Example 23 (Binning the Cloud-Center): Consider the SD-BC with binary input X and binary outputs $Y = X$ and

$$Z = \begin{cases} X & \text{if } S = 0, \\ ? & \text{if } S = 1, \end{cases} \qquad (80)$$

where $S \sim \text{Ber}(p)$, $p \in [0, 1]$ *is independent of* X, *and assume that the encoder conveys only private messages ($R = 0$). By Corollary 5 (with $X \sim \text{Ber}(1/2)$) the capacity*

region $\mathscr{C}_{\text{P-MSI}}$ with F-MSI at the stochastic receiver $\mathcal{Z}$ ($R_{\mathcal{Y}}^{(p)} = 0$) is the set of rate-tuples $(0, R_{\mathcal{Y}}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ that satisfy

$$R_{\mathcal{Y}} \leq 1 \tag{81a}$$
$$R_{\mathcal{Z}} \leq 1 - p \tag{81b}$$
$$R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} \leq 1. \tag{81c}$$

This implies that, irrespective of $p \in [0, 1]$, we can achieve the rate-tuple

$$(0, R_{\mathcal{Y}}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) = (0, p, 1 - p, 0). \tag{82}$$

As we argue shortly, if we do not bin the cloud-center, then we can achieve the rate-tuple (82) only in the degenerate cases where $p \in \{0, 1\}$. This implies that, for every $p \in (0, 1)$, if we do not bin the cloud-center, then we can achieve only a strict subset of the capacity region.

We next show that, if we do not bin the cloud-center, then we can achieve the rate-tuple (82) only in the degenerate cases where $p \in \{0, 1\}$. To this end recall that—without binning the cloud-center—the achievable rate-tuples $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ are the ones that satisfy (2) and (4) for some PMF $p(v, u, x, y, z)$ of the form (3), and where one can w.l.g. restrict X to be a function of (Y, U). Fix some PMF $p(v, u, x, y, z)$ of the form (3) that satisfies that X is a function of (Y, U), and assume that $p > 0$. As we argue shortly, (2c) implies that, for every $p \in (0, 1]$, at least one of the following two holds:

$$R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} < H(X) \leq 1 \quad \text{or} \quad I(U; X|V) = 0. \tag{83}$$

Moreover, it follows from (4b) that

$$R_{\mathcal{Z}}^{(p)} \overset{(a)}{\leq} I(U; Z|V) \tag{84}$$
$$\overset{(b)}{=} I(U; Z, S|V) \tag{85}$$
$$\overset{(c)}{=} I(U; Z|S, V) \tag{86}$$
$$\overset{(d)}{=} (1 - p)I(U; X|V), \tag{87}$$

where (a) is (4b); (b) holds because S is a function of Z; (c) holds because S is independent of (V, U); and (d) follows from (80). From (83) and (87) it follows that, for every $p \in (0, 1]$, at least one of the following two holds:

$$R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} < H(X) \leq 1 \quad \text{or} \quad R_{\mathcal{Z}}^{(p)} = 0. \tag{88}$$

Consequently, for every $p \in (0, 1]$, if $R_{\mathcal{Z}}^{(p)}$ is strictly positive, then $R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)}$ must be strictly smaller than 1. In particular, this implies that if we do not bin the cloud-center, then we can achieve the rate-tuple $(0, R_{\mathcal{Y}}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) = (0, p, 1 - p, 0)$ only in the degenerate cases where $p \in \{0, 1\}$.

To conclude, it remains to show that, for every $p \in (0, 1]$, (2c) implies (83). To this end we observe from (2c) that

$$R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} \overset{(a)}{\leq} I(V; Y) + H(Y|U) + I(U; Z|V) \tag{89}$$
$$\overset{(b)}{\leq} I(X; Y, Z) \tag{90}$$
$$\overset{(c)}{=} H(X) \tag{91}$$
$$\overset{(d)}{\leq} 1, \tag{92}$$

where (a) is (2c); (b) follows from (102) in the proof of Corollary 5, which can be found in Appendix C; (c) holds because $Y = X$; and (d) holds because X is binary. Note that (92) can hold with equality only if the following two equalities hold

$$I(U; Y|Z, V) = I(V; Z|Y) = 0. \tag{93}$$

Indeed, Inequality (d) in the derivation of (102) holds with equality only if (93) holds, and hence Inequality (b) in the derivation of (92) holds with equality only if (93) holds. Note that

$$I(U; Y|Z, V) \overset{(a)}{=} I(U; X|Z, V) \tag{94}$$
$$\overset{(b)}{=} (1 - p)I(U; X|X, V) + p\, I(U; X|Z = ?, V)$$
$$= p\, I(U; X|V), \tag{95}$$

where (a) holds because $Y = X$; and (b) follows from (80). We are now ready to conclude the proof of our claim that, for every $p \in (0, 1]$, (2c) implies (83): We have shown that (2c) implies (92). Moreover, (92) can hold with equality only if (93) holds, and by (95) this implies that, for every $p \in (0, 1]$, (92) can hold with equality only if $I(U; X|V) = 0$. Consequently, (83) holds for every $p \in (0, 1]$.

APPENDIX C
PROOF OF COROLLARY 5

For $U = X$ and $V = Y$ the constraints in (2) and (9) are equivalent. Hence, (9) is an inner bound on the capacity region. We next argue that (9) is also an outer bound on the capacity region. To this end fix any PMF of the form (3) satisfying that X is a function of (Y, U). By (3) U, X, and Z form a Markov chain in that order, and hence

$$I(U; Z) \leq I(X; Z). \tag{96}$$

Moreover,

$$I(V; Y) + H(Y|U) + I(U; Z|V)$$
$$\overset{(a)}{=} H(Y|U) + I(U; Y, Z) - I(U; Y|V)$$
$$\quad - I(U; Z|Y) + I(U; Z|V) \tag{97}$$
$$\overset{(b)}{=} H(Y|U) + I(U; Y, Z) - I(U; Y|V)$$
$$\quad - I(X; Z|Y) + I(U; Z|V) \tag{98}$$
$$\overset{(c)}{=} H(Y|U) + I(U; Y, Z) - I(U; Y, Z|V)$$
$$\quad - I(V; Z|Y) + I(U; Z|V) \tag{99}$$
$$\overset{(d)}{\leq} H(Y|U) + I(U; Y, Z) \tag{100}$$
$$\overset{(e)}{=} H(Y) + I(U; Z|Y) \tag{101}$$
$$\overset{(f)}{=} I(X; Y, Z), \tag{102}$$

where (a) follows from the chain-rule and the fact that, by (3), V, U, and Y form a Markov chain in that order; (b) holds because X is a function of (Y, U) and because, by (3), U, X, and Z form a Markov chain in that order; (c) holds because

$$I(X; Z|Y) - I(U; Z|Y, V) = I(X; Z|Y) - I(X; Z|Y, V)$$
$$= I(V; Z|Y), \tag{103}$$

where we used that X is a function of (Y, U); (d) holds because mutual information is nonnegative and because conditioning cannot increase entropy; (e) follows from the chain-rule; and (f) holds because Y is a function of X, because X is a function of (Y, U), and because of the chain-rule. From (2a), (2b) and (96), as well as (2c) and (102) we conclude that (9) is an outer bound on the capacity region.

APPENDIX D
ANALYSIS OF EXAMPLE 10

- Assume F-MSI at $\mathcal{Y}$ and no MSI at $\mathcal{Z}$ ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} = 0$). The capacity region is given in Corollary 4, and it is not hard to see that the maximum achievable sum-rate $R + R_{\mathcal{Y}} + R_{\mathcal{Z}}$ is

$$\max_{p(u,x,y,z) \in \mathcal{P}_u} \{H(Y|U) + I(U; Z)\}. \tag{104}$$

In the proof of Example 18 (see Equations (135)–(137) in Appendix G), the RHS of (104) is shown to equal the RHS of (18). Further, it is shown that every PMF $p(u, x, y, z) \in \mathcal{P}_u^\star$ satisfies (19). Finally, we obtain (20) by examining (7).

- Assume no MSI ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$). By Theorem 2 the capacity region is the set of rate-tuples $(R, R_{\mathcal{Y}}, R_{\mathcal{Z}})$ satisfying

$$R + R_{\mathcal{Y}} \leq H(Y) \tag{105a}$$
$$R + R_{\mathcal{Z}} \leq I(U; Z) \tag{105b}$$
$$R + R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq I(V; Y) + H(Y|U)$$
$$+ I(U; Z|V) \tag{105c}$$
$$R + R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq H(Y|U) + I(U; Z) \tag{105d}$$
$$2R + R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq I(V; Y) + H(Y|U)$$
$$+ I(U; Z) \tag{105e}$$

for some PMF $p(v, u, x, y, z)$ of the form (3). From (105d) we see that the maximum achievable sum-rate cannot be larger than the maximum sum-rate in (104), and that it can be achieved only if $p(u, x, y, z) \in \mathcal{P}_u$. If we set $R = 0$ (no common message) and let V be deterministic, then we see that (104) is achievable, and hence we obtain that (104) is the maximum achievable sum-rate. Finally, (21) follows from (105e), because the maximum achievable sum-rate is (104).

APPENDIX E
PROOF OF PROPOSITION 14

We assume that $R_{\mathcal{Y}}^{(c)} > 0$, because otherwise the statement is obvious. We will show that the error probability of the described code with rate-limited feedback converges to zero as the blocklength approaches infinity. To this end we introduce the following three error events:

E_0 : Receiver $\mathcal{Y}$ cannot recover $Z^{\alpha n}$ $\qquad$ (106)

E_1 : The messages sent in Phase 1 are not decoded

correctly $\qquad$ (107)

E_2 : The messages sent in Phase 2 (including the

compression index) are not decoded

correctly. $\qquad$ (108)

Using these error events, we can upper-bound the code's error probability by

$$\mathbb{P}[\text{error}] \leq \mathbb{P}[E_0 \cup E_1 \cup E_2] \tag{109}$$
$$= \mathbb{P}[E_2] + \mathbb{P}[E_0|E_2^c] + \mathbb{P}[E_1|E_0^c]. \tag{110}$$

To show that the code's error probability converges to zero, we will show that each term on the RHS of (110) converges to zero. For the third term we have

$$\mathbb{P}[E_1|E_0^c] \to 0 \ (n \to \infty), \tag{111}$$

because $(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Y}}^{(c)}, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}^{(\text{enh})}$. Also, because Event E_2^c occurs only if Receiver $\mathcal{Y}$ recovers the correct compression index, and because a compression index of rate $H(Z|Y)$ is enough for Receiver $\mathcal{Y}$ to recover $Z^{\alpha n}$ using also the side-information $Y^{\alpha n}$, the second term satisfies

$$\mathbb{P}[E_0|E_2^c] \to 0 \ (n \to \infty). \tag{112}$$

Finally, we note that the first term satisfies

$$\mathbb{P}[E_2] \to 0 \ (n \to \infty), \tag{113}$$

because

$$\left(R_{\mathcal{Y}}^{(p)}, \hat{R}_{\mathcal{Y}}^{(c)} + \frac{\alpha}{1-\alpha} H(Z|Y), R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}\right) \in \mathscr{C}_{\text{P-MSI}}, \tag{114}$$

where we used (30) and that $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}$. (The "rate" $\frac{\alpha}{1-\alpha} H(Z|Y)$ accounts for the $\alpha n H(Z|Y)$ compression bits from Phase 1 that have to be sent during the $(1-\alpha)n$ channel uses that Phase 2 comprises.)

APPENDIX F
PROOF OF PROPOSITIONS 15 AND 16

Before we prove Propositions 15 and 16, we sketch our proof of Proposition 15. A first step in the proof is to note that—under the conditions stated in Proposition 15—it is possible to identify a rate-triple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ that satisfies the following two:

1) $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)} + R_{\mathcal{Z}}^{(c)})$ lies on the boundary of the no-feedback capacity region $\mathscr{C}$ of the SD-BC without MSI and in the interior of the capacity region $\mathscr{C}_{\text{enh}}$ of the enhanced BC without MSI.
2) There exists a strictly-positive rate $R_{\mathcal{Y}}^{(c)}$ for which $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ is contained in $\mathscr{C}_{\text{P-MSI}}$.

As we argue next, the above conditions guarantee that, irrespective of $R_{\text{FB}} > 0$, the feedback code of Section IV-B achieves a rate-tuple outside the no-feedback capacity region $\mathscr{C}$ of the SD-BC without MSI. By Remark 7 $\mathscr{C}$ is also the no-feedback capacity region of the SD-BC with P-MSI at Receiver $\mathcal{Y}$ only, and hence it follows that the conditions in Proposition 15 are sufficient for feedback to increase the capacity region of the SD-BC with P-MSI at Receiver $\mathcal{Y}$ only. It remains to show that the feedback code of Section IV-B achieves a rate-tuple outside the no-feedback capacity region $\mathscr{C}$. Condition 1 guarantees that in Phase 1 of the feedback code, when the encoder codes for the enhanced BC, message information can be sent at rates $(\tilde{R}_{\mathcal{Y}}, 0, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)})$ satisfying

$$\tilde{R}_{\mathcal{Y}} > R_{\mathcal{Y}}^{(p)}, \quad \tilde{R}_{\mathcal{Z}}^{(p)} > R_{\mathcal{Z}}^{(p)}, \quad \tilde{R}_{\mathcal{Z}}^{(c)} > R_{\mathcal{Z}}^{(c)}. \tag{115}$$

Recall that the deterministic receiver $\mathcal{Y}$ needs resolution information to decode its Phase-1 message. Condition 2 guarantees that in Phase 2 of the feedback code the encoder can send the required resolution information at rate $R_{\mathcal{Y}}^{(c)}$ and, simultaneously, fresh message-information at rates $(R_{\mathcal{Z}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$. Specifically, Proposition 14 implies that, for every sufficiently-small α, the rate-tuple

$$\alpha(\tilde{R}_{\mathcal{Y}}, 0, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}) + (1-\alpha)(R_{\mathcal{Y}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \quad (116)$$

is achievable. Since $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)} + R_{\mathcal{Z}}^{(c)})$ lies on the boundary of the no-feedback capacity region $\mathscr{C}$, we conclude from (115) that the rate-tuple (116) is not achievable without feedback. This proves that the feedback code of Section IV-B achieves a rate-tuple outside the no-feedback capacity region $\mathscr{C}$.

The proof of Proposition 16 is similar: the main difference is that we set $R_{\mathcal{Z}}^{(c)} = \tilde{R}_{\mathcal{Z}}^{(c)} = 0$.

We next provide the details of the proof of Propositions 15 and 16. We shall use the following lemma:

Lemma 24: Fix a rate-tuple $(R_{\mathcal{Y}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}$. If $R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$ (or, equivalently, $R_{\mathcal{Z}}^{(c)} > 0$), and if for some PMF $p(u, x, y, z)$ of the form (6) we have

$$R_{\mathcal{Y}}^{(p)} < H(Y) \quad (117a)$$
$$R_{\mathcal{Z}}^{(p)} + R_{\mathcal{Z}}^{(c)} \le I(U; Z) \quad (117b)$$
$$R_{\mathcal{Y}}^{(p)} + R_{\mathcal{Z}}^{(p)} + R_{\mathcal{Z}}^{(c)} \le H(Y|U) + I(U; Z) \quad (117c)$$
$$I(U; Y) > 0, \quad (117d)$$

then there exists some positive rate $R_{\mathcal{Y}}^{(c)} > 0$ satisfying

$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}. \quad (118)$$

If $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$ (or, equivalently, $R_{\mathcal{Z}}^{(c)} = 0$), and if for some PMF $p(v, u, x, y, z)$ of the form (3) we have (117) and

$$I(V; Y) - I(V; Z) > 0, \quad (119)$$

then there exists some positive rate $R_{\mathcal{Y}}^{(c)} > 0$ satisfying (118).

Proof of Lemma 24: Suppose that (117) holds for some PMF $p(u, x, y, z)$ of the form (6). A rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ is in $\mathscr{C}_{\text{P-MSI}}$ if it satisfies (2) for some conditional PMF $p(v|u)$ and

$$p(v, u, x, y, z) = p(v|u) \, p(u, x, y, z). \quad (120)$$

By inspection of (2) and (117), we see that this holds if

$$R_{\mathcal{Y}}^{(c)} \le H(Y) - R_{\mathcal{Y}}^{(p)} \quad (121a)$$
$$R_{\mathcal{Y}}^{(c)} \le H(Y|U) + I(U; Z) + I(V; Y)$$
$$\quad - I(V; Z) - R_{\mathcal{Y}}^{(p)} - R_{\mathcal{Z}}^{(p)} \quad (121b)$$
$$R_{\mathcal{Y}}^{(c)} \le H(Y|U) + I(U; Z) + I(V; Y)$$
$$\quad - R_{\mathcal{Y}}^{(p)} - R_{\mathcal{Z}}^{(p)} - R_{\mathcal{Z}}^{(c)}. \quad (121c)$$

And from (117) it follows that if

$$I(V; Z) - I(V; Y) < R_{\mathcal{Z}}^{(c)} \quad \text{and} \quad I(V; Y) > 0, \quad (122)$$

then there exists some $R_{\mathcal{Y}}^{(c)} > 0$ for which the rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfies (121). This proves the claim for the case where $R_{\mathcal{Z}}^{(p)} = R_{\mathcal{Z}}$ (or, equivalently, $R_{\mathcal{Z}}^{(c)} = 0$), because in this case (122) and (119) are equivalent. To prove the claim for the case where $R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$ (or, equivalently, $R_{\mathcal{Z}}^{(c)} > 0$), fix $\varepsilon \in (0, 1)$, let $S \sim \text{Ber}(1 - \varepsilon)$ be independent of (U, X, Y, Z), and choose

$$V = \begin{cases} U & S = 0, \\ ? & S = 1. \end{cases} \quad (123)$$

For this choice of V (122) holds for every sufficiently-small ε, because

$$I(V; Z) = \varepsilon I(U; Z) \le \varepsilon \log |\mathcal{Z}| \quad \text{and}$$
$$0 < I(V; Y) = \varepsilon I(U; Y). \quad (124)$$

This proves the claim for the case where $R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$. $\quad\square$

Note that if $R_{\mathcal{Z}}^{(p)} < R_{\mathcal{Z}}$, then Lemma 24 does not ask for much: By the assumption that $(R_{\mathcal{Y}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}$, there must exist some PMF $p(u, x, y, z)$ of the form (6) for which (117b) and (117c) are satisfied and (117a) holds with nonstrict inequality (this follows from Remark 7 and Corollary 3). Hence, all we are asking for is that $R_{\mathcal{Y}} < H(Y)$ and $I(U; Y) > 0$ hold. (Roughly speaking, all we are asking for is that the transmission of $M_{\mathcal{Z}}$ interfere with the transmission of $M_{\mathcal{Y}}$.)

Proof of Proposition 15 and 16: Assume that there exists some rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ that satisfies the following three conditions[14]:

$$R_{\mathcal{Y}}^{(c)} > 0 \quad (125a)$$
$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}} \quad (125b)$$
$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)} + R_{\mathcal{Z}}^{(c)}) \in (\partial\mathscr{C} \cap (\mathscr{C}_{\text{enh}} \setminus \partial\mathscr{C}_{\text{enh}})). \quad (125c)$$

Now look for rates

$$\tilde{R}_{\mathcal{Y}}^{(p)} > R_{\mathcal{Y}}^{(p)}, \quad \tilde{R}_{\mathcal{Z}}^{(p)} > R_{\mathcal{Z}}^{(p)}, \quad \tilde{R}_{\mathcal{Z}}^{(c)} > R_{\mathcal{Z}}^{(c)} \quad (126)$$

satisfying that $(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(p)} + \tilde{R}_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{enh}}$. By Assumption (125c) such rates exist.

Proposition 14 and Assumptions (125a) and (125b) guarantee that, irrespective of $R_{\text{FB}} > 0$, there exists some $\alpha \in (0, 1)$ for which the rate-tuple

$$\alpha(\tilde{R}_{\mathcal{Y}}^{(p)}, 0, \tilde{R}_{\mathcal{Z}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(c)}) + (1-\alpha)(R_{\mathcal{Y}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \quad (127)$$

is in the feedback capacity region of the SD-BC with P-MSI at Receiver $\mathcal{Y}$. Note that (13), (125c), and (126) imply that the rate-tuple in (127) is not in $\mathscr{C}_{\text{P-MSI}}$ (see also Footnote 14). To conclude the proof of Proposition 15, it thus suffices to show that, under the conditions in Proposition 15, it is possible to find a rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying (125). But this follows from the first part of Lemma 24.

---

[14]In particular, (125c) implies that $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \partial\mathscr{C}_{\text{P-MSI}}^{(R_{\mathcal{Y}}^{(c)}=0)}$, where $\mathscr{C}_{\text{P-MSI}}^{(R_{\mathcal{Y}}^{(c)}=0)} \subset (\mathbb{R}_0^+)^3$ denotes the set of rate-triples $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying $(R_{\mathcal{Y}}^{(p)}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)}) \in \mathscr{C}_{\text{P-MSI}}$ (cf. (13)).

The proof of Proposition 16 is similar. Assume that there exists some rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, 0)$ that satisfies the following three conditions:

$$R_{\mathcal{Y}}^{(c)} > 0 \tag{128a}$$

$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, 0) \in \mathscr{C}_{\text{P-MSI}} \tag{128b}$$

$$(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)}) \in \left(\partial\mathscr{C} \cap (\mathscr{C}_{\text{enh}} \setminus \partial\mathscr{C}_{\text{enh}})\right). \tag{128c}$$

Now look for rates

$$\tilde{R}_{\mathcal{Y}}^{(p)} > R_{\mathcal{Y}}^{(p)}, \quad \tilde{R}_{\mathcal{Z}}^{(p)} > R_{\mathcal{Z}}^{(p)} \tag{129}$$

satisfying that $(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(p)}) \in \mathscr{C}_{\text{enh}}$. By Assumption (128c) such rates exist.

Proposition 14 and Assumptions (128a) and (128b) guarantee that, irrespective of $R_{\text{FB}} > 0$, there exists some $\alpha \in (0, 1)$ for which the rate-pair

$$\alpha(\tilde{R}_{\mathcal{Y}}^{(p)}, \tilde{R}_{\mathcal{Z}}^{(p)}) + (1 - \alpha)(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Z}}^{(p)}) \tag{130}$$

is in the feedback capacity region of the SD-BC with P-MSI at Receiver $\mathcal{Y}$. Note that (128c) and (129) imply that the rate-tuple in (130) is not in $\mathscr{C}$. To conclude the proof of Proposition 16, it thus suffices to show that, under the conditions in Proposition 16, it is possible to find a rate-tuple $(R_{\mathcal{Y}}^{(p)}, R_{\mathcal{Y}}^{(c)}, R_{\mathcal{Z}}^{(p)}, 0)$ satisfying (128). But this follows from the second part of Lemma 24. $\square$

APPENDIX G
ANALYSIS OF EXAMPLE 18

The proof of Example 18 hinges on Propositions 15 and 16, which state sufficient conditions for feedback to increase the capacity region of the SD-BC with and without P-MSI at Receiver $\mathcal{Y}$. It can be roughly outlined as follows: Let $\bar{p} = 1 - p$. For the SD-BC of Example 18, we first parametrize $\partial\mathscr{C} \cap \{R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p}]\}$ by $R_{\mathcal{Z}}$, where $R_{\mathcal{Z}}^*$ denotes the maximum rate $R_{\mathcal{Z}}$ for which we can achieve the sum-rate capacity without MSI, and where $\bar{p}$ is the maximum rate $R_{\mathcal{Z}}$ that is achievable without MSI. We then show that for the boundary point with $R_{\mathcal{Z}} = R_{\mathcal{Z}}^*$ we can satisfy all the conditions in Proposition 15, and that for some boundary point with $R_{\mathcal{Z}} \in (R_{\mathcal{Z}}^*, \bar{p})$ we can satisfy all the conditions in Proposition 16. This allows us to conclude from Propositions 15 and 16 that—on the considered SD-BC—feedback increases the sum-rate capacity with P-MSI at Receiver $\mathcal{Y}$ and the capacity region without MSI.

Recall that the no-feedback capacity region $\mathscr{C}$ of the SD-BC without MSI is the set of rate-tuples satisfying (5) for some PMF of the form (6), and where we can restrict X to be a function of (Y, U). For the SD-BC of Example 18, the fact that X is a function of (Y, U) implies that we can partition the support $\mathcal{U}$ of U into two disjoint sets $\mathcal{U}_0$ and $\mathcal{U}_1$, where

$$p_{X_2, Y|U}(1, 1|u) = 0, \quad \forall u \in \mathcal{U}_0 \quad \text{and}$$
$$p_{X_2, Y|U}(0, 1|u) = 0, \quad \forall u \in \mathcal{U}_1. \tag{131}$$

For each $u \in \mathcal{U}_0$ introduce

$$p_0(u) = p_{X_2, Y|U}(0, 0|u) \tag{132a}$$
$$p_1(u) = p_{X_2, Y|U}(0, 1|u) \tag{132b}$$
$$p_2(u) = p_{X_2, Y|U}(1, 2|u), \tag{132c}$$

and note that $p_0(u) + p_1(u) + p_2(u) = 1$. As we argue next, we can w.l.g. assume that to each $u \in \mathcal{U}_0$ there corresponds some $u' \in \mathcal{U}_1$ satisfying $p_U(u) = p_U(u')$ and

$$p_0(u) = p_{X_2, Y|U}(0, 0|u) = p_{X_2, Y|U}(1, 2|u') \tag{133a}$$
$$p_1(u) = p_{X_2, Y|U}(0, 1|u) = p_{X_2, Y|U}(1, 1|u') \tag{133b}$$
$$p_2(u) = p_{X_2, Y|U}(1, 2|u) = p_{X_2, Y|U}(0, 0|u'), \tag{133c}$$

This allows us to simplify (5b) and (5c) to

$$R_{\mathcal{Z}} \leq \bar{p}\left[1 - 2\sum_{u \in \mathcal{U}_0} p(u)h_{\text{b}}(p_2(u))\right] \tag{134a}$$

$$R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq \bar{p} + 2\sum_{u \in \mathcal{U}_0} p(u)\bigg[p\, h_{\text{b}}(p_2(u)) + (1 - p_2(u))h_{\text{b}}\left(\frac{p_0(u)}{p_0(u) + p_1(u)}\right)\bigg]. \tag{134b}$$

To show that to each $u \in \mathcal{U}_0$ there corresponds some $u' \in \mathcal{U}_1$ satisfying $p_U(u) = p_U(u')$ and (133), we first note that all $u \in \mathcal{U}_0$ and $u' \in \mathcal{U}_1$ satisfying (133) must satisfy $H(X_2|U = u) = H(X_2|U = u')$ and $H(Y|U = u) = H(Y|U = u')$. Using this, that $I(U; Z) = \bar{p}\, I(U; X_2)$, and that entropy is concave, we can now argue that the claim must hold by symmetry.

Having obtained (134), we are now ready to determine the sum-rate capacity. First, note that the RHS of (134b) is maximum only if $p_0(u) = p_1(u) = (1 - p_2(u))/2$, and that in this case (134b) simplifies to

$$R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq \bar{p} + 2\sum_{u \in \mathcal{U}_0} p(u)\bigg[p\, h_{\text{b}}(p_2(u)) + (1 - p_2(u))\bigg]. \tag{135}$$

Because the function $p_2(u) \mapsto p\, h_{\text{b}}(p_2(u)) + (1 - p_2(u))$ is strictly concave in $p_2(u)$, we readily find that the RHS of (135) is maximum if, and only if, (iff) for all $u \in \mathcal{U}_0$ we have

$$p_2(u) = \frac{1}{1 + 2^{1/p}}. \tag{136}$$

Hence, we find that the sum-rate capacity is obtained by evaluating the RHS of (135) for the choice (136), and that every PMF $p(u, x, y, z)$ of the form (6) that achieves the maximum sum-rate must satisfy the following two:

$$0 < I(U; Y) < I(U; Z) \tag{137a}$$

$$R_{\mathcal{Z}} \leq I(U; Z) = \bar{p}\left[1 - h_b\left(\frac{1}{1 + 2^{1/p}}\right)\right] = R_{\mathcal{Z}}^*, \tag{137b}$$

where $R_{\mathcal{Z}}^*$ is the maximum rate $R_{\mathcal{Z}}$ for which we can achieve the sum-rate capacity without MSI.

We next parametrize $\partial\mathscr{C} \cap \{R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p}]\}$ by $R_{\mathcal{Z}}$, i.e., for each $R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p}]$ we determine

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = \max\{R_{\mathcal{Y}} \geq 0 \colon (R_{\mathcal{Y}}, R_{\mathcal{Z}}) \in \mathscr{C}\}. \tag{138}$$

To this end we first show that $R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}}$ is strictly decreasing in $R_{\mathcal{Z}}$. Indeed, because $\mathscr{C}$ is convex we have for all $\epsilon \in (0, R_{\mathcal{Z}} - R_{\mathcal{Z}}^*]$, for $\alpha = \epsilon/(R_{\mathcal{Z}} - R_{\mathcal{Z}}^*)$, and for $\bar{\alpha} = 1 - \alpha$

$$R_{\mathcal{Y}}(R_{\mathcal{Z}} - \epsilon) + R_{\mathcal{Z}} - \epsilon$$
$$\geq \bar{\alpha}(R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}}) + \alpha(R_{\mathcal{Y}}(R_{\mathcal{Z}}^*) + R_{\mathcal{Z}}^*) \tag{139}$$
$$> R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}}, \tag{140}$$

where the last inequality holds because $R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}}$ is maximum if $R_{\mathcal{Z}} = R_{\mathcal{Z}}^*$ and strictly smaller than its maximum if $R_{\mathcal{Z}} > R_{\mathcal{Z}}^*$. As we argue next, each pair $(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}})$ satisfies

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = H(Y|U) \tag{141a}$$
$$R_{\mathcal{Z}} = I(U; Z) \tag{141b}$$

for some PMF $p(u, x, y, z)$ of the form (6). Note that the claim holds for $R_{\mathcal{Z}} = R_{\mathcal{Z}}^*$, and hence we assume that $R_{\mathcal{Z}}^* < R_{\mathcal{Z}}$. Because $(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}}) \in \partial\mathscr{C}$, Corollary 3 implies that for some PMF $p(u, x, y, z)$ of the form (6) we must have

$$R_{\mathcal{Z}} \le I(U; Z) \tag{142}$$
$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = \min\{H(Y|U) + I(U; Z) - R_{\mathcal{Z}}, H(Y)\}. \tag{143}$$

For contradiction, assume that $R_{\mathcal{Z}} < I(U; Z)$. Then,

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}} \overset{(a)}{\ge} \bar{\alpha}\big(H(Y|U) + I(U; Z)\big)$$
$$+ \alpha\big(R_{\mathcal{Y}}(R_{\mathcal{Z}}^*) + R_{\mathcal{Z}}^*\big) \tag{144}$$
$$\overset{(b)}{>} R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}}, \tag{145}$$

where (a) holds for $\alpha = \big(I(U; Z) - R_{\mathcal{Z}}\big)/\big(I(U; Z) - R_{\mathcal{Z}}^*\big)$ and $\bar{\alpha} = 1 - \alpha$, because $R_{\mathcal{Z}} = \bar{\alpha} I(U; Z) + \alpha R_{\mathcal{Z}}^*$, and because the capacity region is convex; and (b) is true because $R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}} \le H(Y|U) + I(U; Z)$, because $\alpha > 0$, and because $R_{\mathcal{Z}}^* < R_{\mathcal{Z}}$ and $R_{\mathcal{Y}}(R_{\mathcal{Z}}) + R_{\mathcal{Z}}$ is strictly decreasing in $R_{\mathcal{Z}}$. This is a contradiction, and hence the claim follows.

From (134), (135), and (141) we obtain that each pair $(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}})$, $R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p}]$ must be of the form

$$R_{\mathcal{Z}} = \bar{p}\left[1 - 2\sum_{u \in \mathcal{U}_0} p(u)h_{\mathrm{b}}\big(p_2(u)\big)\right] \tag{146a}$$

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = \bar{p} + 2\sum_{u \in \mathcal{U}_0} p(u)\Big[p\, h_{\mathrm{b}}\big(p_2(u)\big) + 1$$
$$- p_2(u)\Big] - R_{\mathcal{Z}}. \tag{146b}$$

This can also be written as

$$R_{\mathcal{Z}} = \sum_{u \in \mathcal{U}_0} 2p(u)\bar{p}\Big[1 - h_{\mathrm{b}}\big(p_2(u)\big)\Big] \tag{147a}$$

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = \sum_{u \in \mathcal{U}_0} 2p(u)\Big[h_{\mathrm{b}}\big(p_2(u)\big) + 1 - p_2(u)\Big], \tag{147b}$$

where we used that

$$\sum_{u \in \mathcal{U}_0} 2p(u) = 1. \tag{148}$$

Note that if $p_2(u) > 1/2$, then replacing $p_2(u)$ by $1 - p_2(u)$ in (147) does not affect the value of $R_{\mathcal{Z}}$ but increases that of $R_{\mathcal{Y}}(R_{\mathcal{Z}})$. Hence, for each $u \in \mathcal{U}_0$ satisfying that $p(u) > 0$ we have $p_2(u) \in [0, 1/2]$, and hence we can w.l.g. assume that $p_2(u) \in [0, 1/2]$ holds for all $u \in \mathcal{U}_0$. Since $h_{\mathrm{b}}(\cdot)$ is invertible on $[0, 1/2]$ and $\bar{p}$ is a constant, we can thus write

$$R_{\mathcal{Z}} = \sum_{u \in \mathcal{U}_0} 2p(u)R_{\mathcal{Z}}^{(u)} \tag{149a}$$

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = \sum_{u \in \mathcal{U}_0} 2p(u)\left[2 - \frac{1}{\bar{p}}R_{\mathcal{Z}}^{(u)} - h_{\mathrm{b}}^{-1}\left(1 - \frac{R_{\mathcal{Z}}^{(u)}}{\bar{p}}\right)\right], \tag{149b}$$

where $R_{\mathcal{Z}}^{(u)} = \bar{p}\big[1 - h_{\mathrm{b}}(p_2(u))\big]$. Note that $R_{\mathcal{Z}}^{(u)}$ can assume any value in $[0, \bar{p}]$ depending on $p_2(u) \in [0, 1/2]$. Since $h_{\mathrm{b}}(\cdot)$ is strictly increasing and strictly concave on $[0, 1/2]$, its inverse $h_{\mathrm{b}}^{-1}(\cdot)$ is strictly increasing and strictly convex on $[0, 1]$. In particular, this implies that the mapping

$$x \mapsto 2 - \frac{1}{\bar{p}}x - h_{\mathrm{b}}^{-1}\left(1 - \frac{x}{\bar{p}}\right) \tag{150}$$

is strictly concave on $[0, 1]$. Recalling that $R_{\mathcal{Y}}(R_{\mathcal{Z}})$ is the maximum rate $R_{\mathcal{Y}}$ that is achievable for a given $R_{\mathcal{Z}}$, we thus obtain from Jensen's inequality that for each $u \in \mathcal{U}_0$ satisfying that $p(u) > 0$ we must have $R_{\mathcal{Z}}^{(u)} = R_{\mathcal{Z}}$. This allows us to parametrize $\partial\mathscr{C} \cap \{R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p}]\}$ by $R_{\mathcal{Z}}$:

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = 2 - \frac{1}{\bar{p}}R_{\mathcal{Z}} - h_{\mathrm{b}}^{-1}\left(1 - \frac{R_{\mathcal{Z}}}{\bar{p}}\right),$$
$$R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p}]. \tag{151}$$

With (151) at hand, we next argue that for $p > 1/2$ there exists some rate-pair $(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}})$ with $R_{\mathcal{Z}} < \bar{p}$ satisfying that (5) and (34) of Proposition 16 hold for some PMF $p(u, x, y, z)$ of the form (6) and for $V = U$. Because $(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}}) \in \partial\mathscr{C}$ and because for $V = U$ the PMF $p(v, u, x, y, z)$ is of the form (3), this is almost enough to conclude from Proposition 16 that feedback increases the capacity region without MSI: once we have established the claim, all that remains to be shown is that the identified rate-pair satisfies

$$\big(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}}\big) \in \mathscr{C}_{\mathrm{enh}} \setminus \partial\mathscr{C}_{\mathrm{enh}}. \tag{152}$$

To establish the claim, let $\mathcal{U}_0 = \{u\}$. In particular, this implies that $\mathcal{U}_1 = \{u'\}$, where (133) holds and $p(u) = p(u') = 1/2$. For every $R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p}]$ set $p_2(u) = h_{\mathrm{b}}^{-1}(1 - R_{\mathcal{Z}}/\bar{p})$ and

$$\frac{1 - p_2(u)}{2} = p_{X_2,Y|U}(0, 0|u) = p_{X_2,Y|U}(1, 2|u') \tag{153a}$$
$$\frac{1 - p_2(u)}{2} = p_{X_2,Y|U}(0, 1|u) = p_{X_2,Y|U}(1, 1|u') \tag{153b}$$
$$p_2(u) = p_{X_2,Y|U}(1, 2|u) = p_{X_2,Y|U}(0, 0|u'). \tag{153c}$$

For this choice we have

$$R_{\mathcal{Y}}(R_{\mathcal{Z}}) = H(Y|U) = h_{\mathrm{b}}\big(p_2(u)\big) + 1 - p_2(u) \tag{154a}$$
$$R_{\mathcal{Z}} = I(U; Z) = \bar{p}\Big(1 - h_{\mathrm{b}}\big(p_2(u)\big)\Big), \tag{154b}$$

Moreover, it holds that

$$H(Y) = h_{\mathrm{b}}\left(\frac{1 - p_2(u)}{2}\right) + \frac{1 + p_2(u)}{2}, \tag{155}$$

and from (154a) we thus obtain

$$I(U; Y) = H(Y) - H(Y|U) \tag{156}$$
$$= h_{\mathrm{b}}\left(\frac{1 - p_2(u)}{2}\right) - h_{\mathrm{b}}\big(p_2(u)\big)$$
$$+ \frac{3p_2(u) - 1}{2}. \tag{157}$$

Note that for all $p_2(u) \in \big[0, 1/(1 + 2^{1/p})\big]$ we have

$$R_{\mathcal{Y}} < H(Y). \tag{158}$$

Moreover, if $p > 1/2$, then we obtain for $p_2(u) = 0$ that

$$I(U; Z) = \bar{p} < \frac{1}{2} = I(U; Y). \tag{159}$$

Since $I(U; Y)$ and $I(U; Z)$ are continuous in $p_2(u)$, this implies that there exists an open interval $(0, \kappa) \subset \left(0, 1/(1 + 2^{1/p})\right)$ for which

$$I(U; Y) > I(U; Z), \quad p_2(u) \in (0, \kappa). \tag{160}$$

Indeed, this implies our claim: for $p > 1/2$ there exists a rate-pair $\left(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}}\right)$ with $R_{\mathcal{Z}} < \bar{p}$ for which (5) and (34) of Proposition 16 hold for some PMF $p(u, x, y, z)$ of the form (6) and for $V = U$.

As we argue next, to conclude the analysis of Example 18 it now suffices to show that on the enhanced BC $\left(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}}\right) \in \mathscr{C}_{\text{enh}} \setminus \partial \mathscr{C}_{\text{enh}}$ holds for all $R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p})$. Indeed, we already showed that this is enough to conclude from Proposition 16 that feedback increases the capacity region without MSI. It is, moreover, enough to conclude from Proposition 15 that feedback increases the sum-rate capacity with P-MSI at Receiver $\mathcal{Y}$, because $\left(R_{\mathcal{Y}}(R_{\mathcal{Z}}^*), R_{\mathcal{Z}}^*\right)$ is in $\partial \mathscr{C}$ and satisfies (5) and (32) for some PMF of the form (6), where we used (137a) and (141) to obtain that it satisfies (32).

To conclude, it now remains to show that on the enhanced BC $\left(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}}\right) \in \mathscr{C}_{\text{enh}} \setminus \partial \mathscr{C}_{\text{enh}}$ holds for all $R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p})$. Let $U \sim \text{Ber}(1/2)$ be a binary random variable, and set

$$\frac{1 - p_2(u)}{2} = p_{X_2, Y|U}(0, 0|0) = p_{X_2, Y|U}(1, 2|1) \tag{161a}$$

$$\frac{1 - p_2(u)}{2} = p_{X_2, Y|U}(0, 1|0) = p_{X_2, Y|U}(1, 1|1) \tag{161b}$$

$$\epsilon p_2(u) = p_{X_2, Y|U}(1, 1|0) = p_{X_2, Y|U}(0, 1|1) \tag{161c}$$

$$(1 - \epsilon) p_2(u) = p_{X_2, Y|U}(1, 2|0) = p_{X_2, Y|U}(0, 0|1), \tag{161d}$$

where ϵ and $p_2(u)$ take values in the set $[0, 1]$. If $\epsilon = 0$ and $p_2(u) \in \left[0, 1/(1 + 2^{1/p})\right]$, then, under the above PMF $p(u, x_2, y)$, (5) and (25) both evaluate to (154). Note that

$$I(U; Z) = \bar{p}\left[1 - h_{\text{b}}\left(p_2(u)\right)\right] \tag{162}$$

does not depend on ϵ, but that $I(X; Y, Z|U)$ does. Define

$$f\left(\epsilon, p_2(u)\right) = I(X; Y, Z|U), \tag{163}$$

and note that it satisfiess

$$f\left(\epsilon, p_2(u)\right)$$
$$= I(X; Y, Z|U) \tag{164}$$
$$= \bar{p} H(X_1, X_2|U) + p H(Y|U)$$
$$= \bar{p}\left[h_{\text{b}}\left(p_2(u)\right) + 1 - p_2(u) + p_2(u) h_b(\epsilon)\right]$$
$$\quad - p\left[\frac{1 - p_2(u)}{2} \log\left(\frac{1 - p_2(u)}{2}\right)\right.$$
$$\quad + \left(\frac{1 - (1 - 2\epsilon) p_2(u)}{2}\right) \log\left(\frac{1 - (1 - 2\epsilon) p_2(u)}{2}\right)$$
$$\quad \left. + (1 - \epsilon) p_2(u) \log\left((1 - \epsilon) p_2(u)\right)\right]. \tag{165}$$

Note that for all $p_2(u) \in \left(0, 1/(1 + 2^{1/p})\right]$ we have

$$\frac{\partial f\left(\epsilon, p_2(u)\right)}{\partial \epsilon}$$
$$= \bar{p}\, p_2(u) \log\left(\frac{1 - \epsilon}{\epsilon}\right)$$
$$\quad + p\, p_2(u) \log\left(\frac{2\,\bar{\epsilon}\, p_2(u)}{1 - (1 - 2\epsilon) p_2(u)}\right) \tag{166}$$
$$\to \infty \quad (\epsilon \downarrow 0). \tag{167}$$

In particular, this implies that there exists some $\epsilon > 0$ so that $f\left(\epsilon, p_2(u)\right) > f\left(0, p_2(u)\right)$. Since the capacity region of the enhanced BC contains every rate-pair $(\tilde{R}_{\mathcal{Y}}, \tilde{R}_{\mathcal{Z}})$ that for some $\epsilon > 0$ satisfies

$$\tilde{R}_{\mathcal{Y}} = I(X; Y, Z|U) \tag{168a}$$
$$\tilde{R}_{\mathcal{Z}} = I(U; Z), \tag{168b}$$

this proves our claim that for all $R_{\mathcal{Z}} \in [R_{\mathcal{Z}}^*, \bar{p})$ we have $\left(R_{\mathcal{Y}}(R_{\mathcal{Z}}), R_{\mathcal{Z}}\right) \in \mathscr{C}_{\text{enh}} \setminus \partial \mathscr{C}_{\text{enh}}$.

Remark 25: To prove that feedback can increase the capacity region of the SD-BC without MSI, we used Proposition 16 with the choice $V = U$. This raises the question whether choosing $V \neq U$ can help. It turns out that it can. To see this extend the SD-BC of Example 18 by a parallel channel of capacity larger than $\log 3$, and assume that Receiver $\mathcal{Z}$ additionally observes the output of this parallel channel. (E.g., assume that, in addition to Z, Receiver $\mathcal{Z}$ noiselessly observes an input X_3, which can assume 4 different values.) For the constructed SD-BC, it is easy to see that a rate-tuple satisfying (5) for some PMF of the form (6) is a boundary point of $\mathscr{C}$ only if the auxiliary random variable U comprises a capacity-achieving input to the parallel channel. (E.g., if Receiver $\mathcal{Z}$ observes the pair (Z, X_3), then X_3 must have a uniform prior and be deterministic given U.) But this implies that $I(U; Y) \leq \log 3 < I(U; Z)$. Hence, we cannot invoke Proposition 16 with the choice $V = U$ to show that feedback can increase the capacity region. However, we can invoke Proposition 16 with the following choice of V: choose V to be the random variable that we obtain when we discard the input to the parallel channel from U.

APPENDIX H
PROOF OF THEOREM 19

To prove the theorem, we show that every rate-tuple in the feedback capacity region of the SD-BC with F-MSI at the stochastic receiver $\mathcal{Z}$ satisfies (9) (with $R = 0$) for some PMF of the form (10). Let $Q \sim \text{Unif}[1 : n]$ be independent of $(M_{\mathcal{Y}}, M_{\mathcal{Z}})$, and introduce $(X, Y, Z) = (X_Q, Y_Q, Z_Q)$.

The rate of message $M_{\mathcal{Y}}$ satisfies

$$R_{\mathcal{Y}} - \epsilon_n \overset{(a)}{\leq} \frac{1}{n} I\left(M_{\mathcal{Y}}; Y^n, M_{\mathcal{Z}}^{(c)}\right) \tag{169}$$
$$\overset{(b)}{\leq} H\left(Y_Q | M_{\mathcal{Z}}^{(c)}, Y^{Q-1}, Q\right) \tag{170}$$
$$\overset{(c)}{\leq} H(Y), \tag{171}$$

where (a) follows from Fano's inequality; (b) holds because of the chain-rule, because $M_{\mathcal{Y}}$ and $M_{\mathcal{Z}}^{(c)}$ are independent,

and because conditional entropy is nonnegative; and (c) holds because conditioning cannot increase entropy. The rate of message $M_{\mathcal{Z}}$ satisfies

$$R_{\mathcal{Z}} - \epsilon_n \overset{(a)}{\leq} \frac{1}{n} I(M_{\mathcal{Z}}; Z^n, M_{\mathcal{Y}}) \tag{172}$$

$$\overset{(b)}{\leq} I(M_{\mathcal{Z}}; Z_Q | M_{\mathcal{Y}}, Z^{Q-1}, Q) \tag{173}$$

$$\overset{(c)}{\leq} I(X; Z), \tag{174}$$

where (a) follows form Fano's inequality; (b) holds because of the chain-rule, and because $M_{\mathcal{Y}}$ and $M_{\mathcal{Z}}$ are independent; and (c) holds because conditioning cannot increase entropy, and because Z, X, and $(M_{\mathcal{Y}}, M_{\mathcal{Z}}, Z^{Q-1}, Q)$ form a Markov chain in that order.

We next establish the sum-rate constraint. To this end we first note that if the probability of a decoding error is small, then—with high probability—$M_{\mathcal{Y}}$ and $M_{\mathcal{Z}}^{(p)}$ are computable from $(Y^n, M_{\mathcal{Z}}^{(c)})$ and $(Z^n, M_{\mathcal{Y}})$, respectively, and hence $(M_{\mathcal{Y}}, M_{\mathcal{Z}}^{(p)})$ is computable from $(Y^n, Z^n, M_{\mathcal{Z}}^{(c)})$. We thus obtain

$$R_{\mathcal{Y}} + R_{\mathcal{Z}}^{(p)} - \epsilon_n$$

$$\overset{(a)}{\leq} \frac{1}{n} I\big(M_{\mathcal{Y}}, M_{\mathcal{Z}}^{(p)}; Y^n, Z^n, M_{\mathcal{Z}}^{(c)}\big) \tag{175}$$

$$\overset{(b)}{=} I\big(M_{\mathcal{Y}}, M_{\mathcal{Z}}^{(p)}; Y_Q, Z_Q | M_{\mathcal{Z}}^{(c)}, Y^{Q-1}, Z^{Q-1}, Q\big) \tag{176}$$

$$\overset{(c)}{\leq} I(X; Y, Z), \tag{177}$$

where (a) follows from Fano's inequality; (b) follows from the chain-rule and the independence of $\big(M_{\mathcal{Y}}, M_{\mathcal{Z}}^{(p)}\big)$ and $M_{\mathcal{Z}}^{(c)}$; and (c) holds because conditioning cannot increase entropy, and because (Y, Z), X, and $(M_{\mathcal{Y}}, M_{\mathcal{Z}}, Y^{Q-1}, Z^{Q-1}, Q)$ form a Markov chain in that order.

APPENDIX I
ANALYSIS OF EXAMPLE 20

By Remark 7 and Corollary 3, the no-feedback capacity region without MSI at the stochastic receiver $\mathcal{Z}$ is the set of rate-tuples satisfying (5) for some PMF $p(u, x, y, z)$ of the form (6), where we can w.l.g. restrict X to be a function of (Y, U). Note that for the SD-BC of Example 20

$$I(U; Z) \overset{(a)}{=} I(U; Z, S) \tag{178}$$

$$\overset{(b)}{=} I(U; Z|S) \tag{179}$$

$$\overset{(c)}{=} \bar{p} \, I(U; X) \tag{180}$$

$$\overset{(d)}{=} \bar{p} \, I(U; Y) + \bar{p} \, I(U; X|Y) \tag{181}$$

$$\overset{(e)}{=} \bar{p} \, I(U; Y) + \bar{p} \, H(X|Y) \tag{182}$$

$$= \bar{p} \, I(U; Y) + \bar{p} \sum_{y \in \mathcal{Y}} p(y) H(X|Y = y) \tag{183}$$

$$\overset{(f)}{\leq} \bar{p} \, I(U; Y) + \bar{p} \sum_{y \in \mathcal{Y}} p(y) \log |\mathcal{X}_y|, \tag{184}$$

where (a) holds because S is computable from Z; (b) follows from the chain-rule and the independence of S and U; (c) follows from (35); (d) holds because Y is a function

of X, and because of the chain-rule; (e) holds because X is a function of (Y, U); and (f) holds because $Y = y$ implies that $X \in \mathcal{X}_y$, and because the uniform distribution maximizes entropy. Note that, irrespective of $H(Y)$ and $I(U; Y)$, we can achieve (f) with equality. Using that

$$H(Y|U) + \bar{p} \, I(U; Y) + \bar{p} \sum_{y \in \mathcal{Y}} p(y) \log |\mathcal{X}_y|$$

$$= H(Y) - p \, I(U; Y) + \bar{p} \sum_{y \in \mathcal{Y}} p(y) \log |\mathcal{X}_y|, \tag{185}$$

we conclude that, indeed, the capacity region without feedback is the set of rate-tuples $(R_{\mathcal{Y}}, 0, R_{\mathcal{Z}}^{(p)}, R_{\mathcal{Z}}^{(c)})$ satisfying (36) for some PMF $p(u, x, y, z)$ of the form (6).

Consider now the case with feedback. As we argue next, the feedback capacity region of any SD-BC without MSI at the stochastic receiver $\mathcal{Z}$ ($R_{\mathcal{Y}}^{(p)} = R_{\mathcal{Y}}$) is contained in the set of all rate-tuples satisfying (5a) and (5b) (with $R = 0$) as well as

$$R_{\mathcal{Y}} + R_{\mathcal{Z}} \leq I(X; Y, Z|U) + I(U; Z) \tag{186}$$

for some PMF $p(u, x, y, z)$ of the form (6), irrespective of whether or not the deterministic receiver $\mathcal{Y}$ has MSI ($R_{\mathcal{Z}}^{(p)} \in [0, R_{\mathcal{Z}}]$). Indeed, let $Q \sim \text{Unif}[1 : n]$ be independent of $(M_{\mathcal{Y}}, M_{\mathcal{Z}})$, and introduce

$$U_Q = (M_{\mathcal{Z}}, Y^{Q-1}, Z^{Q-1}), \tag{187a}$$

$$U = (U_Q, Q), \tag{187b}$$

$$X = X_Q, \ Y = Y_Q, \ Z = Z_Q. \tag{187c}$$

Note that—also in the presence of feedback—U, X, and (Y, Z) form a Markov chain in that order, i.e., that their PMF is of the form (6). Using Fano's inequality, it is not hard to show that

$$R_{\mathcal{Y}} - \epsilon_n \leq \frac{1}{n} I(M_{\mathcal{Y}}; Y^n, M_{\mathcal{Z}}) \leq H(Y) \tag{188}$$

$$R_{\mathcal{Z}} - \epsilon_n \leq \frac{1}{n} I(M_{\mathcal{Z}}; Z^n) \leq I(U; Z) \tag{189}$$

$$R_{\mathcal{Y}} + R_{\mathcal{Z}} - \epsilon_n \leq \frac{1}{n} \big[I(M_{\mathcal{Y}}; Y^n, Z^n, M_{\mathcal{Z}}) + I(M_{\mathcal{Z}}; Z^n) \big]$$

$$\leq I(X; Y, Z|U) + I(U; Z), \tag{190}$$

which proves the claim.

We are now ready to conclude that feedback cannot increase the capacity region. To this end we note that for the SD-BC of Example 20

$$I(X; Y, Z|U) + I(U; Z)$$

$$\overset{(a)}{=} H(Y|U) + I(X; Z|Y, U) + I(U; Z) \tag{191}$$

$$\overset{(b)}{=} H(Y|U) + I(X; Z, S|Y, U) + I(U; Z, S) \tag{192}$$

$$\overset{(c)}{=} H(Y|U) + I(X; Z|Y, U, S) + I(U; Z|S) \tag{193}$$

$$\overset{(d)}{=} H(Y|U) + \bar{p} \, H(X|Y, U) + \bar{p} \, I(U; X) \tag{194}$$

$$\overset{(e)}{=} H(Y|U) + \bar{p} \, H(X|Y, U) + \bar{p} \, I(U; Y)$$
$$\quad + \bar{p} \, I(U; X|Y) \tag{195}$$

$$= H(Y) - p \, I(U; Y) + \bar{p} \, H(X|Y, U)$$
$$\quad + \bar{p} \, I(U; X|Y) \tag{196}$$

$$\overset{(f)}{=} H(Y) - p\,I(U;Y) + \bar{p}\,H(X|Y) \qquad (197)$$

$$\overset{(g)}{\leq} H(Y) - p\,I(U;Y) + \bar{p}\sum_{y\in\mathcal{Y}} p(y)\log|\mathcal{X}_y|, \quad (198)$$

where (a) holds because of the chain-rule, and because Y is a function of X; (b) holds because S is computable from Z; (c) follows from the chain-rule and the independence of S and (U,X); (d) follows from (35); (e) holds because Y is a function of X, and because of the chain-rule; (f) follows from the chain-rule; and (g) holds because $Y = y$ implies that $X \in \mathcal{X}_y$, and because the uniform distribution maximizes entropy.

ACKNOWLEDGMENTS

REFERENCES

[1] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[2] S. I. Gel'fand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Problems Inf. Transmiss.*, vol. 16, no. 1, pp. 17–25, Jan./Mar. 1980.

[3] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.

[4] G. Kramer and S. Shamai (Shitz), "Capacity for classes of broadcast channels with receiver side information," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sep. 2007, pp. 313–318.

[5] T. J. Oechtering, M. Wigger, and R. Timo, "Broadcast capacity regions with three receivers and message cognition," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2012, pp. 388–392.

[6] B. Asadi, L. Ong, and S. J. Johnson, "A unified scheme for two-receiver broadcast channels with receiver message side information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 386–390.

[7] Z. Goldfeld, H. H. Permuter, and G. Kramer, "Duality of a source coding problem and the semi-deterministic broadcast channel with rate-limited cooperation," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2285–2307, May 2016.

[8] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.

[9] Y. Steinberg, "Instances of the relay-broadcast channel and cooperation strategies," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 2653–2657.

[10] A. El Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379–381, May 1978.

[11] G. Dueck, "Partial feedback for two-way and broadcast channels," *Inf. Control*, vol. 46, no. 1, pp. 1–15, Jul. 1980.

[12] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 667–671, Jul. 1984.

[13] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 4–21, Jan. 2003.

[14] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1329–1345, Mar. 2013.

[15] R. Venkataramanan and S. S. Pradhan, "An achievable rate region for the broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6175–6191, Oct. 2013.

[16] 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 Trans. Inf. Theory*, vol. 62, no. 4, pp. 2009–2033, Apr. 2016.

[17] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[18] F. M. J. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 313–327, May 1985.

[19] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems Control Theory*, vol. 9, no. 1, pp. 19–31, 1980.

Annina Bracher received the B.Sc. and M.Sc. degrees in electrical engineering (both with distinction) from ETH Zurich in 2010 and 2012 and an additional M.Sc. degree in engineering from Princeton University in 2014. She received the Ph.D. degree in electrical engineering from ETH Zurich in 2016. Dr. Bracher is now a Risk Modeller at Swiss Re.

Michèle Wigger (S'05–M'09–SM'14) received the M.Sc. degree in electrical engineering, with distinction, and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively. In 2009, she was first a post-doctoral fellow at the University of California, San Diego, USA, and then joined Telecom Paris Tech, Paris, France, where she is currently an Associate Professor. Dr. Wigger has held visiting professor appointments at the Technion-Israel Institute of Technology and ETH Zurich. Dr. Wigger has previously served as an Associate Editor of the IEEE COMMUNICATION LETTERS, and is now Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY. She is currently also serving on the Board of Governors of the IEEE Information Theory Society. Dr. Wigger's research interests are in multi-terminal information theory, in particular in distributed source coding and in capacities of networks with states, feedback, user cooperation, or caching.