Identification via the Broadcast Channel

Annina Bracher and Amos Lapidoth
Signal and Information Processing Laboratory
ETH Zurich, Switzerland
E-mail: {bracher,lapidoth}@isi.ee.ethz.ch

***Abstract*—We show that the identification (ID) capacity of the two-receivers broadcast channel is the set of rate pairs satisfying that, for some distribution on the input, each receiver's ID rate does not exceed the mutual information between the input and the output that it observes. The capacity's interior is achieved by codes with deterministic encoders. Our results hold under the average error criterion, which requires that each receiver reliably identify its message if the other receiver's message is uniformly distributed. Key in the proof is a new ID code for the single-user channel.**

I. Introduction

In Shannon's classical transmission problem, the encoder transmits a message from a set $\mathcal{M}$ over a discrete memoryless channel (DMC) $W(y|x)$, and the receiver guesses the transmitted message from the channel outputs. The guess can be any of the $|\mathcal{M}|$ messages in the set $\mathcal{M}$, and the receiver thus faces a hypothesis-testing problem with $|\mathcal{M}|$ hypotheses. Ahlswede and Dueck's identification-via-channels problem [1] is different. Here the encoder sends an identification (ID) message from a set $\mathcal{M}$, and $|\mathcal{M}|$ receiving parties observe the channel outputs. Each party is focused on a different message $m \in \mathcal{M}$. The m-focused party guesses whether or not Message m was sent. It thus faces a hypothesis-testing problem with only two hypotheses.

While in Shannon's problem the number of messages that can be transmitted reliably is exponential, and the transmission rate is defined as the logarithm of the number of transmission messages normalized by the blocklength n, in the ID problem the number of messages that can be identified reliably is double exponential, and the ID rate is defined as the iterated logarithm of the number of ID messages normalized by n. The supremum of all achievable rates is the same for the two problems: both the transmission and the ID capacity equal C, where $C = \max_P I(P,W)$ [1]–[3]. While the transmission capacity is achieved by codes with deterministic encoders, the ID capacity can only be achieved by codes with stochastic encoders. Such encoders associate with each ID message $m \in \mathcal{M}$ a distribution Q_m on the channel inputs. To send Message m, they draw the inputs according to Q_m.

Here we study identification via the broadcast channel (BC) $W(y,z|x)$, where the sender wishes to simultaneously send one distinct ID message to each receiver. We show that the ID capacity of the BC is the set of rate pairs such that for some distribution on the channel input each receiver's ID rate does not exceed the mutual information between the channel input and the channel output that it observes (Theorem 1). The converse that we provide is a strong converse.

The ID capacity of the BC was studied in [4]–[7] under the maximum error criterion, which requires that each receiver reliably identify its message irrespective of the realization of the ID message for the other receiver. Under this criterion, the ID capacity of the BC is still unknown (but see [7] for the case where an additional constraint is imposed on the speed at which the probabilities of error decay to zero). Here, we find the capacity under a different criterion, namely, the average error criterion: We assume independent and uniformly distributed ID messages and require that each receiver reliably identify its message in expectation over the other receiver's message. The resulting ID capacity is typically larger than the set of all rate pairs that are known to be achievable under the maximum error criterion.

We show that codes with deterministic encoders achieve all rate pairs in the interior of the ID capacity of the BC. Note, however, that to each receiver such a deterministic encoder appears to be stochastic since it selects the channel inputs in dependence on the other receiver's uniformly distributed ID message (of positive rate).

Our results extend to the setting where the receivers' ID messages comprise a common and a private part (Theorem 2). Assuming that the private parts are uniformly distributed and independent of each other and of the common part, we require that each receiver reliably identify its message in expectation over the private part of the other receiver's ID message.

We conclude with an inner bound on the ID capacity of the BC with one-sided feedback. It is tight if the outputs are independent conditional on the channel input (Theorem 3).

II. The ID Capacity of the BC

Recall identification via the DMC $W(y|x)$ with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$: Given a set $\mathcal{M}$, a blocklength n, and positive constants λ_1, λ_2, associate with every ID message $m \in \mathcal{M}$ a PMF Q_m on $\mathcal{X}^n$ and a set $\mathcal{D}_m \subset \mathcal{Y}^n$. The tuple $\{Q_m, \mathcal{D}_m\}_m$ is an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code if the maximum probability of missed identification satisfies

$$\max_m (Q_m W^n)(Y^n \notin \mathcal{D}_m) \le \lambda_1, \tag{1}$$

and the maximum probability of wrong identification satisfies

$$\max_{m \neq \hat{m}} (Q_m W^n)(Y^n \in \mathcal{D}_{\hat{m}}) \le \lambda_2. \tag{2}$$

A rate R is achievable if for all positive λ_1, λ_2 and large n there is an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code with $\log\log|\mathcal{M}| \,/n \ge R$. The ID capacity is the supremum of all achievable rates. It was found in [1], [3] to be $\max_P I(P,W)$.

978-1-4799-5186-4/14/$31.00 ©2014 IEEE

We study identification via the BC $W(y,z\,|x)$ with input alphabet $\mathcal{X}$ and output alphabets $\mathcal{Y}$ and $\mathcal{Z}$: Given sets $\mathcal{M}_\mathcal{Y}$ and $\mathcal{M}_\mathcal{Z}$, a blocklength n, and positive constants $\lambda_1^\mathcal{Y}$, $\lambda_2^\mathcal{Y}$, $\lambda_1^\mathcal{Z}$, $\lambda_2^\mathcal{Z}$, associate with every ID message pair $(m_\mathcal{Y}, m_\mathcal{Z}) \in \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}$ a PMF $Q_{m_\mathcal{Y},m_\mathcal{Z}}$ on $\mathcal{X}^n$ and sets $\mathcal{D}^\mathcal{Y}_{m_\mathcal{Y}} \subset \mathcal{Y}^n$ and $\mathcal{D}^\mathcal{Z}_{m_\mathcal{Z}} \subset \mathcal{Z}^n$. The tuple $\{Q_{m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}^\mathcal{Y}_{m_\mathcal{Y}}, \mathcal{D}^\mathcal{Z}_{m_\mathcal{Z}}\}_{m_\mathcal{Y},m_\mathcal{Z}}$ is an $(n, \mathcal{M}_\mathcal{Y}, \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z})$ ID code if the maximum probabilities of missed and wrong identification satisfy

$$\max_{m_\mathcal{Y}} \frac{1}{|\mathcal{M}_\mathcal{Z}|} \sum_{m_\mathcal{Z}\in\mathcal{M}_\mathcal{Z}} (Q_{m_\mathcal{Y},m_\mathcal{Z}} W^n)\big(Y^n \notin \mathcal{D}^\mathcal{Y}_{m_\mathcal{Y}}\big) \le \lambda_1^\mathcal{Y}$$

$$\max_{m_\mathcal{Z}} \frac{1}{|\mathcal{M}_\mathcal{Y}|} \sum_{m_\mathcal{Y}\in\mathcal{M}_\mathcal{Y}} (Q_{m_\mathcal{Y},m_\mathcal{Z}} W^n)\big(Z^n \notin \mathcal{D}^\mathcal{Z}_{m_\mathcal{Z}}\big) \le \lambda_1^\mathcal{Z}$$

$$\max_{m_\mathcal{Y}\neq\hat{m}_\mathcal{Y}} \frac{1}{|\mathcal{M}_\mathcal{Z}|} \sum_{m_\mathcal{Z}\in\mathcal{M}_\mathcal{Z}} (Q_{m_\mathcal{Y},m_\mathcal{Z}} W^n)\Big(Y^n \in \mathcal{D}^\mathcal{Y}_{\hat{m}_\mathcal{Y}}\Big) \le \lambda_2^\mathcal{Y}$$

$$\max_{m_\mathcal{Z}\neq\hat{m}_\mathcal{Z}} \frac{1}{|\mathcal{M}_\mathcal{Y}|} \sum_{m_\mathcal{Y}\in\mathcal{M}_\mathcal{Y}} (Q_{m_\mathcal{Y},m_\mathcal{Z}} W^n)\big(Z^n \in \mathcal{D}^\mathcal{Z}_{\hat{m}_\mathcal{Z}}\big) \le \lambda_2^\mathcal{Z},$$

i.e., if $\{\frac{1}{|\mathcal{M}_\mathcal{Z}|}\sum_{m_\mathcal{Z}} Q_{m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}^\mathcal{Y}_{m_\mathcal{Y}}\}_{m_\mathcal{Y}}$ is an $(n, \mathcal{M}_\mathcal{Y}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y})$ ID code for the marginal channel $W_\mathcal{Y}(y\,|x) = \sum_z W(y,z\,|x)$ and $\{\frac{1}{|\mathcal{M}_\mathcal{Y}|}\sum_{m_\mathcal{Y}} Q_{m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}^\mathcal{Z}_{m_\mathcal{Z}}\}_{m_\mathcal{Z}}$ an $(n, \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z})$ ID code for $W_\mathcal{Z}(z\,|x) = \sum_y W(y,z\,|x)$. (In contrast, the maximum error criterion in [4]–[7] requires for all $m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}$ that $\{Q_{m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}^\mathcal{Y}_{m_\mathcal{Y}}\}_{m_\mathcal{Y}}$ be an $(n, \mathcal{M}_\mathcal{Y}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y})$ ID code for $W_\mathcal{Y}(y\,|x)$, and for all $m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y}$ that $\{Q_{m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}^\mathcal{Z}_{m_\mathcal{Z}}\}_{m_\mathcal{Z}}$ be an $(n, \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z})$ ID code for $W_\mathcal{Z}(z\,|x)$.) A rate pair $(R_\mathcal{Y}, R_\mathcal{Z})$ is achievable if for all positive $\lambda_1^\mathcal{Y}$, $\lambda_2^\mathcal{Y}$, $\lambda_1^\mathcal{Z}$, $\lambda_2^\mathcal{Z}$ and large n there is an $(n, \mathcal{M}_\mathcal{Y}, \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z})$ ID code with $\log\log|\mathcal{M}_\mathcal{Y}|\,/n \ge R_\mathcal{Y}$ and $\log\log|\mathcal{M}_\mathcal{Z}|\,/n \ge R_\mathcal{Z}$. The ID capacity is the closure of the set of achievable rate pairs. Our main result is:

Theorem 1: The ID capacity of the BC $W(y,z\,|x)$ consists of all rate pairs $(R_\mathcal{Y}, R_\mathcal{Z})$ that satisfy for some PMF P on $\mathcal{X}$

$$R_\mathcal{Y} \le I(P, W_\mathcal{Y}) \text{ and } R_\mathcal{Z} \le I(P, W_\mathcal{Z})\,. \tag{3}$$

Its interior is achieved by codes with deterministic encoders.

The proof is deferred to Section V. Here we sketch its direct part. Fix a PMF P on $\mathcal{X}$ and a rate pair $(R_\mathcal{Y}, R_\mathcal{Z})$ satisfying (3). By possibly relabeling the receivers, we can assume w.l.g. that $\mathcal{Y}$ is the "strong receiver" and $\mathcal{Z}$ is the "weak receiver" in the sense that $I(P, W_\mathcal{Y}) \ge I(P, W_\mathcal{Z})$. By possibly increasing $R_\mathcal{Y}$, we can now assume $R_\mathcal{Y} \ge R_\mathcal{Z}$. The blocklength-$n$ transmission is partitioned into two phases: Phase 1 of length $n-\sqrt{n}$ and Phase 2 of length $\sqrt{n}$. We want the weak receiver to be able to reliably identify its ID message $M_\mathcal{Z}$ based on the output symbols that it observes in Phase 1. Moreover, we want the transmitted sequence in Phase 1 to be uniformly distributed over a set of size $2^{nR_\mathcal{Y}}$ and the strong receiver to be able to recover it based on the output symbols that it observes in Phase 1. Put differently, Phase 1 should convey the ID message $M_\mathcal{Z}$ to the weak receiver and also establish common-randomness of rate $R_\mathcal{Y}$ between the encoder and the strong receiver. In Phase 2, we only require that the strong receiver be able to recover the transmitted sequence in Phase 2 and—using the common-randomness it obtained in Phase 1—to identify the ID message $M_\mathcal{Y}$. Note that Phase 2 corresponds to Phase 2 of the common-randomness ID code of [8] and is thus feasible. To prove that also Phase 1 is feasible, we construct in Section III an ID code of rate $R_\mathcal{Z}$ for the weak receiver that has the following property: Provided that the ID message for the weak receiver is drawn uniformly over its support, the distribution that the encoding induces on the channel inputs is uniform over the codebook of size $2^{nR_\mathcal{Y}}$ of some reliable transmission code for the strong receiver. Since the transmission code is reliable, the strong receiver can recover the transmitted sequence in Phase 1. Moreover, since we consider the average error criterion and assume that the ID messages are independent of each other and uniformly distributed, the transmission in Phase 1 is uniformly distributed over a set of size $2^{nR_\mathcal{Y}}$ irrespective of the realization of the ID message for the strong receiver. Thus, Phase 1 indeed establishes common-randomness of rate $R_\mathcal{Y}$ between the encoder and the strong receiver. The new ID code for the weak receiver is key in the proof: In contrast to existing ID codes for the single-user channel, it allows the transmission in Phase 1 to be drawn from a set of size $2^{nR_\mathcal{Y}}$ even if $R_\mathcal{Y}$ is larger than $I(P, W_\mathcal{Z})$.

III. A New ID Code for the DMC

We prove the existence of the ID code for the weak receiver using randomization and show that the following random ID code is with high probability reliable for the DMC $W(y\,|x)$. Fix a PMF P on $\mathcal{X}$, let $R < I(P,W)$ be the ID rate, let n be the blocklength, and let $\mathcal{M}$ be the message set of size $|\mathcal{M}| = 2^{2^{n\tilde{R}}}$. Fix $\tilde{R}$, R_P such that $\tilde{R} < R_P$ and $R < \tilde{R} < I(P,W)$. Draw 2^{nR_P} elements of $\mathcal{X}^n$ independently according to the product PMF P^n, and place them in a pool $\boldsymbol{\mathcal{P}}$. Label the elements of the pool using a set $\mathcal{V}$ of cardinality 2^{nR_P} with $\mathbf{P}(v)$ denoting the n-tuple in the pool labeled by $v \in \mathcal{V}$. For each $m \in \mathcal{M}$ generate a bin $\boldsymbol{\mathcal{B}}_m$ with $2^{n\tilde{R}}$ labels from $\mathcal{V}$ drawn independently and uniformly over the set $\mathcal{V}$, and index the labels in each bin using the set $\mathcal{I} = \{1, \ldots, 2^{n\tilde{R}}\}$ with $V_m(i)$ denoting the label in Bin $\boldsymbol{\mathcal{B}}_m$ indexed by $i \in \mathcal{I}$. Denote the n-tuple $\mathbf{P}(V_m(i))$ by $\mathbf{B}_m(i)$. Reveal the pool $\boldsymbol{\mathcal{P}}$ and bins $\{\boldsymbol{\mathcal{B}}_m\}_m$ to all parties. To send ID Message $m \in \mathcal{M}$, the encoder picks I uniformly over $\mathcal{I}$ and transmits $\mathbf{B}_m(I)$, i.e.,

$$\boldsymbol{Q}_m(\mathbf{x}) = \frac{1}{|\mathcal{I}|} \sum_i \mathbb{1}_{\mathbf{x}=\mathbf{B}_m(i)}. \tag{4}$$

Denote the ID message by M, the picked index by I, the label $V_M(I)$ of the transmitted pool element by V, the inputs $\mathbf{B}_M(I)$ by X^n, and the outputs by Y^n. Let $\epsilon > 0$ satisfy $2\epsilon H(P\times W) < I(P,W) - \tilde{R}$. For $\hat{m} \in \mathcal{M}$ the $\hat{m}$-focused party identifies $\hat{m}$ iff at least one n-tuple in Bin $\boldsymbol{\mathcal{B}}_{\hat{m}}$ is jointly typical with the outputs, i.e., for $\mathcal{T}_\epsilon^{(n)} \triangleq \mathcal{T}_\epsilon^{(n)}(P\times W)$

$$\boldsymbol{\mathcal{D}}_{\hat{m}} = \Big\{\mathbf{y} \in \mathcal{Y}^n \colon \exists i \in \mathcal{I} \text{ s.t. } (\mathbf{B}_{\hat{m}}(i)\,, \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}\Big\}. \tag{5}$$

We next show that the ID code is reliable. Let $\mathbb{P}$ be the distribution on the code, message, index, label, inputs, and outputs. Denote by $\mathbb{E}$ expectation under $\mathbb{P}$. Subscripts indicate

that some RVs assume the values of the subscripts, e.g., for $M = m$ denote expectation under $\mathbb{P}_m$ by $\mathbb{E}_m$. We show that the maximum probabilities of missed and wrong identification converge to 0 in probability over the code's realization. This implies for all positive λ_1, λ_2 and large n that $\{\boldsymbol{Q}_m, \boldsymbol{\mathcal{D}}_m\}_m$ is with high probability an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code.

Consider first missed identification. For $m \in \mathcal{M}$ let $\boldsymbol{\mathcal{I}}_m = \{i \in \mathcal{I} \colon V_m(i) \notin \{V_m(j) : j < i\}\}$ and $\boldsymbol{\mathcal{I}}_m^c = \mathcal{I} \setminus \boldsymbol{\mathcal{I}}_m$. Then,

$$\begin{aligned}
&\mathbb{P}_m\Big[(X^n, Y^n) \notin \mathcal{T}_\epsilon^{(n)} \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\Big] \\
&\overset{(a)}{=} \mathbb{E}_m\Big[\mathbb{E}_m\Big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}, I\Big] \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\Big] \\
&\overset{(b)}{=} \sum_{i\in\mathcal{I}} \mathbb{E}_m\Big[\mathbb{1}_{I=i}\, \mathbb{E}_m\Big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}, I\Big] \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\Big] \\
&\overset{(c)}{=} \sum_{i\in\mathcal{I}} \mathbb{E}_m\Big[\mathbb{1}_{I=i}\, \mathbb{E}_{m,i}\Big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\Big] \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\Big] \\
&\overset{(d)}{=} \sum_{i\in\mathcal{I}} \mathbb{E}_m[\mathbb{1}_{I=i}]\, \mathbb{E}_{m,i}\Big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \Big| \mathbf{B}_m(i)\Big] \qquad (6) \\
&\overset{(e)}{\leq} \frac{|\boldsymbol{\mathcal{I}}_m^c|}{|\mathcal{I}|} + \frac{1}{|\mathcal{I}|} \sum_{i\in\boldsymbol{\mathcal{I}}_m} \mathbb{E}_{m,i}\Big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \Big| \mathbf{B}_m(i)\Big], \qquad (7)
\end{aligned}$$

where (a) is due to the tower property, (b) holds since $\sum_{i\in\mathcal{I}} \mathbb{1}_{I=i} = 1$, expectation is linear, and $\mathbb{1}_{I=i}$ is $\sigma(I)$-measurable, (c) is true because $\mathbb{1}_{I=i}$ is 0 if $I \neq i$, (d) holds since $\mathbb{E}_{m,i}\big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\big]$ is $\sigma(\mathbf{B}_m(i))$-measurable and I is independent of $(\boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\})$, and (e) is true because $\mathcal{I} = \boldsymbol{\mathcal{I}}_m^c \cup \boldsymbol{\mathcal{I}}_m$, the indicator function is at most 1, and $\mathbb{E}_m[\mathbb{1}_{I=i}] = 1/|\mathcal{I}|$. Let $0 < \mu < \tilde{R} - R$ and $\delta_n = \max\{2\,|\mathcal{I}|\,/\,|\mathcal{V}|\,, 2^{-n\mu/2}\}$. Using Azuma's inequality for supermartingales we find that

$$\mathbb{P}\left[\frac{|\boldsymbol{\mathcal{I}}_m^c|}{|\mathcal{I}|} \geq \delta_n\right] \leq e^{-|\mathcal{I}|\left(\delta_n - \frac{|\mathcal{I}|-1}{|\mathcal{V}|}\right)^2/2} \leq e^{-|\mathcal{I}|2^{-n\mu-3}}. \qquad (8)$$

For $\boldsymbol{\mathcal{I}}_m = \mathcal{I}_m$ the RVs $\mathbf{B}_m(i)\,, i \in \mathcal{I}_m$ are IID. Hence, so are the RVs $\mathbb{E}_{m,i}\big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \big| \mathbf{B}_m(i)\big], i \in \mathcal{I}_m$. Moreover,

$$\begin{aligned}
&\mathbb{E}_{\mathcal{I}_m}\Big[\mathbb{E}_{m,i}\Big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \Big| \mathbf{B}_m(i)\Big]\Big] \\
&\quad\overset{(a)}{=} \mathbb{P}_{\mathcal{I}_m,m,i}\Big[(X^n, Y^n) \notin \mathcal{T}_\epsilon^{(n)}\Big] \\
&\quad\overset{(b)}{=} P_X^n \times W^n\Big((X^n, Y^n) \notin \mathcal{T}_\epsilon^{(n)}\Big) \triangleq \beta_n, \qquad (9)
\end{aligned}$$

where (a) is a consequence of the tower property, and (b) is true because under distribution $\mathbb{P}_{\mathcal{I}_m,m,i}$ the channel inputs $X^n = \mathbf{B}_m(i)$ are IID P. Hence, Höffding's inequality and (9) imply for $\alpha_n = \max\{2\beta_n, 2^{-n\mu/2}\sqrt{1/\,(1-\delta_n)}\}$

$$\begin{aligned}
&\mathbb{P}_{\mathcal{I}_m}\left[\frac{1}{|\mathcal{I}|} \sum_{i\in\mathcal{I}_m} \mathbb{E}_{m,i}\Big[\mathbb{1}_{(X^n,Y^n)\notin\mathcal{T}_\epsilon^{(n)}} \Big| \mathbf{B}_m(i)\Big] \geq \alpha_n \frac{|\mathcal{I}_m|}{|\mathcal{I}|}\right] \\
&\leq e^{-2|\mathcal{I}_m|(\alpha_n-\beta_n)^2} \leq e^{-|\mathcal{I}_m|2^{-n\mu-1}/(1-\delta_n)}. \qquad (10)
\end{aligned}$$

For $\kappa_n = \alpha_n + \delta_n$, equations (7), (8), and (10) imply

$$\begin{aligned}
&\mathbb{P}\Big[\mathbb{P}_m\Big[(X^n, Y^n) \notin \mathcal{T}_\epsilon^{(n)} \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\Big] \geq \kappa_n\Big] \\
&\quad\leq e^{-|\mathcal{I}|2^{-n\mu-3}} + e^{-|\mathcal{I}|2^{-n\mu-1}}. \qquad (11)
\end{aligned}$$

In particular, the generated ID code satisfies

$$\begin{aligned}
&\mathbb{P}\big[\exists m \in \mathcal{M} \colon \mathbb{P}_m[Y^n \notin \boldsymbol{\mathcal{D}}_m | \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}] \geq \kappa_n\big] \\
&\overset{(a)}{\leq} \sum_{m\in\mathcal{M}} \mathbb{P}\Big[\mathbb{P}_m\Big[(X^n, Y^n) \notin \mathcal{T}_\epsilon^{(n)} \Big| \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}\Big] \geq \kappa_n\Big] \\
&\overset{(b)}{\leq} 2\,|\mathcal{M}|\, e^{-|\mathcal{I}|2^{-n\mu-3}} \overset{(c)}{\to} 0\,(n \to \infty)\,, \qquad (12)
\end{aligned}$$

where (a) holds by definition of $\boldsymbol{\mathcal{D}}_m$ and the union bound, (b) is due to (11), and (c) follows from $|\mathcal{M}| = 2^{2^{nR}}$, $|\mathcal{I}| = 2^{n\tilde{R}}$, and $\mu < \tilde{R} - R$. Since $\kappa_n \to 0$ as $n \to \infty$, the maximum probability of missed identification converges to 0.

Consider now wrong identification. For $m, \hat{m} \in \mathcal{M}$ distinct let $\boldsymbol{\mathcal{I}}_{m\hat{m}} = \{i \in \boldsymbol{\mathcal{I}}_m \colon V_m(i) \notin \boldsymbol{\mathcal{B}}_{\hat{m}}\}$ and $\tilde{\boldsymbol{\mathcal{I}}}_{m\hat{m}}^c = \{i \in \mathcal{I} \colon V_m(i) \in \boldsymbol{\mathcal{B}}_{\hat{m}}\}$. Then,

$$\begin{aligned}
&\mathbb{P}_m[Y^n \in \boldsymbol{\mathcal{D}}_{\hat{m}} | \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}] \\
&\overset{(a)}{=} \sum_{i\in\mathcal{I}} \mathbb{E}_m[\mathbb{1}_{I=i}]\, \mathbb{E}_{m,i}[\,\mathbb{1}_{Y^n\in\boldsymbol{\mathcal{D}}_{\hat{m}}} | \,\mathbf{B}_m(i)\,, \boldsymbol{\mathcal{D}}_{\hat{m}}] \\
&\overset{(b)}{\leq} \frac{|\boldsymbol{\mathcal{I}}_{m\hat{m}}^c|}{|\mathcal{I}|} + \frac{1}{|\mathcal{I}|} \sum_{i\in\boldsymbol{\mathcal{I}}_{m\hat{m}}} \mathbb{E}_{m,i}[\,\mathbb{1}_{Y^n\in\boldsymbol{\mathcal{D}}_{\hat{m}}} | \,\mathbf{B}_m(i)\,, \boldsymbol{\mathcal{D}}_{\hat{m}}], \qquad (13)
\end{aligned}$$

where (a) and (6) follow similarly, and (b) holds since $\mathcal{I} = \boldsymbol{\mathcal{I}}_{m\hat{m}}^c \cup \boldsymbol{\mathcal{I}}_{m\hat{m}}$ and the indicator function is at most 1. Observe that $|\boldsymbol{\mathcal{I}}_{m\hat{m}}^c| \leq |\boldsymbol{\mathcal{I}}_m^c| + \left|\tilde{\boldsymbol{\mathcal{I}}}_{m\hat{m}}^c\right|$ and

$$\begin{aligned}
\mathbb{P}\left[\frac{|\tilde{\boldsymbol{\mathcal{I}}}_{m\hat{m}}^c|}{|\mathcal{I}|} \geq \delta_n\right] &\overset{(a)}{=} \sum_{\mathcal{B}_{\hat{m}}} \mathbb{P}[\boldsymbol{\mathcal{B}}_{\hat{m}} = \mathcal{B}_{\hat{m}}]\, \mathbb{P}_{\mathcal{B}_{\hat{m}}}\left[\sum_{j\in\mathcal{I}} \frac{\mathbb{1}_{V_m(i)\in\mathcal{B}_{\hat{m}}}}{|\mathcal{I}|} \geq \delta_n\right] \\
&\overset{(b)}{\leq} e^{-2|\mathcal{I}|\left(\delta_n - \frac{|\mathcal{I}|}{|\mathcal{V}|}\right)^2} \leq e^{-|\mathcal{I}|2^{-n\mu-1}}, \qquad (14)
\end{aligned}$$

where (a) holds by the law of total probability and (b) by Höffding's inequality (the RVs $\mathbb{1}_{V_m(i)\in\mathcal{B}_{\hat{m}}}, i \in \mathcal{I}$ are IID with mean at most $|\mathcal{I}|/|\mathcal{V}| < \delta_n$). For $(\boldsymbol{\mathcal{D}}_{\hat{m}}, \boldsymbol{\mathcal{I}}_{m\hat{m}}) = (\mathcal{D}_{\hat{m}}, \mathcal{I}_{m\hat{m}})$ the RVs $\mathbb{E}_{m,i}[\,\mathbb{1}_{Y^n\in\boldsymbol{\mathcal{D}}_{\hat{m}}} | \,\mathbf{B}_m(i)\,, \boldsymbol{\mathcal{D}}_{\hat{m}}]\,, i \in \mathcal{I}_{m\hat{m}}$ are IID and

$$\begin{aligned}
&\mathbb{E}_{\mathcal{D}_{\hat{m}},\mathcal{I}_{m\hat{m}}}\big[\mathbb{E}_{m,i}[\,\mathbb{1}_{Y^n\in\boldsymbol{\mathcal{D}}_{\hat{m}}} | \,\mathbf{B}_m(i)\,, \boldsymbol{\mathcal{D}}_{\hat{m}}]\big] \\
&\quad\overset{(a)}{=} \mathbb{P}_{\mathcal{D}_{\hat{m}},\mathcal{I}_{m\hat{m}},m,i}[Y^n \in \boldsymbol{\mathcal{D}}_{\hat{m}}] \overset{(b)}{=} \sum_{\mathbf{y}\in\mathcal{D}_{\hat{m}}} (PW)^n(\mathbf{y}) \\
&\quad\overset{(c)}{\leq} 2^{-n\left(I(P,W) - \tilde{R} - 2\epsilon H(P\times W)\right)} \triangleq \gamma_n, \qquad (15)
\end{aligned}$$

where (a) is due to the tower property, (b) is true because under distribution $\mathbb{P}_{\mathcal{D}_{\hat{m}},\mathcal{I}_{m\hat{m}},m,i}$ the channel inputs $X^n = \mathbf{B}_m(i)$ are IID P and because of the properties of typical sequences, and (c) holds for n large. Hence, Höffding's inequality implies for $\eta_n = \max\{2\gamma_n, 2^{-n\mu/2}\sqrt{1/\,(1-2\delta_n)}\}$

$$\begin{aligned}
&\mathbb{P}_{\mathcal{I}_{m\hat{m}}}\left[\frac{1}{|\mathcal{I}|} \sum_{i\in\boldsymbol{\mathcal{I}}_{m\hat{m}}} \mathbb{E}_{m,i}[\,\mathbb{1}_{Y^n\in\boldsymbol{\mathcal{D}}_{\hat{m}}} | \,\mathbf{B}_m(i)\,, \boldsymbol{\mathcal{D}}_{\hat{m}}] \geq \eta_n \frac{|\mathcal{I}_{m\hat{m}}|}{|\mathcal{I}|}\right] \\
&\leq e^{-|\mathcal{I}_{m\hat{m}}|2^{-n\mu-1}/(1-2\delta_n)}. \qquad (16)
\end{aligned}$$

For $\omega_n = 2\delta_n + \eta_n$, equations (8), (13), (14), and (16) imply

$$\begin{aligned}
&\mathbb{P}\big[\mathbb{P}_m[Y^n \in \boldsymbol{\mathcal{D}}_{\hat{m}} | \boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}] \geq \omega_n\big] \\
&\quad\leq e^{-|\mathcal{I}|2^{-n\mu-3}} + e^{-|\mathcal{I}|2^{-n\mu-1}} + e^{-|\mathcal{I}|2^{-n\mu-1}}. \qquad (17)
\end{aligned}$$

In particular, the generated ID code satisfies

$$\begin{aligned}&\mathbb{P}\big[\exists m,\hat{m}\in\mathcal{M},\, m\neq\hat{m}\colon \mathbb{P}_m[Y^n\in\boldsymbol{\mathcal{D}}_{\hat{m}}|\,\boldsymbol{\mathcal{P}},\{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}]\geq\omega_n\big]\\&\overset{(a)}{\leq}\sum_{m\in\mathcal{M}}\sum_{\hat{m}\in\mathcal{M}\setminus\{m\}}\mathbb{P}\big[\mathbb{P}_m[Y^n\in\boldsymbol{\mathcal{D}}_{\hat{m}}|\,\boldsymbol{\mathcal{P}},\{\boldsymbol{\mathcal{B}}_{\tilde{m}}\}]\geq\omega_n\big]\\&\overset{(b)}{\leq}3\,|\mathcal{M}|^2\,e^{-|\mathcal{I}|2^{-n\mu-3}}\overset{(c)}{\to}0\,(n\to\infty)\,,\end{aligned}\tag{18}$$

where (a) is a consequence of the union bound, (b) is due to (17), and (c) is true because $|\mathcal{M}|=2^{2^{n\boldsymbol{R}}}$, $|\mathcal{I}|=2^{n\tilde{R}}$, and $\mu<\tilde{R}-R$. Since $\omega_n\to 0$ as $n\to\infty$, the maximum probability of wrong identification converges to 0.

IV. The Common-Randomness ID Code for the DMC

The ID code for the strong receiver only differs from the common-randomness code of [8] insofar as the common-randomness is drawn uniformly over the pool $\boldsymbol{\mathcal{P}}$ of Section III: For a DMC $W(y|x)$, fix a PMF P on $\mathcal{X}$ and R, R_P satisfying $R<R_P<I(P,W)$. Let the message set $\mathcal{M}$ satisfy $|\mathcal{M}|=2^{2^{n\boldsymbol{R}}}$, and let (f,ϕ) be an $\big(\sqrt{n},R,\epsilon_{\sqrt{n}}\big)$ transmission code for the DMC $W(y|x)$, i.e., $\epsilon_{\sqrt{n}}=\max_{u\in\mathcal{U}}W^{\sqrt{n}}\big(Y^{\sqrt{n}}\notin\phi^{-1}(u)\big|f(u)\big)$ for some set $\mathcal{U}$ of cardinality $2^{\sqrt{n}R}$. Since $R<I(P,W)$ we can assume $\epsilon_{\sqrt{n}}\to 0$ as $n\to\infty$. Having fixed P and R_P, generate the pool $\boldsymbol{\mathcal{P}}$ as in the previous section, and label its elements using a set $\mathcal{V}$ of cardinality $2^{nR_{\boldsymbol{P}}}$ with $\mathbf{P}(v)$ denoting the codeword labeled by $v\in\mathcal{V}$. For $v\in\mathcal{V}$ and $m\in\mathcal{M}$ draw $U_v(m)$ independently and uniformly over $\mathcal{U}$. Reveal the pool $\boldsymbol{\mathcal{P}}$ and sequences $\{f(U_v(m))\}_{v,m}$ to all parties. The encoder draws V uniformly over $\mathcal{V}$. In Phase 1 it transmits the common-randomness $\mathbf{P}(V)$, and in Phase 2 it transmits $f(U_V(M))$, where M is the ID message. In Phase 1 the $\hat{m}$-focused receiving party forms an estimate $\hat{V}$ of the common-randomness. Based on the outputs that the $\hat{m}$-focused receiving party observes in Phase 2, it forms an estimate of $f(U_V(M))$ and guesses that Message $\hat{m}$ was sent if this estimate coincides with $f\big(U_{\hat{V}}(\hat{m})\big)$. More precisely, ID message $m\in\mathcal{M}$ is associated with the PMF

$$\boldsymbol{Q}_m(\mathbf{x})=\frac{1}{|\mathcal{V}|}\sum_v\mathbb{1}_{\mathbf{x}=\mathbf{P}(v)\circ f(U_v(m))}\tag{19}$$

on $\mathcal{X}^{n+\sqrt{n}}$ and, for some $\epsilon>0$ satisfying $3\epsilon H(P\times W)<I(P,W)-R_P$ and $\mathcal{T}_\epsilon^{(n)}\triangleq\mathcal{T}_\epsilon^{(n)}(P\times W)$, with the set

$$\begin{aligned}\boldsymbol{\mathcal{D}}_m=\Big\{\mathbf{y}\in\mathcal{Y}^{n+\sqrt{n}}\colon\exists v\in\mathcal{V}\text{ s.t. }&(\mathbf{P}(v)\,,y^n)\in\mathcal{T}_\epsilon^{(n)}\\&\text{and }\phi\big(y_{n+1}^{n+\sqrt{n}}\big)=U_v(m_{\mathcal{Y}})\Big\}.\end{aligned}\tag{20}$$

For $X^n=\mathbf{P}(v)$ let E_v be the event that $(\mathbf{P}(v)\,,Y^n)\notin\mathcal{T}_\epsilon^{(n)}$ or $(\mathbf{P}(\hat{v})\,,Y^n)\in\mathcal{T}_\epsilon^{(n)}$ for some $\hat{v}\neq v$. Since V is independent of $\boldsymbol{\mathcal{P}}$ and drawn uniformly over $\mathcal{V}$, the error probability in Phase 1 is $K=\frac{1}{|\mathcal{V}|}\sum_v\mathbb{P}[E_v|\boldsymbol{\mathcal{P}}]$. The proof of the channel coding theorem implies that $\mathbb{E}[K]$ converges to 0 as $n\to\infty$, and for $\zeta>0$ Markov's inequality implies that $K\leq\zeta\,\mathbb{E}[K]$ with probability at least $1-1/\zeta$. From the analysis in [8] it now follows that $\{\boldsymbol{Q}_m,\boldsymbol{\mathcal{D}}_m\}_m$ is for all positive λ_1, λ_2 and large n with high probability an $(n+\sqrt{n},\mathcal{M},\lambda_1,\lambda_2)$ ID code.

V. Proof of Theorem 1

We now outline converse and direct part of Theorem 1:

Converse: Suppose $\big\{Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}},\mathcal{D}^{\mathcal{Y}}_{m_{\mathcal{Y}}},\mathcal{D}^{\mathcal{Z}}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Y}},m_{\mathcal{Z}}}$ is an $\big(n,\mathcal{M}_{\mathcal{Y}},\mathcal{M}_{\mathcal{Z}},\lambda_1^{\mathcal{Y}},\lambda_2^{\mathcal{Y}},\lambda_1^{\mathcal{Z}},\lambda_2^{\mathcal{Z}}\big)$ ID code with $\lambda_1^{\mathcal{Y}}+\lambda_2^{\mathcal{Y}}+\lambda_1^{\mathcal{Z}}+\lambda_2^{\mathcal{Z}}<1$, and denote by P_{X^n} the empirical type of the inputs X^n. With a slight modification of [3], we obtain that an $(n,\mathcal{M},\lambda_1,\lambda_2)$ ID code for $W(y|x)$ satisfies for every $\epsilon>0$

$$\frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}Q_m\big(I(P_{X^n},W)\geq R-\epsilon\big)\geq 1-\lambda_1-\lambda_2-2^{2^{n(\boldsymbol{R}-\epsilon)}-2^{n\boldsymbol{R}}}.$$

Since $\{Q_{m_{\mathcal{Y}}},\mathcal{D}_{m_{\mathcal{Y}}}\}_{m_{\mathcal{Y}}}$ is an $\big(n,\mathcal{M},\lambda_1^{\mathcal{Y}},\lambda_2^{\mathcal{Y}}\big)$ ID code for $W_{\mathcal{Y}}$ and $\{Q_{m_{\mathcal{Z}}},\mathcal{D}_{m_{\mathcal{Z}}}\}_{m_{\mathcal{Z}}}$ an $\big(n,\mathcal{M},\lambda_1^{\mathcal{Z}},\lambda_2^{\mathcal{Z}}\big)$ ID code for $W_{\mathcal{Z}}$

$$\begin{aligned}&\frac{\sum_{m_{\mathcal{Y}},m_{\mathcal{Z}}}Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}}\big(I(P_{X^n},W_\nu)\geq R_\nu-\epsilon,\,\nu\in\{\mathcal{Y},\mathcal{Z}\}\big)}{|\mathcal{M}_{\mathcal{Y}}|\,|\mathcal{M}_{\mathcal{Z}}|}\\&\geq 1-\sum_{\nu\in\{\mathcal{Y},\mathcal{Z}\}}\left(\sum_{i\in\{1,2\}}\lambda_i^\nu+2^{2^{n(\boldsymbol{R}_\nu-\epsilon)}-2^{n\boldsymbol{R}_\nu}}\right)>0\end{aligned}$$

for n large. Hence, for a rate pair $(R_{\mathcal{Y}},R_{\mathcal{Z}})$ to be achievable there must for every $\epsilon>0$ and n large be an n-tuple $\mathbf{x}\in\mathcal{X}^n$ such that $R_\nu\leq I(P_{\mathbf{x}},W_\nu)+\epsilon$ holds for each $\nu\in\{\mathcal{Y},\mathcal{Z}\}$.

Direct Part: Fix a PMF P on $\mathcal{X}$ and positive rates $R_{\mathcal{Y}},R_{\mathcal{Z}}$ for which the inequalities (3) are strict. Suppose w.l.g. that $R_{\mathcal{Z}}<I(P,W_{\mathcal{Z}})\leq I(P,W_{\mathcal{Y}})$, i.e., $\mathcal{Y}$ is the strong receiver.

Code Generation: For $\nu\in\{\mathcal{Y},\mathcal{Z}\}$ let $\mathcal{M}_\nu$ satisfy $|\mathcal{M}_\nu|=2^{2^{n\boldsymbol{R}_\nu}}$. Fix R_P, $\tilde{R}_{\mathcal{Z}}$ such that $R_{\mathcal{Y}}<R_P<I(P,W_{\mathcal{Y}})$ and $R_{\mathcal{Z}}<\tilde{R}_{\mathcal{Z}}<\min\{R_P,I(P,W_{\mathcal{Z}})\}$. Draw $2^{nR_{\boldsymbol{P}}}$ elements of $\mathcal{X}^n$ independently according to the product PMF P^n and place them in a pool $\boldsymbol{\mathcal{P}}$. Label the elements of $\boldsymbol{\mathcal{P}}$ using a set $\mathcal{V}$ of cardinality $2^{nR_{\boldsymbol{P}}}$ with $\mathbf{P}(v)$ denoting the n-tuple labeled by $v\in\mathcal{V}$. For each $m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}$ generate a bin $\boldsymbol{\mathcal{B}}_{m_{\mathcal{Z}}}$ with $2^{n\tilde{R}_{\mathcal{Z}}}$ labels from $\mathcal{V}$ drawn independently and uniformly over $\mathcal{V}$, and index the labels in each bin using the set $\mathcal{I}=\{1,\dots,2^{n\tilde{R}_{\mathcal{Z}}}\}$ with $V_{m_{\mathcal{Z}}}(i)$ denoting the label in Bin $\boldsymbol{\mathcal{B}}_{m_{\mathcal{Z}}}$ indexed by $i\in\mathcal{I}$. Denote the n-tuple $\mathbf{P}(V_{m_{\mathcal{Z}}}(i))$ by $\mathbf{B}_{m_{\mathcal{Z}}}(i)$. For each $m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}$ draw an index $I_{m_{\mathcal{Y}}}$ independently and uniformly over $\mathcal{I}$. Let (f,ϕ) be an $(\sqrt{n},R_{\mathcal{Y}},\epsilon_{\sqrt{n}})$ transmission code for $W_{\mathcal{Y}}(y|x)$, i.e., $\epsilon_{\sqrt{n}}=\max_{u\in\mathcal{U}}W_{\mathcal{Y}}^{\sqrt{n}}\big(Y^{\sqrt{n}}\notin\phi^{-1}(u)\big|f(u)\big)$ for some set $\mathcal{U}$ of cardinality $2^{\sqrt{n}R_{\mathcal{Y}}}$. Since $R_{\mathcal{Y}}<I(P,W_{\mathcal{Y}})$ we can assume $\epsilon_{\sqrt{n}}\to 0$ as $n\to\infty$. For $v\in\mathcal{V}$ and $m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}$ draw $U_v(m_{\mathcal{Y}})$ independently and uniformly over $\mathcal{U}$. Reveal the pool $\boldsymbol{\mathcal{P}}$, bins $\{\boldsymbol{\mathcal{B}}_{m_{\mathcal{Z}}}\}_{m_{\mathcal{Z}}}$, indices $\{I_{m_{\mathcal{Y}}}\}_{m_{\mathcal{Y}}}$, and sequences $\{f(U_v(m_{\mathcal{Y}}))\}_{v,m_{\mathcal{Y}}}$ to all parties.

Encoding: The 0-1 valued PMF on $\mathcal{X}^{n+\sqrt{n}}$ associated with the ID message pair $(m_{\mathcal{Y}},m_{\mathcal{Z}})\in\mathcal{M}_{\mathcal{Y}}\times\mathcal{M}_{\mathcal{Z}}$ is

$$\boldsymbol{Q}_{m_{\mathcal{Y}},m_{\mathcal{Z}}}(\mathbf{x})=\mathbb{1}_{\mathbf{x}=\mathbf{B}_{m_{\mathcal{Z}}}(I_{m_{\mathcal{Y}}})\circ f\left(U_{V_{m_{\mathcal{Z}}}(I_{m_{\mathcal{Y}}})}(m_{\mathcal{Y}})\right)}.\tag{21}$$

Decoding: Denote the ID message for Receiver $\mathcal{Y}$ by $M_{\mathcal{Y}}$, the ID message for Receiver $\mathcal{Z}$ by $M_{\mathcal{Z}}$, the index $I_{M_{\mathcal{Y}}}$ by I, the label $V_{M_{\mathcal{Z}}}(I)$ of the transmitted pool element by V, the inputs $\mathbf{B}_{M_{\mathcal{Z}}}(I)\circ f(U_V(M_{\mathcal{Y}}))$ by $X^{n+\sqrt{n}}$, the outputs at Receiver $\mathcal{Y}$ by $Y^{n+\sqrt{n}}$, and the outputs at Receiver $\mathcal{Z}$ by $Z^{n+\sqrt{n}}$. Let $\epsilon>0$ satisfy $3\epsilon H(P\times W)<I(P,W_{\mathcal{Y}})-R_P$ and

$2\epsilon H(P\times W)<I(P,W_{\mathcal{Z}})-\tilde{R}_{\mathcal{Z}}$. Denote $\mathcal{T}_\epsilon^{(n)}(P\times W_{\mathcal{Y}})$ and $\mathcal{T}_\epsilon^{(n)}(P\times W_{\mathcal{Z}})$ by $\mathcal{T}_\epsilon^{(n)}$. Associate Message $\hat{m}_{\mathcal{Y}}$ with

$$\boldsymbol{\mathcal{D}}^{\mathcal{Y}}_{\hat{m}_{\mathcal{Y}}}=\Big\{\mathbf{y}\in\mathcal{Y}^{n+\sqrt{n}}\colon \exists v\in\mathcal{V} \text{ s.t. } (\mathbf{P}(v),y^n)\in\mathcal{T}_\epsilon^{(n)} \text{ and } \phi\big(y_{n+1}^{n+\sqrt{n}}\big)=U_v(\hat{m}_{\mathcal{Y}})\Big\}, \tag{22}$$

and Message $\hat{m}_{\mathcal{Z}}$ with the set $\boldsymbol{\mathcal{D}}^{\mathcal{Z}}_{\hat{m}_{\mathcal{Z}}}=\tilde{\boldsymbol{\mathcal{D}}}^{\mathcal{Z}}_{\hat{m}_{\mathcal{Z}}}\times\mathcal{Z}^{\sqrt{n}}$, where

$$\tilde{\boldsymbol{\mathcal{D}}}^{\mathcal{Z}}_{\hat{m}_{\mathcal{Z}}}=\Big\{\mathbf{z}\in\mathcal{Z}^n\colon \exists i\in\mathcal{I} \text{ s.t. } (\mathbf{B}_{\hat{m}_{\mathcal{Z}}}(i),\mathbf{z})\in\mathcal{T}_\epsilon^{(n)}\Big\}. \tag{23}$$

a) Reliability: In the following analysis convergence is understood in probability over the code's realization.

Let $\boldsymbol{Q}_{m_{\mathcal{Y}}}(\mathbf{x})=\frac{1}{|\mathcal{M}_{\mathcal{Z}}|}\sum_{m_{\mathcal{Z}}}\boldsymbol{Q}_{m_{\mathcal{Y}},m_{\mathcal{Z}}}(\mathbf{x})$, and draw $M_{\mathcal{Z}}$ uniformly over $\mathcal{M}_{\mathcal{Z}}$. Due to Höffding's inequality and the union bound the total variation distance between the PMF of $V_{M_{\mathcal{Z}}}(I_{m_{\mathcal{Y}}})$ and $\mathrm{Unif}(\mathcal{V})$ converges to 0 uniformly over $m_{\mathcal{Y}}$. This also holds for $\boldsymbol{Q}_{m_{\mathcal{Y}}}$ and $\frac{1}{|\mathcal{V}|}\sum_v\mathbb{1}_{\mathbf{x}=\mathbf{P}(v)\circ f(U_v(m_{\mathcal{Y}}))}$ since

$$\boldsymbol{Q}_{m_{\mathcal{Y}}}(\mathbf{x})=\frac{1}{|\mathcal{M}_{\mathcal{Z}}|}\sum_{m_{\mathcal{Z}}}\mathbb{1}_{\mathbf{x}=\mathbf{P}\left(V_{m_{\mathcal{Z}}}\left(I_{m_{\mathcal{Y}}}\right)\right)\circ f\left(U_{V_{m_{\mathcal{Z}}}\left(I_{m_{\mathcal{Y}}}\right)}(m_{\mathcal{Y}})\right)}.$$

Hence, $\big\{\boldsymbol{Q}_{m_{\mathcal{Y}}},\boldsymbol{\mathcal{D}}^{\mathcal{Y}}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}}}$ converges to the code of Section IV and is thus for all positive $\lambda_1^{\mathcal{Y}}$, $\lambda_2^{\mathcal{Y}}$ and large n with high probability an $\big(n+\sqrt{n},\mathcal{M}_{\mathcal{Y}},\lambda_1^{\mathcal{Y}},\lambda_2^{\mathcal{Y}}\big)$ ID code for $W_{\mathcal{Y}}(y\,|x)$.

Let $\boldsymbol{Q}_{m_{\mathcal{Z}}}(\mathbf{x})=\frac{1}{|\mathcal{M}_{\mathcal{Y}}|}\sum_{m_{\mathcal{Y}}}\boldsymbol{Q}_{m_{\mathcal{Y}},m_{\mathcal{Z}}}(\mathbf{x})$, draw $M_{\mathcal{Y}}$ uniformly over $\mathcal{M}_{\mathcal{Y}}$, and denote by $\tilde{\boldsymbol{Q}}_{m_{\mathcal{Z}}}$ the PMF on the first n channel inputs given $M_{\mathcal{Z}}=m_{\mathcal{Z}}$, i.e., for $\mathbf{x}\in\mathcal{X}^n$

$$\tilde{\boldsymbol{Q}}_{m_{\mathcal{Z}}}(\mathbf{x})=\sum_{\hat{\mathbf{x}}\in\mathcal{X}^{\sqrt{n}}}\boldsymbol{Q}_{m_{\mathcal{Z}}}(\mathbf{x}\circ\hat{\mathbf{x}})=\frac{1}{|\mathcal{M}_{\mathcal{Y}}|}\sum_{m_{\mathcal{Y}}}\mathbb{1}_{\mathbf{x}=\mathbf{B}_{m_{\mathcal{Z}}}\left(I_{m_{\mathcal{Y}}}\right)}.$$

Höffding's inequality implies that the total variation distance between the PMF of $I_{M_{\mathcal{Y}}}$ and $\mathrm{Unif}(\mathcal{I})$ converges to 0. Hence, the distance between $\tilde{\boldsymbol{Q}}_{m_{\mathcal{Z}}}$ and $\frac{1}{|\mathcal{I}|}\sum_i\mathbb{1}_{\mathbf{x}=\mathbf{B}_{m_{\mathcal{Z}}}(i)}$ converges to 0 uniformly over $m_{\mathcal{Z}}$, and $\big\{\tilde{\boldsymbol{Q}}_{m_{\mathcal{Z}}},\tilde{\boldsymbol{\mathcal{D}}}^{\mathcal{Z}}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}}$ converges to the code of Section III. For all positive $\lambda_1^{\mathcal{Z}}$, $\lambda_2^{\mathcal{Z}}$ and large n it is thus with high probability an $\big(n,\mathcal{M}_{\mathcal{Z}},\lambda_1^{\mathcal{Z}},\lambda_2^{\mathcal{Z}}\big)$ ID code for $W_{\mathcal{Z}}(z\,|x)$, which implies that $\big\{\boldsymbol{Q}_{m_{\mathcal{Z}}},\boldsymbol{\mathcal{D}}^{\mathcal{Z}}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}}$ is with high probability an $\big(n+\sqrt{n},\mathcal{M}_{\mathcal{Z}},\lambda_1^{\mathcal{Z}},\lambda_2^{\mathcal{Z}}\big)$ ID code for $W_{\mathcal{Z}}(z\,|x)$.

To conclude, observe that for each $\nu\in\{\mathcal{Y},\mathcal{Z}\}$ the rate $\log\log|\mathcal{M}_\nu|/(n+\sqrt{n})$ converges to R_ν as $n\to\infty$.

VI. Discussion and Extensions

It is interesting to compare Theorem 1 to the results of [4]–[7] for identification via the BC under the maximum error criterion. If we require that the maximum probabilities of missed and wrong identification decay like n^{-6} or faster, then the maximum error ID capacity equals the common-randomness capacity of the BC [7, Theorem 11]; it is contained in (3); and the containment can be strict: The common-randomness capacity of the degraded BC is the region $\mathcal{R}^{\{i\}}$ of [5], which for example is strictly smaller than (3) if the marginal channels are binary symmetric with different transition probabilities.

We conclude this paper with two extensions of Theorem 1. Suppose first that the ID messages share a common part in the sense that for $\nu\in\{\mathcal{Y},\mathcal{Z}\}$ the message for Receiver ν takes value in the set $\mathcal{M}\times\mathcal{M}_\nu$, and the part with support $\mathcal{M}$ is the same for both receivers. Associate with each $(m,m_{\mathcal{Y}},m_{\mathcal{Z}})\in\mathcal{M}\times\mathcal{M}_{\mathcal{Y}}\times\mathcal{M}_{\mathcal{Z}}$ a PMF $Q_{m,m_{\mathcal{Y}},m_{\mathcal{Z}}}$ on $\mathcal{X}^n$ and sets $\mathcal{D}^{\mathcal{Y}}_{m,m_{\mathcal{Y}}}\subset\mathcal{Y}^n$ and $\mathcal{D}^{\mathcal{Z}}_{m,m_{\mathcal{Z}}}\subset\mathcal{Z}^n$. Call $\big\{Q_{m,m_{\mathcal{Y}},m_{\mathcal{Z}}},\mathcal{D}^{\mathcal{Y}}_{m,m_{\mathcal{Y}}},\mathcal{D}^{\mathcal{Z}}_{m,m_{\mathcal{Z}}}\big\}_{m,m_{\mathcal{Y}},m_{\mathcal{Z}}}$ an $\big(n,\mathcal{M},\mathcal{M}_{\mathcal{Y}},\mathcal{M}_{\mathcal{Z}},\lambda_1^{\mathcal{Y}},\lambda_2^{\mathcal{Y}},\lambda_1^{\mathcal{Z}},\lambda_2^{\mathcal{Z}}\big)$ ID code if $\big\{Q_{m,m_\nu},\mathcal{D}^{\nu}_{m,m_\nu}\big\}_{m,m_\nu}$ is for each $\nu\in\{\mathcal{Y},\mathcal{Z}\}$ an $(n,\mathcal{M}\times\mathcal{M}_\nu,\lambda_1^\nu,\lambda_2^\nu)$ ID code for W_ν, where $Q_{m,m_{\mathcal{Y}}}=\frac{1}{|\mathcal{M}_{\mathcal{Z}}|}\sum_{m_{\mathcal{Z}}}Q_{m,m_{\mathcal{Y}},m_{\mathcal{Z}}}$ and $Q_{m,m_{\mathcal{Z}}}=\frac{1}{|\mathcal{M}_{\mathcal{Y}}|}\sum_{m_{\mathcal{Y}}}Q_{m,m_{\mathcal{Y}},m_{\mathcal{Z}}}$. A rate triple $(R,R_{\mathcal{Y}},R_{\mathcal{Z}})$ is achievable if for all positive $\lambda_1^{\mathcal{Y}}$, $\lambda_2^{\mathcal{Y}}$, $\lambda_1^{\mathcal{Z}}$, $\lambda_2^{\mathcal{Z}}$ and large n there is an $\big(n,\mathcal{M},\mathcal{M}_{\mathcal{Y}},\mathcal{M}_{\mathcal{Z}},\lambda_1^{\mathcal{Y}},\lambda_2^{\mathcal{Y}},\lambda_1^{\mathcal{Z}},\lambda_2^{\mathcal{Z}}\big)$ ID code with $\log\log|\mathcal{M}|/n\geq R$ and $\log\log|\mathcal{M}_\nu|/n\geq R_\nu$ for $\nu\in\{\mathcal{Y},\mathcal{Z}\}$. The ID capacity is the closure of the set of achievable rate triples. It can be characterized as follows:

Theorem 2: If the messages share a common part, then the ID capacity of the BC $W(y,z\,|x)$ consists of all rate triples $(R,R_{\mathcal{Y}},R_{\mathcal{Z}})$ that satisfy for some PMF P on $\mathcal{X}$

$$R,R_{\mathcal{Y}}\leq I(P,W_{\mathcal{Y}}) \text{ and } R,R_{\mathcal{Z}}\leq I(P,W_{\mathcal{Z}}). \tag{24}$$

Its interior is achieved by codes with deterministic encoders.

In a different setting without common message part but with one-sided feedback from Receiver $\mathcal{Y}$, we can use the code of Section III for $\mathcal{Z}$ and the feedback code of [8] for $\mathcal{Y}$: Choose $I(P,W_{\mathcal{Y}})<R_P$ and $R_{\mathcal{Y}}<H(PW_{\mathcal{Y}})\,\mathbb{1}_{\max_{\tilde{P}} I(\tilde{P},W_{\mathcal{Y}})>0}$, and pick the transmission sequence in Phase 2 as a function of $M_{\mathcal{Y}}$ and the common-randomness Y^n. On account of [9] the distribution of Y^n converges to the PMF $(PW_{\mathcal{Y}})^n$, which in [8] is the distribution of Y^n. In particular, we obtain:

Theorem 3: The ID capacity of the BC $W(y,z\,|x)$ with one-sided feedback from Receiver $\mathcal{Y}$ includes all rate pairs $(R_{\mathcal{Y}},R_{\mathcal{Z}})$ that satisfy for some PMF P on $\mathcal{X}$

$$R_{\mathcal{Y}}\leq H(PW_{\mathcal{Y}})\,\mathbb{1}_{\max_{\tilde{P}} I(\tilde{P},W_{\mathcal{Y}})>0} \text{ and } R_{\mathcal{Z}}\leq I(P,W_{\mathcal{Z}}).$$

The bound is achieved by codes with deterministic encoders and tight if $W(y,z\,|x)=W_{\mathcal{Y}}(y\,|x)\,W_{\mathcal{Z}}(z\,|x)$ for all x, y, z.

References

[1] R. Ahlswede and G. Dueck, "Identification via channels," *IEEE Trans. Inf. Theory*, IT-35, No. 1, pp. 15–29, Jan. 1989.

[2] C. E. Shannon, "A mathematical theory of communication," *The Bell System Tech. J.*, Vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.

[3] T. S. Han and S. Verdú, "New results in the theory of identification via channels," *IEEE Trans. Inf. Theory*, IT-38, No. 1, pp. 14–25, Jan. 1992.

[4] B. Verboven and E. C. van der Meulen, "Capacity bounds for identification via broadcast channels that are optimal for the determination broadcast channel," *IEEE Trans. Inf. Theory*, IT-36, No. 6, pp. 1197–1205, Nov. 1990.

[5] I. Bilik and Y. Steinberg, "Inner and outer bounds on the identification capacity region of the degraded broadcast channel," *Proc. of IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 146–146, June 2001.

[6] Y. Oohama, "Converse coding theorem for identification via general degraded broadcast channels," *Proc. of IEEE Int. Symp. on Inf. Theory (ISIT)*, p. 226, July 2003.

[7] R. Ahlswede, "General theory of information transfer: updated," *Discrete Applied Mathematics*, Vol. 156, No. 9, pp. 1348–1388, May 2008,

[8] R. Ahlswede and G. Dueck, "Identification in the presence of feedback-a discovery of new capacity formulas," *IEEE Trans. Inf. Theory*, IT-35, No. 1, pp. 30–36, Jan. 1989.

[9] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, IT-39, No. 3, pp. 752–772, May 1993.