# Identification via the Broadcast Channel

Annina Bracher and Amos Lapidoth, *Fellow, IEEE*

***Abstract*— The identification (ID) capacity region of the two-receiver broadcast channel (BC) is shown to be the set of rate-pairs for which, 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. Moreover, the capacity region's interior is achieved by codes with deterministic encoders. The results are obtained under the average-error criterion, which requires that each receiver reliably identify its message whenever the message intended for the other receiver is drawn at random. They hold also for channels whose transmission capacity region is to-date unknown. Key to the proof is a new ID code construction for the single-user channel. An extension to the three-receiver BC is also discussed: an inner bound on the ID capacity region is obtained, and that is shown to be in some cases tight.**

***Index Terms*— Channel capacity, identification via channels, broadcast channel.**

## I. Introduction

In SHANNON'S classical transmission problem the encoder transmits a message from a message set $\mathcal{M}$ of size $|\mathcal{M}|$ over a discrete memoryless channel (DMC) $W(y|x)$, and the receiver guesses the transmitted message based on the channel's 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. Loosely speaking, we say that a transmission scheme is reliable if, irrespective of the transmitted message $m$, the receiver guesses correctly with high probability. 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 receiving party* must guess whether or not Message $m'$ was sent. It thus faces a hypothesis-testing problem with only two hypotheses. Loosely speaking, we say that an identification scheme is reliable if, for every possible transmitted ID message $m \in \mathcal{M}$ and for every $m' \in \mathcal{M}$ (possibly equal to $m$), the $m'$-focused receiving party guesses correctly with high probability. That is, if $m'$ equals the transmitted ID message $m$, then the $m'$-focused receiving party guesses with high probability that $m'$ was sent, and otherwise it guesses with high probability that $m'$ was not sent.[1]

In Shannon's problem the number of messages that can be transmitted reliably is exponential in the number of channel uses, and the transmission rate is thus defined as the logarithm of the number of transmission messages normalized by the blocklength $n$. In Ahlswede and Dueck's ID problem the number of identifiable messages is double-exponential, and the ID rate is thus defined as the *iterated* logarithm of the number of ID messages normalized by $n$. The suprema of achievable rates for the two problems are identical: both the transmission and the ID capacity equal $C$, where $C = \max_P I(P, W)$ [1]–[3].

The two problems also differ in the role of randomization at the encoder. Whether or not stochastic encoders are allowed does not influence the transmission capacity. However, stochastic encoders are essential for achieving the ID capacity. Such encoders associate with each ID message a distribution on the channel-input sequence and send ID Message $m$ by generating the channel-input sequence according to the distribution associated with $m$. If we only allow deterministic encoders, then the number of identifiable messages grows only exponentially in the blocklength.[2] Throughout this paper we allow stochastic encoders, but for our main achievability result (Theorem 10) they are unnecessary.

The present paper studies identification via a two-receiver broadcast channel (BC) $W(y, z|x)$ whose transmitting terminal is Terminal $\mathcal{X}$ and whose receiving terminals are $\mathcal{Y}$ and $\mathcal{Z}$. The sender wishes to send two ID messages, one to each receiving terminal. The received sequence at Terminal $\mathcal{Y}$ is observed by different parties, each of which is focused—among all the possible ID messages intended for Terminal $\mathcal{Y}$—on a different ID message. Likewise for Terminal $\mathcal{Z}$. We show that the ID capacity region of the BC is the set of rate-pairs for which, 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 10). The converse we provide is a strong converse.

Manuscript received March 27, 2016; revised November 22, 2016; accepted February 1, 2017. Date of publication February 24, 2017; date of current version May 18, 2017. This paper was presented in part at the 2014 IEEE International Symposium on Information Theory.

A. Bracher is with Swiss Reinsurance Company Ltd., CH-8022 Zurich, Switzerland (e-mail: annina_bracher@swissre.com).

A. Lapidoth is with the Signal and Information Processing Laboratory, ETH Zurich, CH-8092 Zurich, Switzerland (e-mail: lapidoth@isi.ee.ethz.ch).

Communicated by S. S. Pradhan, Associate Editor for Shannon Theory.

Color versions of one or more of the figures as well as additional material are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2017.2674669

[1] The corresponding error events are called *missed identification* and *wrong identification*: a missed identification occurs if $m' = m$ and the $m'$-focused receiving party guesses that $m'$ was not sent, and a wrong identification occurs if $m' \neq m$ and the $m'$-focused receiving party guesses that $m'$ was sent. The identification scheme is reliable if the maximum probabilities of missed and wrong identification are small, where the maximum is w.r.t. $m$ for the probability of missed identification and w.r.t. the distinct pair $m, m'$ for the probability of wrong identification.

[2] For ID codes with deterministic encoders, the ID rate is defined as the logarithm of the number of ID messages normalized by $n$, and the supremum of all achievable ID rates is the logarithm of the number of distinct probability mass functions (PMFs) $W(\cdot|x)$ on the channel output that are induced by the different channel-input symbols $x \in \mathcal{X}$ [1].

Our results are obtained under the average-error criterion. Under this criterion, the ID messages $M_{\mathcal{Y}}$ and $M_{\mathcal{Z}}$ to the two receiving terminals are assumed to be independent with each being uniform over its message set ($\mathcal{M}_{\mathcal{Y}}$ or $\mathcal{M}_{\mathcal{Z}}$), and each receiver must identify the message intended for it reliably in expectation over the ID message intended for the other receiving terminal. Loosely speaking, we thus say that an identification scheme is reliable under the average-error criterion if the following two requirements are met: 1) for all (possibly equal) $m_{\mathcal{Y}}, m'_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$, if the ID message that is sent to Terminal $\mathcal{Y}$ is $m_{\mathcal{Y}}$ and the ID message that is sent to Terminal $\mathcal{Z}$ is drawn uniformly over $\mathcal{M}_{\mathcal{Z}}$, then the $m'_{\mathcal{Y}}$-focused receiving party guesses correctly with high probability whether or not $m_{\mathcal{Y}}$ is equal to $m'_{\mathcal{Y}}$; and 2) likewise for all $m_{\mathcal{Z}}, m'_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$.[3]

Identification via the BC was previously studied in [4]–[7] under a different criterion, namely, the maximum-error criterion. Under this criterion each receiver must identify its message reliably irrespective of the realization of the ID message intended for the other receiver. Loosely speaking, we thus say that an identification scheme is reliable under the maximum-error criterion if for all transmitted ID message-pairs $(m_{\mathcal{Y}}, m_{\mathcal{Z}}) \in \mathcal{M}_{\mathcal{Y}} \times \mathcal{M}_{\mathcal{Z}}$ the following two requirements are met: 1) for every $m'_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ (possibly equal to $m_{\mathcal{Y}}$), the $m'_{\mathcal{Y}}$-focused receiving party guesses correctly with high probability whether or not $m_{\mathcal{Y}}$ is equal to $m'_{\mathcal{Y}}$; and 2) likewise for every $m'_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$.

The maximum-error ID capacity region of the BC is to-date unknown.[4] Clearly, the average-error ID capacity region is an outer bound, but whether this bound is tight is unknown. To-date, the best known inner bound on the maximum-error ID capacity region of the BC is the "common-randomness capacity region" of the BC [7]. This inner bound is achieved by a common-randomness ID code, which—like that of [16] for the DMC—uses a transmission code to establish common randomness between the encoder and each decoder. The average-error ID capacity region of the BC typically exceeds this inner bound [8, Remark 2.4.3], but this, of course, does not imply that it exceeds the maximum-error ID capacity region. We do know that the capacity regions differ when only deterministic encoders are allowed, because, unlike the maximum-error ID capacity region (or, for that matter, the single-user channel), all rate-pairs in the interior of the average-error ID capacity region can be achieved by deterministic encoders (Remark 12). This is perhaps not surprising, because to each receiver such a deterministic encoder appears stochastic: the transmitted sequence depends not only on the ID message addressed to it but also on the random ID message (of positive rate) addressed to the other terminal.

To derive our capacity region, we introduce a new capacity-achieving ID code construction for the single-user channel. Our coding scheme for the BC builds on this by making it appear to each receiver as though we were using an instance of the new single-user ID code on its marginal channel. We next describe the new single-user coding scheme, which is reminiscent of [1] but with an important twist that is key to our results. We then describe our scheme for the BC.

For a DMC $W(y|x)$ the new scheme can be described as follows: Fix an input distribution $P$, an ID rate $R < I(P, W)$, and some blocklength $n$. The scheme associates with each ID message $m$ a multiset we call "the $m$-th bin" and whose elements are $n$-tuples (not necessarily distinct) of channel inputs.[5] To send the $m$-th ID message, the (stochastic) encoder sends a random element of this bin. At the receiver's side, the $m'$-focused receiving party guesses that $m'$ was sent if at least one element of the $m'$-th bin is jointly typical with the received $n$-tuple of channel outputs. To construct the bins, we use a random coding argument, with each bin having expected size $e^{n\tilde{R}}$, where $\tilde{R}$ exceeds the ID rate $R$, but is smaller than $I(P, W)$,

$$R < \tilde{R} < I(P, W). \tag{1}$$

The bins are constructed at random from a size $e^{nR_{\mathcal{P}}}$ multiset that we call "pool" and whose elements are $n$-length input sequences. Here $R_{\mathcal{P}}$ can be any number exceeding $\tilde{R}$, possibly even exceeding $I(P, W)$, so, by (1),

$$\tilde{R} < I(P, W) \quad \text{and} \quad R < \tilde{R} < R_{\mathcal{P}}. \tag{2}$$

We construct every bin by randomly selecting its elements from the pool, with the $n$-tuples in the pool being selected for inclusion in the $m$-th bin independently each with probability $e^{-n(R_{\mathcal{P}} - \tilde{R})}$. Since the pool is of size $e^{nR_{\mathcal{P}}}$, each bin is a multiset of expected size $e^{n\tilde{R}}$. The elements of the pool are drawn independently $\sim P^n$. As we shall see, the generated ID code is with high probability reliable (Section II).

Our above scheme is reminiscent of the one in [1]: every ID message is associated with a bin, and in both schemes the bins are chosen at random from a pool. The main difference is that in our scheme the pool need not constitute a codebook that is reliable in Shannon's sense. Indeed, our pool is of size $e^{nR_{\mathcal{P}}}$, where $R_{\mathcal{P}}$ can exceed $I(P, W)$ or even $C$. This flexibility in choosing $R_{\mathcal{P}}$ will be critical on the BC.

[3]The average-error criterion for identification via the BC should not be confused with the average-error criterion for identification via the DMC. On the DMC the average-error criterion requires that for every $m' \in \mathcal{M}$ the probability of wrong identification associated with the pair $m, m'$ be small on average over all possible realizations $m \neq m'$ of the transmitted ID message. Han and Verdú showed that under this criterion the ID capacity is infinite whenever $C > 0$ [3]. This holds because the stochastic encoder can associate the same distribution on the channel-input sequence with an infinite number of ID messages while guaranteeing that the probability of missed identification and the average (but not the maximum) probability of wrong identification be small at each receiving party. The average-error criterion for the BC, which we consider in this paper, is different: for Terminal $\mathcal{Y}$ it requires that the probability of wrong identification associated with any distinct pair $m_{\mathcal{Y}}, m'_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ be small; the term "average" refers to the fact that the probabilities of missed and wrong identification at Terminal $\mathcal{Y}$ are defined on average over all possible realizations $m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$ of the ID message that is sent to Terminal $\mathcal{Z}$. Likewise for Terminal $\mathcal{Z}$.

[4]But see [7] and [8, Sec. 2.4] for the case when an additional constraint is imposed on the decay to zero as a function of the blocklength of the probability of error.

[5]A multiset is a generalized set that allows multiple instances of its elements, e.g., $\{1, 2, 3, 4\}$ and $\{1, 1, 2, 3, 4, 4, 4\}$ are different multisets. The size of a multiset is the number of elements that it contains. The size of the multiset $\{1, 2, 3, 4\}$ is thus four and that of $\{1, 1, 2, 3, 4, 4, 4\}$ is seven. If $X$ is chosen uniformly at random from a multiset, then $\mathbb{P}[X = x]$ is proportional to the number of instances of $x$ in the set. For example, if $X$ is chosen uniformly at random from the multiset $\{1, 1, 2, 3, 4, 4, 4\}$, then $\mathbb{P}[X = 1] = 2/7$.

The scheme we propose for the BC $W(y,z|x)$ is motivated by the single-user scheme. Denote by $W_{\mathcal{Y}}(y|x) = \sum_z W(y,z|x)$ and $W_{\mathcal{Z}}(z|x) = \sum_y W(y,z|x)$ the marginal channels. Fix an input distribution $P$, positive ID rates

$$0 < R_{\mathcal{Y}} < I(P, W_{\mathcal{Y}}),$$
$$0 < R_{\mathcal{Z}} < I(P, W_{\mathcal{Z}}),$$

and some blocklength $n$. We first consider the receivers' side, because in their decoding the receivers follow the single-user scheme. Like the single-user scheme, the scheme for the BC associates with each ID message $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ a multiset we call the $m_{\mathcal{Y}}$-th bin and whose elements are $n$-tuples of channel inputs, and likewise with each ID message $m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$. The $m'_{\mathcal{Y}}$-focused receiving party at Terminal $\mathcal{Y}$ guesses that $m'_{\mathcal{Y}}$ was sent if at least one element of the $m'_{\mathcal{Y}}$-th bin is jointly typical with the sequence it observes, and likewise at Terminal $\mathcal{Z}$. The encoding, however, is different from the single-user scheme. In fact, our encoder for the BC is deterministic: it maps each ID message-pair $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$ to an $n$-tuple of channel inputs we call the "$(m_{\mathcal{Y}}, m_{\mathcal{Z}})$-codeword." (The $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$-codeword is in the intersection of the $m_{\mathcal{Y}}$-th and the $m_{\mathcal{Z}}$-th bin, whenever the intersection is not empty.) We design the codewords and the bins using a random coding argument.

Our goal in designing the codewords and the bins is that to each receiver it would appear as though its intended ID message were sent over its marginal channel using the single-user scheme. More precisely, we want the following to hold: 1) if the ID message that is sent to Terminal $\mathcal{Y}$ is $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ and the ID message that is sent to Terminal $\mathcal{Z}$ is drawn uniformly over $\mathcal{M}_{\mathcal{Z}}$, then the transmitted codeword is nearly uniformly distributed over the $m_{\mathcal{Y}}$-th bin (in terms of Total-Variation distance); and 2) likewise for $m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$. If 1) and 2) hold, then to each receiver it nearly appears as though we were using an instance of the new single-user ID code on its marginal channel: if we view the ID message that is sent to Terminal $\mathcal{Z}$ as uniformly-drawn, then the encoder communicates with Terminal $\mathcal{Y}$ "essentially" using our reliable single-user scheme, and likewise with Terminal $\mathcal{Z}$. To prove that the design goal can be met, we shall use a random coding argument.

The bins are constructed as in the single-user scheme: We construct all the bins—those associated with an ID message $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ or $m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$—from a multiset we call pool. The pool has size $e^{nR_{\mathcal{P}}}$, and each bin associated with an ID message $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ or $m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$ has expected size $e^{n\tilde{R}_{\mathcal{Y}}}$ or $e^{n\tilde{R}_{\mathcal{Z}}}$, respectively. The pool and the bins are generated as in the single-user construction, and $R_{\mathcal{P}}$, $\tilde{R}_{\mathcal{Y}}$, and $\tilde{R}_{\mathcal{Z}}$ meet similar constraints, so

$$\tilde{R}_{\mathcal{Y}} < I(P, W_{\mathcal{Y}}) \quad \text{and} \quad R_{\mathcal{Y}} < \tilde{R}_{\mathcal{Y}} < R_{\mathcal{P}},$$
$$\tilde{R}_{\mathcal{Z}} < I(P, W_{\mathcal{Z}}) \quad \text{and} \quad R_{\mathcal{Z}} < \tilde{R}_{\mathcal{Z}} < R_{\mathcal{P}}.$$

Additionally, we impose the constraint

$$R_{\mathcal{P}} < \tilde{R}_{\mathcal{Y}} + \tilde{R}_{\mathcal{Z}}. \tag{3}$$

(The constraints can all be met, because $R_{\mathcal{Y}}$ and $R_{\mathcal{Z}}$, and thus also $I(P, W_{\mathcal{Y}})$ and $I(P, W_{\mathcal{Z}})$, are positive.) The additional constraint (3) has no counterpart in the single-user setting. It restricts the size of the pool in order to guarantee that with high probability the $m_{\mathcal{Y}}$-th and the $m_{\mathcal{Z}}$-th bin intersect and that consequently the $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$-codeword will be in both bins. If the $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$-codeword is not in this intersection, then, to at least one of the two receivers, it won't appear as though the $n$-tuple of channel inputs were drawn uniformly over the bin associated with its intended ID message. And if this happens to too many pairs $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$, our scheme will fail.

As to the design of the codewords, if the $m_{\mathcal{Y}}$-th and the $m_{\mathcal{Z}}$-th bin intersect, then we draw the $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$-codeword uniformly at random from the intersection, and otherwise we draw it uniformly at random from the pool. As we shall see, the generated ID code meets our design goals with high probability (see Section III-A; key to the proof is that the size of each bin is exponential in $n$ while the cardinalities of $\mathcal{M}_{\mathcal{Y}}$ and $\mathcal{M}_{\mathcal{Z}}$ are double-exponential).

The flexibility afforded by our single-user scheme to choose a pool of size $e^{nR_{\mathcal{P}}}$, where $R_{\mathcal{P}}$ can be larger than $I(P, W_{\mathcal{Y}})$ or $I(P, W_{\mathcal{Z}})$, is crucial to our BC scheme. To see why, consider for now a BC $W(y,z|x)$ and an input distribution $P$ for which

$$I(P, W_{\mathcal{Z}}) < I(P, W_{\mathcal{Y}}).$$

If the pool had been of size $e^{nR_{\mathcal{P}}}$ for some $R_{\mathcal{P}} \le I(P, W_{\mathcal{Z}})$, then at most $\exp\big(\exp\big(nI(P, W_{\mathcal{Z}})\big)\big)$ different bins could have been constructed from the pool, and the BC scheme would have thus failed for $R_{\mathcal{Y}} > I(P, W_{\mathcal{Z}})$, because in this case the number of possible ID messages intended for Receiver $\mathcal{Y}$ would have exceeded the number of different bins. The pool rate $R_{\mathcal{P}}$ must therefore exceed $I(P, W_{\mathcal{Z}})$, and hence the pool cannot consist of a codebook that is reliable in the Shannon sense on the marginal channel $W_{\mathcal{Z}}(z|x)$. It is the possibility of choosing $R_{\mathcal{P}} > I(P, W_{\mathcal{Z}})$ that allows our BC scheme to achieve every rate-pair $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$ satisfying

$$0 < R_{\mathcal{Y}} < I(P, W_{\mathcal{Y}}) \quad \text{and} \quad 0 < R_{\mathcal{Z}} < I(P, W_{\mathcal{Z}}), \tag{4}$$

even when $R_{\mathcal{Y}} > I(P, W_{\mathcal{Z}})$.

The average-error criterion, which we consider in this paper, is suitable whenever the receivers' ID messages are independent and uniform over their supports. As we shall see, we can adapt our coding scheme to solve for the capacity region of a more general scenario where the receivers' ID messages are not independent but have a common part. In this scenario the ID message intended for Terminal $\mathcal{Y}$ is a tuple comprising a private message of rate $R_{\mathcal{Y}}$ and a common message of rate $R$, and likewise for Terminal $\mathcal{Z}$.[6] The common messages are identical, and the private messages are independent, uniformly distributed on their supports, and independent of the common message. We assume that all rates are positive and require that each receiver identify its message reliably in expectation over the other receiver's private message. For this scenario, we show that the ID capacity region of the BC is the set of rate-triples $(R, R_{\mathcal{Y}}, R_{\mathcal{Z}})$ satisfying

$$0 < R, R_{\mathcal{Y}} < I(P, W_{\mathcal{Y}}) \quad \text{and} \quad 0 < R, R_{\mathcal{Z}} < I(P, W_{\mathcal{Z}}) \tag{5}$$

[6]One can view the common-message setting of the transmission problem via the BC as a scenario where the encoder conveys one message to each receiver, but each receiver's message comprises a private and a common part.

for some input distribution $P$ (Theorem 28).[7] Comparing (5) and (4) we see that the common message appears to come for free at all rates up to $\min\{I(P, W_{\mathcal{Y}}), I(P, W_{\mathcal{Z}})\}$. A reason for this is that the ID rate of a pair of ID messages is not equal to the sum of the messages' ID rates.

We also have extensions to the BC with more than two receivers and the two-receiver BC with one-sided feedback: We inner-bound the ID capacity region of the three-receiver BC (Theorem 24) and show that the bound is tight if no receiver is "much more capable" than the other two (see Corollary 26 for more details). In [8, Sec. 2.5.3] we establish the ID capacity region of the two-receiver BC with one-sided feedback for the case where the channel outputs are independent conditional on the channel input [8, Corollary 2.5.15].

The rest of this paper is structured as follows. We conclude this section by introducing some notation and with the concentration inequalities that we shall need. Section II is dedicated to the new ID code for the DMC. Section III studies identification via the BC. The extensions are presented in Section IV, and the paper concludes with a brief summary.

### A. Notation and Terminology

On the single-user channel we denote the channel-input alphabet by $\mathcal{X}$ and the channel-output alphabet by $\mathcal{Y}$. On the two-receiver BC $\mathcal{X}$ is the channel-input alphabet, $\mathcal{Y}$ is the channel-output alphabet at Terminal $\mathcal{Y}$, and $\mathcal{Z}$ is the channel-output alphabet at Terminal $\mathcal{Z}$. All these alphabets are finite. We write $(\mathcal{X}, W(y|x), \mathcal{Y})$ or $W(y|x)$ for a DMC of transition law $W(y|x)$ and $(\mathcal{X}, W(y, z|x), \mathcal{Y}\times\mathcal{Z})$ or $W(y, z|x)$ for a BC of transition law $W(y, z|x)$. We denote the marginal channel of the BC $W(y, z|x)$ to Terminal $\mathcal{Y}$ by $W_{\mathcal{Y}}(y|x)$, i.e., $W_{\mathcal{Y}}(y|x) = \sum_z W(y, z|x)$; and likewise $W_{\mathcal{Z}}(z|x) = \sum_y W(y, z|x)$.

Random variables are denoted by upper-case letters and their realizations or the elements of their supports by lower-case letters, e.g., $Y$ denotes the random output of the DMC and $y \in \mathcal{Y}$ a value it may take. The terms *pool* and *bin* are used for indexed multisets of $n$-tuples from $\mathcal{X}^n$. Pools and bins are denoted by calligraphic letters and in boldface if they are random, e.g., $\boldsymbol{\mathcal{P}}$ denotes a random pool and $\mathcal{P}$ a possible realization. Sequences are in bold lower- or upper-case letters depending on whether they are deterministic or random, e.g., $\mathbf{P}(j)$ denotes the $j$-th $n$-tuple in the random pool $\boldsymbol{\mathcal{P}}$, and $\mathbf{x}$ is an $n$-tuple from $\mathcal{X}^n$. The positive integer $n \in \mathbb{N}$ stands for the blocklength, and, unless otherwise specified, sequences are of length $n$. We denote the positive real numbers by $\mathbb{R}^+$ and the nonnegative real numbers by $\mathbb{R}_0^+$, so $\mathbb{R}_0^+ = \mathbb{R}^+ \cup \{0\}$.

Variables that occur at Time $i$ have the subscript $i$, so $Y_i$ is the Time-$i$ channel-output. Sequences of variables that occur in the time-range $j$ to $i$ bear a subscript $j$ and a superscript $i$, where the subscript $j = 1$ may be dropped, e.g., $Y_4^5$ denotes the fourth and fifth output, and $Y^n$ denotes all the outputs through Time $n$.

[7]The assumption that $R > 0$ is not needed; it only ensures that there is a common message. The assumption that $R_{\mathcal{Y}}$, $R_{\mathcal{Z}} > 0$ is, however, needed: if $R_{\mathcal{Y}}$, say, is zero, then the imposed average-error criterion will turn into a maximum-error criterion for Receiver $\mathcal{Z}$.

The set of PMFs on $\mathcal{X}$ is denoted $\mathscr{P}(\mathcal{X})$, and its generic element $P$. If the input $X$ to the channel $W(y|x)$ is of PMF $P$, then $P \times W$ denotes the joint distribution of $X$ and the channel output $Y$

$$(P \times W)(x, y) = P(x)\, W(y|x), \quad (x, y) \in \mathcal{X} \times \mathcal{Y},$$

and $PW$ denotes the corresponding $Y$-marginal

$$\begin{aligned}(PW)(y) &= \sum_{x\in\mathcal{X}} (P \times W)(x, y) \\ &= \sum_{x\in\mathcal{X}} P(x)\, W(y|x), \quad y \in \mathcal{Y}.\end{aligned}$$

The set of $\epsilon$-typical sequences of length $n$ w.r.t. $P$ is denoted $\mathcal{T}_\epsilon^{(n)}(P)$, i.e.,

$$\mathcal{T}_\epsilon^{(n)}(P) = \left\{\mathbf{x} \in \mathcal{X}^n \colon \left|\frac{N(x|\mathbf{x})}{n} - P(x)\right| \le \epsilon P(x),\ \forall\, x \in \mathcal{X}\right\},$$

where $N(x|\mathbf{x})$ is the number of components of the $n$-tuple $\mathbf{x}$ that equal $x$. We often write $\mathcal{T}_\epsilon^{(n)}$ instead of $\mathcal{T}_\epsilon^{(n)}(P)$ when $P$ is clear from the context. The empirical type of an $n$-tuple $\mathbf{x} \in \mathcal{X}^n$ is denoted $P_{\mathbf{x}}$, so $P_{\mathbf{x}}(x) = N(x|\mathbf{x})/n$, $x \in \mathcal{X}$, and $\mathcal{T}_P^{(n)}$ is the set of all elements of $\mathcal{X}^n$ whose empirical type is $P$. We denote the set of $n$-types on $\mathcal{X}^n$ by $\Gamma^{(n)}$, so

$$\Gamma^{(n)} = \left\{P \in \mathscr{P}(\mathcal{X}) \colon \mathcal{T}_P^{(n)} \neq \emptyset\right\}.$$

For a given DMC $W(y|x)$ and for every $\mathbf{x} \in \mathcal{X}^n$ and $P \in \mathscr{P}(\mathcal{X})$, we denote by $\mathcal{T}_\epsilon^{(n)}(P \times W|\mathbf{x})$ the set of $n$-tuples $\mathbf{y} \in \mathcal{Y}^n$ that are jointly $\epsilon$-typical with $\mathbf{x}$ w.r.t. $P \times W$, i.e.,

$$\mathcal{T}_\epsilon^{(n)}(P \times W|\mathbf{x}) = \left\{\mathbf{y} \in \mathcal{Y}^n \colon (\mathbf{x}, \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}(P \times W)\right\}.$$

Similarly, for a given BC $W(y, z|x)$, $\mathcal{T}_\epsilon^{(n)}(P \times W_{\mathcal{Y}}|\mathbf{x})$ is the set of $n$-tuples $\mathbf{y} \in \mathcal{Y}^n$ that are jointly $\epsilon$-typical with $\mathbf{x}$ w.r.t. $P \times W_{\mathcal{Y}}$, i.e.,

$$\mathcal{T}_\epsilon^{(n)}(P \times W_{\mathcal{Y}}|\mathbf{x}) = \left\{\mathbf{y} \in \mathcal{Y}^n \colon (\mathbf{x}, \mathbf{y}) \in \mathcal{T}_\epsilon^{(n)}(P \times W_{\mathcal{Y}})\right\};$$

and $\mathcal{T}_\epsilon^{(n)}(P \times W_{\mathcal{Z}}|\mathbf{x})$ is the set of $n$-tuples $\mathbf{z} \in \mathcal{Z}^n$ that are jointly $\epsilon$-typical with $\mathbf{x}$ w.r.t. $P \times W_{\mathcal{Z}}$.

A generic probability measure on a measurable space $(\Omega, \mathcal{F})$ is denoted $\mathbb{P}$. If $\mathbb{P}_1$ and $\mathbb{P}_2$ are two probability measures on the same measurable space $(\Omega, \mathcal{F})$, then the Total-Variation distance $d(\mathbb{P}_1, \mathbb{P}_2)$ between $\mathbb{P}_1$ and $\mathbb{P}_2$ is

$$d(\mathbb{P}_1, \mathbb{P}_2) = \sup_{\mathcal{A}\in\mathcal{F}} \mathbb{P}_1[\mathcal{A}] - \mathbb{P}_2[\mathcal{A}].$$

We shall only encounter measurable spaces $(\Omega, \mathcal{F})$ for which $\Omega$ is finite and $\mathcal{F} = 2^\Omega$. On such spaces

$$d(\mathbb{P}_1, \mathbb{P}_2) = \frac{1}{2}\sum_{\omega\in\Omega} \left|\mathbb{P}_1(\omega) - \mathbb{P}_2(\omega)\right|.$$

### B. Some Useful Bounds

We use the following multiplicative Chernoff bounds (see, e.g., [9, Ths. 4.4 and 4.5])[8]:

*Proposition 1: If $S_1, \ldots, S_n$ are independent binary random variables and*

$$\mu = \mathbb{E}\left[\sum_{i=1}^{n} S_i\right],$$

*then for all $0 < \delta < 1$*

$$\mathbb{P}\left[\sum_{i=1}^{n} S_i \le (1-\delta)\mu\right] \le \exp\left\{-\frac{\delta^2\mu}{2}\right\}, \tag{6a}$$

$$\mathbb{P}\left[\sum_{i=1}^{n} S_i \ge (1+\delta)\mu\right] \le \exp\left\{-\frac{\delta^2\mu}{3}\right\}, \tag{6b}$$

*and for all $\delta \ge 1$*

$$\mathbb{P}\left[\sum_{i=1}^{n} S_i \ge (1+\delta)\mu\right] \le \exp\left\{-\frac{\delta\mu}{3}\right\}. \tag{7}$$

We make frequent use of Hoeffding's inequality:

*Proposition 2: [10, Th. 2] If $S_1, \ldots, S_n$ are independent random variables satisfying $S_i \in [a_i, b_i]$, $i \in \{1, \ldots, n\}$, where $a_i$, $b_i \in \mathbb{R}$, then for all $t > 0$*

$$\mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n}\big(S_i - \mathbb{E}[S_i]\big) \ge t\right] \le \exp\left\{-\frac{2n^2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right\}. \tag{8}$$

More general versions of this inequality can be found in [11, Corollary 2.4.7] or [12, Th. 3.24].

## II. A Capacity-Achieving ID Code for the DMC

In this section we present our capacity-achieving ID code for the DMC $\big(\mathcal{X}, W(y|x), \mathcal{Y}\big)$. We begin with the basic definitions of an ID code [1] and with the capacity theorem.

*Definition 3: Fix a finite set $\mathcal{M}$, a blocklength $n \in \mathbb{N}$, and positive constants $\lambda_1$, $\lambda_2$. Associate with every ID message $m \in \mathcal{M}$ a PMF $Q_m$ on $\mathcal{X}^n$ and an ID set $\mathcal{D}_m \subset \mathcal{Y}^n$. The collection of tuples $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{M}}$ is an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code for the DMC $W(y|x)$ if the maximum probability of missed identification*

$$p_{\text{missed-ID}} = \max_{m\in\mathcal{M}} (Q_m W^n)(Y^n \notin \mathcal{D}_m) \tag{9}$$

*and the maximum probability of wrong identification*

$$p_{\text{wrong-ID}} = \max_{m\in\mathcal{M}} \max_{m'\neq m} (Q_m W^n)(Y^n \in \mathcal{D}_{m'}) \tag{10}$$

*satisfy*

$$p_{\text{missed-ID}} \le \lambda_1, \tag{11}$$

$$p_{\text{wrong-ID}} \le \lambda_2. \tag{12}$$

[8]The bound (7) is not stated in [9]. It is, however, a direct consequence of [9, Th. 4.4] and the fact that

$$e^{\delta}/(1+\delta)^{1+\delta} < e^{-\delta/3}, \quad \delta \ge 1.$$

*A rate $R$ is achievable if for every positive $\lambda_1$ and $\lambda_2$ and for every sufficiently-large blocklength $n$ there exists an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code for the DMC with*

$$\begin{cases} \frac{1}{n}\log\log|\mathcal{M}| \ge R & \text{if } R > 0, \\ |\mathcal{M}| = 1 & \text{if } R = 0. \end{cases}$$

*The ID capacity $C$ of the DMC is the supremum of all achievable rates.*

The ID capacity was established in [1] and [3]: Ahlswede and Dueck [1] proved the direct part and a soft converse, which holds for error probabilities that decay exponentially in the blocklength. The strong converse, which holds for all probabilities of missed and wrong identification satisfying $\lambda_1 + \lambda_2 < 1$, is due to Han and Verdú [3].

*Theorem 4 [1, Th. 1] and [3, Th. 2] : The ID capacity $C$ of the DMC $W(y|x)$ is*

$$C = \max_{P} I(P, W). \tag{13}$$

Fix any positive ID rate $R$ satisfying

$$0 < R < \max_{P} I(P, W), \tag{14}$$

and let $\mathcal{M}$ be a size-$\exp(\exp(nR))$ set of possible ID messages. We assume that $\max_P I(P, W)$ is positive, because rate $R = 0$ is always achievable (see Definition 3). We next describe our random code construction and show that, for every positive $\lambda_1$ and $\lambda_2$ and for every sufficiently-large blocklength $n$, it produces with high probability an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code for the DMC $W(y|x)$.

*Code Generation:* Choose a PMF $P$ on $\mathcal{X}$ for which

$$R < I(P, W),$$

and fix an expected bin rate $\tilde{R}$ and a pool rate $R_{\mathcal{P}}$ satisfying

$$R < \tilde{R} < I(P, W) \quad \text{and} \quad \tilde{R} < R_{\mathcal{P}}. \tag{15}$$

Draw $e^{nR_{\mathcal{P}}}$ $n$-tuples $\sim P^n$ independently and place them in a pool $\boldsymbol{\mathcal{P}}$. Index the $n$-tuples in the pool by the elements of a size-$e^{nR_{\mathcal{P}}}$ set $\mathcal{V}$, e.g., $\{1, \ldots, e^{nR_{\mathcal{P}}}\}$, and denote by $\mathbf{P}(v)$ the $n$-tuple in $\boldsymbol{\mathcal{P}}$ that is indexed by $v \in \mathcal{V}$. Associate with each ID message $m \in \mathcal{M}$ an index-set $\boldsymbol{\mathcal{V}}_m$ and a bin $\boldsymbol{\mathcal{B}}_m$ as follows. Select each element of $\mathcal{V}$ for inclusion in $\boldsymbol{\mathcal{V}}_m$ independently with probability $e^{-n(R_{\mathcal{P}}-\tilde{R})}$, and let Bin $\boldsymbol{\mathcal{B}}_m$ be the multiset that contains all the $n$-tuples in the pool that are indexed by $\boldsymbol{\mathcal{V}}_m$,

$$\boldsymbol{\mathcal{B}}_m = \big\{\mathbf{P}(v),\ v \in \boldsymbol{\mathcal{V}}_m\big\}.$$

(Bin $\boldsymbol{\mathcal{B}}_m$ is thus of expected size $e^{n\tilde{R}}$.)

Reveal the pool $\boldsymbol{\mathcal{P}}$, the index-sets $\{\boldsymbol{\mathcal{V}}_m\}_{m\in\mathcal{M}}$, and the corresponding bins $\{\boldsymbol{\mathcal{B}}_m\}_{m\in\mathcal{M}}$ to all parties. The encoding and decoding are determined by

$$\boldsymbol{\mathcal{C}} = \big(\boldsymbol{\mathcal{P}}, \{\boldsymbol{\mathcal{V}}_m\}_{m\in\mathcal{M}}\big). \tag{16}$$

For the purpose of illustration, the pool and the bins are depicted in Figure 1. As mentioned in Section I, our code is similar to the one in [1]: every ID message is associated with a bin, and in both schemes the bins are chosen at random from a pool. The main difference is that in our scheme the pool

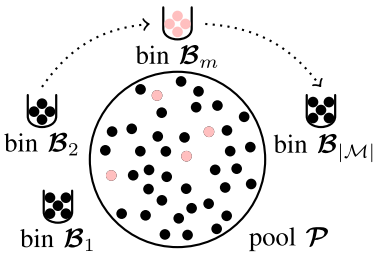


Fig. 1. ID code construction for the DMC.

need not constitute a codebook that is reliable in Shannon's sense. Indeed, our pool is of size $e^{nR_{\mathcal{P}}}$, where $R_{\mathcal{P}}$ can exceed $I(P,W)$ or even $C$.

*Encoding:* To send ID Message $m \in \mathcal{M}$, the encoder draws some $V$ uniformly at random from $\boldsymbol{\mathcal{V}}_m$ and transmits the sequence $\mathbf{P}(V)$. ID Message $m$ is thus associated with the PMF

$$\boldsymbol{Q}_m(\mathbf{x}) = \frac{1}{|\boldsymbol{\mathcal{V}}_m|} \sum_{v \in \boldsymbol{\mathcal{V}}_m} \mathbb{1}_{\mathbf{x}=\mathbf{P}(v)}, \quad \mathbf{x} \in \mathcal{X}^n, \quad \boldsymbol{\mathcal{V}}_m \neq \emptyset. \tag{17}$$

If $\boldsymbol{\mathcal{V}}_m$ is empty, then the encoder chooses $V = v^\star$ and transmits $\mathbf{P}(v^\star)$, where $v^\star$ is an arbitrary but fixed element of $\mathcal{V}$, so

$$\boldsymbol{Q}_m(\mathbf{x}) = \mathbb{1}_{\mathbf{x}=\mathbf{P}(v^\star)}, \quad \mathbf{x} \in \mathcal{X}^n, \quad \boldsymbol{\mathcal{V}}_m = \emptyset. \tag{18}$$

*Decoding:* In this section $\mathcal{T}_\epsilon^{(n)}$ is short for $\mathcal{T}_\epsilon^{(n)}(P \times W)$, and the function $\delta(\cdot)$ maps every nonnegative real number $u$ to $uH(P \times W)$. The decoders choose $\epsilon > 0$ sufficiently small so that $2\delta(\epsilon) < I(P,W) - \tilde{R}$. The $m'$-focused party guesses that $m'$ was sent if, and only if, (iff) for some index $v \in \boldsymbol{\mathcal{V}}_{m'}$ the $n$-tuple $\mathbf{P}(v)$ in Bin $\boldsymbol{\mathcal{B}}_{m'}$ is jointly $\epsilon$-typical with the channel-output sequence $Y^n$, i.e., iff $\big(\mathbf{P}(v), Y^n\big) \in \mathcal{T}_\epsilon^{(n)}$ for some $v \in \boldsymbol{\mathcal{V}}_{m'}$. The set $\boldsymbol{\mathcal{D}}_{m'}$ of output sequences that result in the guess "$m'$ was sent" is thus

$$\boldsymbol{\mathcal{D}}_{m'} = \Big\{\mathbf{y} \in \mathcal{Y}^n : \exists\, v \in \boldsymbol{\mathcal{V}}_{m'} \text{ s.t. } \big(\mathbf{P}(v), \mathbf{y}\big) \in \mathcal{T}_\epsilon^{(n)}\Big\} \tag{19}$$

$$= \bigcup_{v \in \boldsymbol{\mathcal{V}}_{m'}} \mathcal{T}_\epsilon^{(n)}\big(P \times W \big| \mathbf{P}(v)\big). \tag{20}$$

*Analysis of the Probabilities of Missed and Wrong Identification:* We first note that $\boldsymbol{\mathcal{C}}$ (together with the fixed blocklength $n$, the fixed element $v^\star$ of $\mathcal{V}$, and the chosen $\epsilon$) fully specifies the encoding and guessing rules. That is, the randomly constructed ID code $\{\boldsymbol{Q}_m, \boldsymbol{\mathcal{D}}_m\}_{m \in \mathcal{M}}$ is fully specified by $\boldsymbol{\mathcal{C}}$. Let $\mathbb{P}$ be the distribution of $\boldsymbol{\mathcal{C}}$, and let $\mathbb{E}$ denote expectation w.r.t. $\mathbb{P}$. Subscripts indicate conditioning on the event that some of the chance variables assume the values indicated by the subscripts, e.g., $\mathbb{P}_{\boldsymbol{\mathcal{V}}_m}$ denotes the distribution conditional on $\boldsymbol{\mathcal{V}}_m = \mathcal{V}_m$, and $\mathbb{E}_{\boldsymbol{\mathcal{V}}_m}$ denotes the expectation w.r.t. $\mathbb{P}_{\boldsymbol{\mathcal{V}}_m}$.

The maximum probabilities of missed and wrong identification of the randomly constructed ID code $\{\boldsymbol{Q}_m, \boldsymbol{\mathcal{D}}_m\}_{m \in \mathcal{M}}$ are the random variables

$$P_{\text{missed-ID}} = \max_{m \in \mathcal{M}} (\boldsymbol{Q}_m W^n)(Y^n \notin \boldsymbol{\mathcal{D}}_m), \tag{21a}$$

$$P_{\text{wrong-ID}} = \max_{m \in \mathcal{M}} \max_{m' \neq m} (\boldsymbol{Q}_m W^n)(Y^n \in \boldsymbol{\mathcal{D}}_{m'}). \tag{21b}$$

They are fully specified by $\boldsymbol{\mathcal{C}}$. How we upper-bound these probabilities depends on the size of the index-sets and of their pairwise intersections. For every distinct pair $m, m' \in \mathcal{M}$ denote the intersection of the index-sets $\boldsymbol{\mathcal{V}}_m$ and $\boldsymbol{\mathcal{V}}_{m'}$ by $\boldsymbol{\mathcal{V}}_{m,m'}$, so

$$\boldsymbol{\mathcal{V}}_{m,m'} = \boldsymbol{\mathcal{V}}_m \cap \boldsymbol{\mathcal{V}}_{m'}. \tag{22}$$

The expected size of $\boldsymbol{\mathcal{V}}_{m,m'}$ is $e^{n(2\tilde{R}-R_{\mathcal{P}})}$ $(= e^{nR_{\mathcal{P}}} e^{-2n(R_{\mathcal{P}}-\tilde{R})})$ and is thus, by (15), exponentially smaller than the expected size of the index-sets $\boldsymbol{\mathcal{V}}_m$ and $\boldsymbol{\mathcal{V}}_{m'}$, which is $e^{n\tilde{R}}$. The following lemma upper-bounds the probability that the size of the index-sets deviates from its mean $e^{n\tilde{R}}$ or that the pairwise intersections are large compared to $e^{n\tilde{R}}$. To state the lemma, we first introduce the set $\mathcal{G}_\mu$ comprising the realizations $\{\mathcal{V}_m\}_{m \in \mathcal{M}}$ of the index-sets $\{\boldsymbol{\mathcal{V}}_m\}_{m \in \mathcal{M}}$ satisfying that for every distinct pair $m, m' \in \mathcal{M}$ the following three inequalities hold:

$$|\mathcal{V}_m| > (1-\delta_n)e^{n\tilde{R}}, \tag{23a}$$

$$|\mathcal{V}_{m'}| < (1+\delta_n)e^{n\tilde{R}}, \tag{23b}$$

$$|\mathcal{V}_{m,m'}| < e^{n(\tilde{R}-\mu/2)+\log 2}, \tag{23c}$$

where $\mu$ is fixed and satisfies

$$0 < \mu < \min\big\{R_{\mathcal{P}} - \tilde{R}, \tilde{R} - R\big\}, \tag{24}$$

and

$$\delta_n = e^{-n\mu/2}. \tag{25}$$

*Lemma 5: The probability that $\{\boldsymbol{\mathcal{V}}_m\}_{m \in \mathcal{M}}$ is not in $\mathcal{G}_\mu$ converges to zero as the blocklength $n$ tends to infinity:*

$$\lim_{n \to \infty} \mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_m\}_{m \in \mathcal{M}} \notin \mathcal{G}_\mu\big] = 0. \tag{26}$$

*Proof:* See Appendix A. □

To prove that for every choice of $\lambda_1, \lambda_2 > 0$ and $n$ sufficiently large the collection of tuples $\{\boldsymbol{Q}_m, \boldsymbol{\mathcal{D}}_m\}_{m \in \mathcal{M}}$ is with high probability an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code for the DMC $W(y|x)$, we prove the following stronger result:

*Claim 6: The maximum probability of missed identification, $P_{\text{missed-ID}}$, and the maximum probability of wrong identification, $P_{\text{wrong-ID}}$, of the randomly constructed ID code $\{\boldsymbol{Q}_m, \boldsymbol{\mathcal{D}}_m\}_{m \in \mathcal{M}}$ converge in probability to zero exponentially in the blocklength $n$, i.e.,*

$$\exists\, \tau > 0 \text{ s.t.}$$

$$\lim_{n \to \infty} \mathbb{P}\big[\max\{P_{\text{missed-ID}}, P_{\text{wrong-ID}}\} \geq e^{-n\tau}\big] = 0. \tag{27}$$

*Proof:* Fix some $\mu$ satisfying (24), and choose $\delta_n$ as in (25). We upper-bound $P_{\text{missed-ID}}$ and $P_{\text{wrong-ID}}$ differently depending on whether or not $\{\boldsymbol{\mathcal{V}}_v\}$ is in $\mathcal{G}_\mu$, where $\{\boldsymbol{\mathcal{V}}_v\}$ is short for $\{\boldsymbol{\mathcal{V}}_v\}_{v \in \mathcal{M}}$. If $\{\boldsymbol{\mathcal{V}}_v\} \notin \mathcal{G}_\mu$, then we upper-bound them by one to obtain for every $\tau > 0$

$$\begin{aligned}
&\mathbb{P}\big[\max\{P_{\text{missed-ID}}, P_{\text{wrong-ID}}\} \geq e^{-n\tau}\big] \\
&\quad \leq \mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_v\} \notin \mathcal{G}_\mu\big] \\
&\qquad + \sum_{\{\mathcal{V}_v\} \in \mathcal{G}_\mu} \mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_v\} = \{\mathcal{V}_v\}\big] \\
&\qquad \times \mathbb{P}_{\{\mathcal{V}_v\}}\big[\max\{P_{\text{missed-ID}}, P_{\text{wrong-ID}}\} \geq e^{-n\tau}\big].
\end{aligned} \tag{28}$$

By Lemma 5 the first term on the RHS converges to zero as the blocklength $n$ tends to infinity, and it thus suffices to show that

$$\exists\, \tau > 0 \text{ s.t.} \quad \lim_{n\to\infty} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[\max\{P_{\text{missed-ID}}, P_{\text{wrong-ID}}\} \geq e^{-n\tau}\big] = 0. \tag{29}$$

*Remark 7: As we shall see,* (29) *does indeed hold, and we could have therefore simplified our random code construction considerably by drawing only the pool $\boldsymbol{\mathcal{P}}$ at random while fixing the index-sets $\{\mathcal{V}_\upsilon\} \in \mathcal{G}_\mu$. This is correct, but the main purpose of our random code construction for the DMC is to pave the way for the one for the BC, and there we shall need to draw the index-sets at random.*

Henceforth we assume that $n$ is large enough so that the following two inequalities hold:

$$(1-\delta_n)e^{n\tilde{R}} \geq 1, \tag{30a}$$

$$\delta_n + e^{-n\mu/2+\log 2} \leq 1/2, \tag{30b}$$

where $\delta_n$ is defined in (25). (This is possible, because $\delta_n$ converges to zero as $n$ tends to infinity and $\tilde{R},\ \mu > 0$.)

To establish (29), we first show that

$$\exists\, \tau > 0 \text{ s.t.} \quad \lim_{n\to\infty} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[P_{\text{missed-ID}} \geq e^{-n\tau}\big] = 0, \tag{31}$$

and we then show that

$$\exists\, \tau > 0 \text{ s.t.} \quad \lim_{n\to\infty} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[P_{\text{wrong-ID}} \geq e^{-n\tau}\big] = 0. \tag{32}$$

The Union-of-Events bound, (31), and (32) imply (29) and hence (27).

To conclude the proof, it remains to establish (31) and (32). We start by establishing (31). To this end fix any realization $\{\mathcal{V}_\upsilon\}$ in $\mathcal{G}_\mu$. Rather than directly upper-bounding the maximum over $m \in \mathcal{M}$ of $(\boldsymbol{Q}_m W^n)(Y^n \notin \boldsymbol{\mathcal{D}}_m)$ under $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$, we first consider $(\boldsymbol{Q}_m W^n)(Y^n \notin \boldsymbol{\mathcal{D}}_m)$ for a fixed $m \in \mathcal{M}$. (This $\sigma(\boldsymbol{\mathcal{C}})$-measurable random variable with support $[0,1]$ can be viewed as the probability—associated with the randomly constructed ID code—that the $m$-focused party erroneously guesses that $m$ was not sent.) By (23a) (which holds because $\{\mathcal{V}_\upsilon\} \in \mathcal{G}_\mu$) and (30a), $\mathcal{V}_m$ is nonempty, and $\boldsymbol{Q}_m$ is hence given by (17). This implies that $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$-almost-surely the random variable $(\boldsymbol{Q}_m W^n)(Y^n \notin \boldsymbol{\mathcal{D}}_m)$ is upper-bounded by

$$\begin{aligned}
&(\boldsymbol{Q}_m W^n)(Y^n \notin \boldsymbol{\mathcal{D}}_m) \\
&\overset{(a)}{=} \sum_{\mathbf{x}\in\mathcal{X}^n} \frac{1}{|\mathcal{V}_m|} \sum_{\upsilon\in\mathcal{V}_m} \mathbb{1}_{\mathbf{x}=\mathbf{P}(\upsilon)}\, W^n(Y^n \notin \boldsymbol{\mathcal{D}}_m|\mathbf{x}) &(33)\\
&\overset{(b)}{\leq} \frac{1}{|\mathcal{V}_m|} \sum_{\upsilon\in\mathcal{V}_m} W^n\Big(Y^n \notin \mathcal{T}_\epsilon^{(n)}\big(P\times W\big|\mathbf{P}(\upsilon)\big)\Big|\mathbf{P}(\upsilon)\Big), &(34)
\end{aligned}$$

where $(a)$ follows from (17); and $(b)$ follows from (20), which implies that $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$-almost-surely

$$\mathcal{T}_\epsilon^{(n)}\big(P\times W\big|\mathbf{P}(\upsilon)\big) \subseteq \boldsymbol{\mathcal{D}}_m, \quad \upsilon \in \mathcal{V}_m.$$

There is an inequality in $(b)$, because the $m$-focused party may guess correctly even if $\mathbf{y}$ is not jointly typical with $\mathbf{P}(\upsilon)$: it also guesses correctly when $\mathbf{y}$ is jointly typical with $\mathbf{P}(\upsilon')$ for some $\upsilon'$ in $\mathcal{V}_m$ other than $\upsilon$.

Let

$$\beta_n = (P\times W)^n\Big((X^n, Y^n) \notin \mathcal{T}_\epsilon^{(n)}\Big), \tag{35a}$$

$$\alpha_n = \max\big\{2\beta_n, e^{-n\mu/2}\big\}, \tag{35b}$$

and note that (35b) implies that

$$\alpha_n - \beta_n \geq e^{-n\mu/2}/2. \tag{36}$$

Moreover, since $\beta_n$ decays exponentially and $\mu > 0$, there must exist a positive constant $\tau > 0$ and some $\eta_0 \in \mathbb{N}$ for which

$$\alpha_n \leq e^{-n\tau}, \quad n \geq \eta_0. \tag{37}$$

Under $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$ the $[0,1]$-valued random variables

$$\left\{W^n\Big(Y^n \notin \mathcal{T}_\epsilon^{(n)}\big(P\times W\big|\mathbf{P}(\upsilon)\big)\Big|\mathbf{P}(\upsilon)\Big)\right\}_{\upsilon\in\mathcal{V}}$$

are IID and have mean $\beta_n$, because the pool was drawn independently of the index-sets, so $\big\{\mathbf{P}(\upsilon)\big\}_{\upsilon\in\mathcal{V}}$ are IID $\sim P^n$ also under $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$. Consequently, Hoeffding's inequality (Proposition 2) implies that

$$\begin{aligned}
&\mathbb{P}\{\mathcal{V}_\upsilon\}\left[\frac{1}{|\mathcal{V}_m|}\sum_{\upsilon\in\mathcal{V}_m} W^n\Big(Y^n \notin \mathcal{T}_\epsilon^{(n)}\big(P\times W\big|\mathbf{P}(\upsilon)\big)\Big|\mathbf{P}(\upsilon)\Big) \geq \alpha_n\right] \\
&\leq e^{-2|\mathcal{V}_m|(\alpha_n-\beta_n)^2} &(38)\\
&\leq \exp\Big\{-(1-\delta_n)e^{n(\tilde{R}-\mu)-\log 2}\Big\}, \quad \{\mathcal{V}_\upsilon\} \in \mathcal{G}_\mu, &(39)
\end{aligned}$$

where in the second inequality we used (23a) (which holds because $\{\mathcal{V}_\upsilon\} \in \mathcal{G}_\mu$) and (36). Having obtained (39) for every fixed $m$, we are now ready to tackle the maximum over $m$ and prove (31): for every $\tau > 0$ and $\eta_0 \in \mathbb{N}$ satisfying (37) and for all $n$ exceeding $\eta_0$

$$\begin{aligned}
&\max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[P_{\text{missed-ID}} \geq e^{-n\tau}\big] \\
&\overset{(a)}{\leq} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[P_{\text{missed-ID}} \geq \alpha_n\big] &(40)\\
&\overset{(b)}{=} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[\exists m\in\mathcal{M}\colon (\boldsymbol{Q}_m W^n)(Y^n \notin \boldsymbol{\mathcal{D}}_m) \geq \alpha_n\big] &(41)\\
&\overset{(c)}{\leq} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \sum_{m\in\mathcal{M}} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[(\boldsymbol{Q}_m W^n)(Y^n \notin \boldsymbol{\mathcal{D}}_m) \geq \alpha_n\big] &(42)\\
&\overset{(d)}{\leq} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{G}_\mu} \sum_{m\in\mathcal{M}} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\left[\frac{1}{|\mathcal{V}_m|}\sum_{\upsilon\in\mathcal{V}_m} \times W^n\Big(Y^n \notin \mathcal{T}_\epsilon^{(n)}\big(P\times W\big|\mathbf{P}(\upsilon)\big)\Big|\mathbf{P}(\upsilon)\Big) \geq \alpha_n\right] &(43)\\
&\overset{(e)}{\leq} \sum_{m\in\mathcal{M}} \exp\Big\{-(1-\delta_n)e^{n(\tilde{R}-\mu)-\log 2}\Big\} &(44)\\
&\overset{(f)}{\leq} |\mathcal{M}| \exp\Big\{-e^{n(\tilde{R}-\mu)-2\log 2}\Big\} &(45)\\
&\overset{(g)}{\to} 0\ (n\to\infty), &(46)
\end{aligned}$$

where $(a)$ holds by (37), because $n$ exceeds $\eta_0$; $(b)$ follows from (21a); $(c)$ follows from the Union-of-Events bound; $(d)$ follows from (34); $(e)$ holds by (39); $(f)$ follows from (30b), which implies that $\delta_n \leq 1/2$; and $(g)$ holds because $|\mathcal{M}| = \exp(\exp(nR))$ and $\mu < \tilde{R} - R$.

Having established (31), it remains to establish (32) in order to conclude the proof. To this end fix any realization $\{\mathcal{V}_v\}$ in $\mathcal{G}_\mu$. We begin by upper-bounding $(\boldsymbol{Q}_m W^n)(Y^n \in \boldsymbol{\mathcal{D}}_{m'})$ under $\mathbb{P}_{\{\mathcal{V}_v\}}$ for fixed distinct $m, m' \in \mathcal{M}$. Later we will maximize over such $m, m'$. (The $\sigma(\mathcal{C})$-measurable random variable $(\boldsymbol{Q}_m W^n)(Y^n \in \boldsymbol{\mathcal{D}}_{m'})$ with support $[0,1]$ can be viewed as the probability—associated with the randomly constructed ID code—that the $m'$-focused party erroneously guesses that $m'$ was sent when in fact $m$ was sent.) By (23a) (which holds because $\{\mathcal{V}_v\} \in \mathcal{G}_\mu$) and (30a), $\mathcal{V}_m$ is nonempty, and $\boldsymbol{Q}_m$ is hence given by (17). This implies that $\mathbb{P}_{\{\mathcal{V}_v\}}$-almost-surely the random variable $(\boldsymbol{Q}_m W^n)(Y^n \in \boldsymbol{\mathcal{D}}_{m'})$ is upper-bounded by

$$
\begin{aligned}
&(\boldsymbol{Q}_m W^n)(Y^n \in \boldsymbol{\mathcal{D}}_{m'}) \\
&\quad \overset{(a)}{=} \sum_{\mathbf{x} \in \mathcal{X}^n} \frac{1}{|\mathcal{V}_m|} \sum_{v \in \mathcal{V}_m} \mathbb{1}_{\mathbf{x}=\mathbf{P}(v)}\, W^n(Y^n \in \boldsymbol{\mathcal{D}}_{m'} | \mathbf{x}) && (47) \\
&\quad = \frac{1}{|\mathcal{V}_m|} \sum_{v \in \mathcal{V}_m} W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'} \big| \mathbf{P}(v)\big) && (48) \\
&\quad \overset{(b)}{\leq} \frac{|\mathcal{V}_{m,m'}|}{|\mathcal{V}_m|} \\
&\qquad + \frac{1}{|\mathcal{V}_m|} \sum_{v \in \mathcal{V}_m \setminus \mathcal{V}_{m,m'}} W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'} \big| \mathbf{P}(v)\big), && (49)
\end{aligned}
$$

where $(a)$ follows from (17); and $(b)$ holds because

$$W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'} \big| \mathbf{P}(v)\big) \leq 1, \quad v \in \mathcal{V}.$$

We consider the two terms on the RHS of (49) separately, beginning with $|\mathcal{V}_{m,m'}|/|\mathcal{V}_m|$. Because $\{\mathcal{V}_v\} \in \mathcal{G}_\mu$,

$$\frac{|\mathcal{V}_{m,m'}|}{|\mathcal{V}_m|} \overset{(a)}{<} \frac{e^{n(\tilde{R}-\mu/2)+\log 2}}{(1-\delta_n)e^{n\tilde{R}}} \overset{(b)}{\leq} e^{-n\mu/2+2\log 2}, \tag{50}$$

where $(a)$ follows from (23a) and (23c); and $(b)$ follows from (30b), which implies that $\delta_n \leq 1/2$. We next consider the second term in (49), namely,

$$\frac{1}{|\mathcal{V}_m|} \sum_{v \in \mathcal{V}_m \setminus \mathcal{V}_{m,m'}} W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'} \big| \mathbf{P}(v)\big).$$

The cardinality of $\boldsymbol{\mathcal{D}}_{m'}$ is $\mathbb{P}_{\{\mathcal{V}_v\}}$-almost-surely upper-bounded by

$$
\begin{aligned}
|\boldsymbol{\mathcal{D}}_{m'}| &\overset{(a)}{=} \Big| \bigcup_{v \in \mathcal{V}_{m'}} \mathcal{T}_\epsilon^{(n)}\big(P \times W \big| \mathbf{P}(v)\big) \Big| \\
&\leq \sum_{v \in \mathcal{V}_{m'}} \Big| \mathcal{T}_\epsilon^{(n)}\big(P \times W \big| \mathbf{P}(v)\big) \Big| && (51) \\
&\overset{(b)}{\leq} (1+\delta_n) e^{n(\tilde{R}+H(W|P)+\delta(\epsilon))}, && (52)
\end{aligned}
$$

where $(a)$ follows from (20); and $(b)$ follows from

$$\Big| \mathcal{T}_\epsilon^{(n)}(P \times W | \mathbf{x}) \Big| \leq e^{n(H(W|P)+\delta(\epsilon))}, \quad \mathbf{x} \in \mathcal{X}^n,$$

and from (23b) (which holds because $\{\mathcal{V}_v\} \in \mathcal{G}_\mu$).

Let

$$\gamma_n = (1+\delta_n) e^{-n(I(P,W)-\tilde{R}-2\delta(\epsilon))}, \tag{53a}$$

$$\kappa_n = \max\big\{2\gamma_n, e^{-n\mu/2}\big\}, \tag{53b}$$

and note that (53b) implies that

$$\kappa_n - \gamma_n \geq e^{-n\mu/2}/2. \tag{54}$$

Fix a realization $\mathcal{D}_{m'}$ of $\boldsymbol{\mathcal{D}}_{m'}$ for which $\mathbb{P}_{\{\mathcal{V}_v\}}[\boldsymbol{\mathcal{D}}_{m'} = \mathcal{D}_{m'}] > 0$. From (20) it follows that all output sequences in $\mathcal{D}_{m'}$ are of approximate type $PW$, i.e., that

$$\mathcal{D}_{m'} \subseteq \mathcal{T}_\epsilon^{(n)}(PW). \tag{55}$$

And from (52) it follows that

$$|\mathcal{D}_{m'}| \leq (1+\delta_n) e^{n(\tilde{R}+H(W|P)+\delta(\epsilon))}. \tag{56}$$

The next computation is under $\mathbb{P}_{\{\mathcal{V}_v\},\mathcal{D}_{m'}}$, where we condition not only on $\{\boldsymbol{\mathcal{V}}_v\} = \{\mathcal{V}_v\}$ but also on $\boldsymbol{\mathcal{D}}_{m'} = \mathcal{D}_{m'}$. The $n$-tuples in the pool $\big\{\mathbf{P}(v)\big\}_{v \in \mathcal{V} \setminus \mathcal{V}_{m'}}$ that are not indexed by $\mathcal{V}_{m'}$ are IID $\sim P^n$ also under $\mathbb{P}_{\{\mathcal{V}_v\},\mathcal{D}_{m'}}$, because the pool was drawn independently of the index-sets, and because by (20) $\boldsymbol{\mathcal{D}}_{m'}$ depends only on $\big\{\mathbf{P}(v)\big\}_{v \in \boldsymbol{\mathcal{V}}_{m'}}$. Hence, under $\mathbb{P}_{\{\mathcal{V}_v\},\mathcal{D}_{m'}}$ the $[0,1]$-valued random variables

$$\Big\{ W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'} \big| \mathbf{P}(v)\big) \Big\}_{v \in \mathcal{V} \setminus \mathcal{V}_{m'}}$$

are IID of mean

$$
\begin{aligned}
&\mathbb{E}_{\{\mathcal{V}_v\},\mathcal{D}_{m'}}\Big[ W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'} \big| \mathbf{P}(v)\big) \Big] \\
&\quad \overset{(a)}{=} \sum_{\mathbf{y} \in \mathcal{D}_{m'}} (PW)^n(\mathbf{y}) && (57) \\
&\quad \overset{(b)}{\leq} |\mathcal{D}_{m'}|\, e^{-n(H(PW)-\delta(\epsilon))} && (58) \\
&\quad \overset{(c)}{\leq} (1+\delta_n) e^{-n(I(P,W)-\tilde{R}-2\delta(\epsilon))} && (59) \\
&\quad \overset{(d)}{=} \gamma_n, && (60)
\end{aligned}
$$

where $(a)$ holds because $\boldsymbol{\mathcal{D}}_{m'} = \mathcal{D}_{m'}$ and $\big\{\mathbf{P}(v)\big\}_{v \in \mathcal{V} \setminus \mathcal{V}_{m'}}$ are IID $\sim P^n$ under $\mathbb{P}_{\{\mathcal{V}_v\},\mathcal{D}_{m'}}$; $(b)$ holds because

$$(PW)^n(\mathbf{y}) \leq e^{-n(H(PW)-\delta(\epsilon))}, \quad \mathbf{y} \in \mathcal{T}_\epsilon^{(n)}(PW),$$

and by (55); $(c)$ follows from (56); and $(d)$ holds by (53a). Consequently, Hoeffding's inequality (Proposition 2) implies that

$$
\begin{aligned}
&\mathbb{P}_{\{\mathcal{V}_v\},\mathcal{D}_{m'}}\Bigg[ \frac{1}{|\mathcal{V}_m \setminus \mathcal{V}_{m,m'}|} \sum_{v \in \mathcal{V}_m \setminus \mathcal{V}_{m,m'}} \\
&\qquad \times\ W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'} \big| \mathbf{P}(v)\big) \geq \kappa_n \Bigg] \\
&\quad \overset{(a)}{\leq} \exp\big\{ -2\,|\mathcal{V}_m \setminus \mathcal{V}_{m,m'}|\,(\kappa_n - \gamma_n)^2 \big\} && (61) \\
&\quad \overset{(b)}{\leq} \exp\big\{ -|\mathcal{V}_m \setminus \mathcal{V}_{m,m'}|\, e^{-n\mu-\log 2} \big\} && (62) \\
&\quad \overset{(c)}{\leq} \exp\Big\{ -e^{n(\tilde{R}-\mu)-2\log 2} \Big\}, \\
&\qquad \{\mathcal{V}_v\} \in \mathcal{G}_\mu,\ \mathbb{P}_{\{\mathcal{V}_v\}}[\boldsymbol{\mathcal{D}}_{m'} = \mathcal{D}_{m'}] > 0, && (63)
\end{aligned}
$$

where $(a)$ holds because $\mathcal{V}_m \setminus \mathcal{V}_{m,m'}$ is a subset of $\mathcal{V} \setminus \mathcal{V}_{m'}$; $(b)$ follows from (54); and $(c)$ follows from

$$|\mathcal{V}_m \setminus \mathcal{V}_{m,m'}| \overset{(d)}{>} (1-\delta_n)e^{n\tilde{R}} - e^{n(\tilde{R}-\mu/2)+\log 2} \tag{64}$$

$$\overset{(e)}{\geq} e^{n\tilde{R}-\log 2}, \tag{65}$$

where $(d)$ is due to (23a) and (23c) (which hold because $\{\mathcal{V}_v\} \in \mathcal{G}_\mu$), and $(e)$ is due to (30b). By (63) and because $|\mathcal{V}_m \setminus \mathcal{V}_{m,m'}| \leq |\mathcal{V}_m|$, the probability that the second term in (49) exceeds $\kappa_n$ is upper-bounded by

$$\mathbb{P}_{\{\mathcal{V}_v\}}\Bigg[\frac{1}{|\mathcal{V}_m|}\sum_{v\in\mathcal{V}_m\setminus\mathcal{V}_{m,m'}} W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'}\big|\mathbf{P}(v)\big) \geq \kappa_n\Bigg]$$

$$\leq \mathbb{P}_{\{\mathcal{V}_v\}}\Bigg[\frac{1}{|\mathcal{V}_m \setminus \mathcal{V}_{m,m'}|}\sum_{v\in\mathcal{V}_m\setminus\mathcal{V}_{m,m'}} \times W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'}\big|\mathbf{P}(v)\big) \geq \kappa_n\Bigg] \tag{66}$$

$$= \sum_{\mathcal{D}_{m'}} \mathbb{P}_{\{\mathcal{V}_v\}}[\boldsymbol{\mathcal{D}}_{m'} = \mathcal{D}_{m'}]\,\mathbb{P}_{\{\mathcal{V}_v\},\mathcal{D}_{m'}}\Bigg[\frac{1}{|\mathcal{V}_m \setminus \mathcal{V}_{m,m'}|} \times \sum_{v\in\mathcal{V}_m\setminus\mathcal{V}_{m,m'}} W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'}\big|\mathbf{P}(v)\big) \geq \kappa_n\Bigg] \tag{67}$$

$$\leq \exp\Big\{-e^{n(\tilde{R}-\mu)-2\log 2}\Big\}, \quad \{\mathcal{V}_v\} \in \mathcal{G}_\mu. \tag{68}$$

Having obtained (49), (50), and (68) for every fixed distinct $m$, $m'$, we are now ready to tackle the maximum over $m$, $m'$ and prove (32): Let

$$\omega_n = e^{-n\mu/2+2\log 2} + \kappa_n, \tag{69}$$

and note that, by (53), because $\mu > 0$, because $\delta_n$ converges to zero as $n$ tends to infinity, and because $2\delta(\epsilon) < I(P,W) - \tilde{R}$, there must exist a positive constant $\tau > 0$ and some $\eta_0 \in \mathbb{N}$ for which

$$\omega_n \leq e^{-n\tau}, \quad n \geq \eta_0. \tag{70}$$

For every $\tau > 0$ and $\eta_0 \in \mathbb{N}$ satisfying (70) and for all $n$ exceeding $\eta_0$

$$\max_{\{\mathcal{V}_v\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_v\}}\big[P_{\text{wrong-ID}} \geq e^{-n\tau}\big]$$

$$\overset{(a)}{\leq} \max_{\{\mathcal{V}_v\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_v\}}\big[P_{\text{wrong-ID}} \geq \omega_n\big] \tag{71}$$

$$\overset{(b)}{=} \max_{\{\mathcal{V}_v\}\in\mathcal{G}_\mu} \mathbb{P}_{\{\mathcal{V}_v\}}\big[\exists\, m, m' \in \mathcal{M},\, m \neq m' : (\boldsymbol{\mathcal{Q}}_m W^n)(Y^n \in \boldsymbol{\mathcal{D}}_{m'}) \geq \omega_n\big] \tag{72}$$

$$\overset{(c)}{\leq} \max_{\{\mathcal{V}_v\}\in\mathcal{G}_\mu} \sum_{m\in\mathcal{M}}\sum_{m'\neq m} \mathbb{P}_{\{\mathcal{V}_v\}}\big[(\boldsymbol{\mathcal{Q}}_m W^n)(Y^n \in \boldsymbol{\mathcal{D}}_{m'}) \geq \omega_n\big] \tag{73}$$

$$\overset{(d)}{\leq} \max_{\{\mathcal{V}_v\}\in\mathcal{G}_\mu} \sum_{m\in\mathcal{M}}\sum_{m'\neq m} \mathbb{P}_{\{\mathcal{V}_v\}}\Bigg[\frac{|\mathcal{V}_{m,m'}|}{|\mathcal{V}_m|} + \frac{1}{|\mathcal{V}_m|}\sum_{v\in\mathcal{V}_m\setminus\mathcal{V}_{m,m'}} W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'}\big|\mathbf{P}(v)\big) \geq \omega_n\Bigg] \tag{74}$$

$$\overset{(e)}{\leq} \max_{\{\mathcal{V}_v\}\in\mathcal{G}_\mu} \sum_{m\in\mathcal{M}}\sum_{m'\neq m} \mathbb{P}_{\{\mathcal{V}_v\}}\Bigg[\frac{|\mathcal{V}_{m,m'}|}{|\mathcal{V}_m|} \geq e^{-n\mu/2+2\log 2}\Bigg] + \max_{\{\mathcal{V}_v\}\in\mathcal{G}_\mu} \sum_{m\in\mathcal{M}}\sum_{m'\neq m} \mathbb{P}_{\{\mathcal{V}_v\}}\Bigg[\frac{1}{|\mathcal{V}_m|}\sum_{v\in\mathcal{V}_m\setminus\mathcal{V}_{m,m'}} \times W^n\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'}\big|\mathbf{P}(v)\big) \geq \kappa_n\Bigg] \tag{75}$$

$$\overset{(f)}{\leq} |\mathcal{M}|^2 \exp\Big\{-e^{n(\tilde{R}-\mu)-2\log 2}\Big\} \tag{76}$$

$$\overset{(g)}{\to} 0\,(n\to\infty), \tag{77}$$

where $(a)$ holds by (70), because $n$ exceeds $\eta_0$; $(b)$ follows from (21b); $(c)$ follows from the Union-of-Events bound; $(d)$ follows from (49); $(e)$ follows from (69) and the Union-of-Events bound; $(f)$ holds by (50) and (68); and $(g)$ holds because $|\mathcal{M}| = \exp(\exp(nR))$ and $\mu < \tilde{R} - R$. □

## III. Identification via the BC

In this section we establish the ID capacity region of the two-receiver BC $\big(\mathcal{X}, W(y,z|x), \mathcal{Y}\times\mathcal{Z}\big)$ under the average-error criterion, which requires that each receiver identify the message intended for it reliably in expectation over the uniform ID message intended for the other receiver. We begin with the basic definitions of an average-error ID code for the BC $W(y,z|x)$:

*Definition 8: Fix finite sets $\mathcal{M}_\mathcal{Y}$ and $\mathcal{M}_\mathcal{Z}$, a blocklength $n \in \mathbb{N}$, 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 $\mathcal{X}^n$, with every $m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y}$ an ID set $\mathcal{D}_{m_\mathcal{Y}} \subset \mathcal{Y}^n$, and with every $m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}$ an ID set $\mathcal{D}_{m_\mathcal{Z}} \subset \mathcal{Z}^n$. The collection of tuples $\big\{Q_{m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}_{m_\mathcal{Y}}, \mathcal{D}_{m_\mathcal{Z}}\big\}_{(m_\mathcal{Y},m_\mathcal{Z})\in\mathcal{M}_\mathcal{Y}\times\mathcal{M}_\mathcal{Z}}$ 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 for the BC $W(y,z|x)$ if the maximum probabilities of missed identification at Terminals $\mathcal{Y}$ and $\mathcal{Z}$*

$$p^\mathcal{Y}_{\text{missed-ID}} = \max_{m_\mathcal{Y}\in\mathcal{M}_\mathcal{Y}} \frac{1}{|\mathcal{M}_\mathcal{Z}|} \times \sum_{m_\mathcal{Z}\in\mathcal{M}_\mathcal{Z}} \big(Q_{m_\mathcal{Y},m_\mathcal{Z}} W^n\big)\big(Y^n \notin \mathcal{D}_{m_\mathcal{Y}}\big), \tag{78a}$$

$$p^\mathcal{Z}_{\text{missed-ID}} = \max_{m_\mathcal{Z}\in\mathcal{M}_\mathcal{Z}} \frac{1}{|\mathcal{M}_\mathcal{Y}|} \times \sum_{m_\mathcal{Y}\in\mathcal{M}_\mathcal{Y}} \big(Q_{m_\mathcal{Y},m_\mathcal{Z}} W^n\big)\big(Z^n \notin \mathcal{D}_{m_\mathcal{Z}}\big) \tag{78b}$$

*satisfy*

$$p^\mathcal{Y}_{\text{missed-ID}} \leq \lambda_1^\mathcal{Y}, \tag{79a}$$

$$p^\mathcal{Z}_{\text{missed-ID}} \leq \lambda_1^\mathcal{Z}, \tag{79b}$$

*and the maximum probabilities of wrong identification at Terminals $\mathcal{Y}$ and $\mathcal{Z}$*

$$p^\mathcal{Y}_{\text{wrong-ID}} = \max_{m_\mathcal{Y}\in\mathcal{M}_\mathcal{Y}} \max_{m'_\mathcal{Y}\neq m_\mathcal{Y}} \frac{1}{|\mathcal{M}_\mathcal{Z}|} \times \sum_{m_\mathcal{Z}\in\mathcal{M}_\mathcal{Z}} \big(Q_{m_\mathcal{Y},m_\mathcal{Z}} W^n\big)\big(Y^n \in \mathcal{D}_{m'_\mathcal{Y}}\big), \tag{80a}$$

$$p^{\mathcal{Z}}_{\text{wrong-ID}} = \max_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}} \max_{m'_{\mathcal{Z}}\neq m_{\mathcal{Z}}} \frac{1}{|\mathcal{M}_{\mathcal{Y}}|} \times \sum_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}} \big(Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}}W^n\big)\big(Z^n\in\mathcal{D}_{m'_{\mathcal{Z}}}\big) \quad (80b)$$

*satisfy*

$$p^{\mathcal{Y}}_{\text{wrong-ID}} \le \lambda_2^{\mathcal{Y}}, \quad (81a)$$

$$p^{\mathcal{Z}}_{\text{wrong-ID}} \le \lambda_2^{\mathcal{Z}}. \quad (81b)$$

*A rate-pair $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$ is achievable if for every positive $\lambda_1^{\mathcal{Y}}$, $\lambda_2^{\mathcal{Y}}$, $\lambda_1^{\mathcal{Z}}$, and $\lambda_2^{\mathcal{Z}}$ and for every sufficiently-large blocklength $n$ there exists 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 for the BC with*

$$\begin{cases} \frac{1}{n}\log\log|\mathcal{M}_{\mathcal{Y}}| \ge R_{\mathcal{Y}} & \text{if } R_{\mathcal{Y}} > 0,\\ |\mathcal{M}_{\mathcal{Y}}| = 1 & \text{if } R_{\mathcal{Y}} = 0, \end{cases}$$

$$\begin{cases} \frac{1}{n}\log\log|\mathcal{M}_{\mathcal{Z}}| \ge R_{\mathcal{Z}} & \text{if } R_{\mathcal{Z}} > 0,\\ |\mathcal{M}_{\mathcal{Z}}| = 1 & \text{if } R_{\mathcal{Z}} = 0. \end{cases}$$

*The ID capacity region $\mathcal{C}$ of the BC is the closure of the set of all achievable rate-pairs.*

Equivalently, we can define an ID code for the BC $W(y,z|x)$ as follows:

*Remark 9: Given a collection of PMFs*

$$\big\{Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}}\big\}_{(m_{\mathcal{Y}},m_{\mathcal{Z}})\in\mathcal{M}_{\mathcal{Y}}\times\mathcal{M}_{\mathcal{Z}}}$$

*on $\mathcal{X}^n$, define the mixture PMFs on $\mathcal{X}^n$*

$$Q_{m_{\mathcal{Y}}} = \frac{1}{|\mathcal{M}_{\mathcal{Z}}|}\sum_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}} Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}}, \quad m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}, \quad (82a)$$

$$Q_{m_{\mathcal{Z}}} = \frac{1}{|\mathcal{M}_{\mathcal{Y}}|}\sum_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}} Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}}, \quad m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}. \quad (82b)$$

*The collection of tuples*

$$\big\{Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}}, \mathcal{D}_{m_{\mathcal{Y}}}, \mathcal{D}_{m_{\mathcal{Z}}}\big\}_{(m_{\mathcal{Y}},m_{\mathcal{Z}})\in\mathcal{M}_{\mathcal{Y}}\times\mathcal{M}_{\mathcal{Z}}}$$

*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 for the BC $W(y,z|x)$ iff the following two requirements are met: 1) $\big\{Q_{m_{\mathcal{Y}}}, \mathcal{D}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}}$ is an $\big(n, \mathcal{M}_{\mathcal{Y}}, \lambda_1^{\mathcal{Y}}, \lambda_2^{\mathcal{Y}}\big)$ ID code for the marginal channel $W_{\mathcal{Y}}(y|x)$; and 2) $\big\{Q_{m_{\mathcal{Z}}}, \mathcal{D}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ is an $\big(n, \mathcal{M}_{\mathcal{Z}}, \lambda_1^{\mathcal{Z}}, \lambda_2^{\mathcal{Z}}\big)$ ID code for $W_{\mathcal{Z}}(z|x)$.*

Our main result is a single-letter characterization of the ID capacity region of the BC:

*Theorem 10: The ID capacity region $\mathcal{C}$ of the BC $W(y,z|x)$ is the set of all rate-pairs $(R_{\mathcal{Y}}, R_{\mathcal{Z}}) \in (\mathbb{R}_0^+)^2$ that for some PMF $P$ on $\mathcal{X}$ satisfy*

$$R_{\mathcal{Y}} \le I(P, W_{\mathcal{Y}}), \quad (83a)$$

$$R_{\mathcal{Z}} \le I(P, W_{\mathcal{Z}}). \quad (83b)$$

We prove the direct part in Section III-A and the converse part in Section III-B. In fact, we shall establish the following stronger results:

*Remark 11: The ID capacity region $\mathcal{C}$ of the BC $W(y,z|x)$ is achievable even if we require that the maximum probabilities of missed and wrong identification decay exponentially in the blocklength $n$. And for all sufficiently-large $n$, rate-pairs outside this region can be achieved only if $\lambda_1^{\mathcal{Y}}+\lambda_2^{\mathcal{Y}}+\lambda_1^{\mathcal{Z}}+\lambda_2^{\mathcal{Z}} \ge 1$.*

*Proof:* This follows from Claims 14 and 15 ahead. □

In contrast to transmission via the BC, Theorem 10 implies that for identification via the BC there is no trade-off between Receiver $\mathcal{Y}$ and Receiver $\mathcal{Z}$'s rate. An intuitive explanation for this is that in transmission via the BC the message to the other receiver hurts because it is like noise, whereas here this effect is offset by the benefits afforded by randomization.

Recall that to achieve the ID capacity of a DMC requires stochastic encoders; deterministic encoders cannot achieve any positive ID rate [1]. On the BC this is not true:

*Remark 12: Every rate-pair in the interior of the ID capacity region $\mathcal{C}$ of the BC $W(y,z|x)$ can be achieved using ID codes with deterministic encoders.*

*Proof:* The encoder we construct in Section III-A ahead to prove the direct part of Theorem 10 is deterministic: it maps every ID message-pair to a channel-input sequence that is fully determined by the random code construction. □

As a corollary to Theorem 10, we next observe that the ID capacity region of the BC is convex. This requires proof, because the ID rate is the iterated logarithm of the number of ID messages normalized by the blocklength $n$, and we therefore cannot invoke a time-sharing argument [4, Remark 2].

*Corollary 13: The ID capacity region of the BC $W(y,z|x)$ is convex.*

*Proof:* This readily follows from the fact that the rate region in Theorem 10 is convex. (That the rate region in Theorem 10 is convex is proved in [8, Proof of Corollary 2.3.6].) □

We next prove Theorem 10: Section III-A establishes the direct part and Section III-B a strong converse.

### A. The Direct Part of Theorem 10

In this section we prove the direct part of Theorem 10 by fixing any input distribution $P \in \mathscr{P}(\mathcal{X})$ and any positive ID rate-pair $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$ satisfying

$$0 < R_{\mathcal{Y}} < I(P, W_{\mathcal{Y}}), \quad (84a)$$

$$0 < R_{\mathcal{Z}} < I(P, W_{\mathcal{Z}}) \quad (84b)$$

and showing that the rate-pair $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$ is achievable. We assume that both $I(P, W_{\mathcal{Y}})$ and $I(P, W_{\mathcal{Z}})$ are positive; when they are not, the result follows from Theorem 4. Let $\mathcal{M}_{\mathcal{Y}}$ be a size-$\exp(\exp(nR_{\mathcal{Y}}))$ set of possible ID messages for Terminal $\mathcal{Y}$, and let $\mathcal{M}_{\mathcal{Z}}$ be a size-$\exp(\exp(nR_{\mathcal{Z}}))$ set of possible ID messages for Terminal $\mathcal{Z}$. We next describe our random code construction and show that, for every positive $\lambda_1^{\mathcal{Y}}$, $\lambda_2^{\mathcal{Y}}$, $\lambda_1^{\mathcal{Z}}$, and $\lambda_2^{\mathcal{Z}}$ and for every sufficiently-large blocklength $n$, it produces with high probability 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 for the BC $W(y,z|x)$. The scheme that we propose builds on our code construction for the single-user channel in Section II by making it appear to each receiver as though we were using an instance of the single-user ID code on its marginal channel.

*Code Generation:* Fix an expected bin rate $\tilde{R}_{\mathcal{Y}}$ for Terminal $\mathcal{Y}$, an expected bin rate $\tilde{R}_{\mathcal{Z}}$ for Terminal $\mathcal{Z}$, and a pool rate $R_{\mathcal{P}}$ satisfying

$$R_{\mathcal{Y}} < \tilde{R}_{\mathcal{Y}} < I(P, W_{\mathcal{Y}}), \quad (85a)$$

$$R_{\mathcal{Z}} < \tilde{R}_{\mathcal{Z}} < I(P, W_{\mathcal{Z}}), \tag{85b}$$
$$\tilde{R}_{\mathcal{Y}} < R_{\mathcal{P}}, \tag{85c}$$
$$\tilde{R}_{\mathcal{Z}} < R_{\mathcal{P}}, \tag{85d}$$
$$R_{\mathcal{P}} < \tilde{R}_{\mathcal{Y}} + \tilde{R}_{\mathcal{Z}}. \tag{85e}$$

This is possible by (84). Draw $e^{nR_{\mathcal{P}}}$ $n$-tuples $\sim P^n$ independently and place them in a pool $\boldsymbol{\mathcal{P}}$. Index the $n$-tuples in the pool by the elements of a size-$e^{nR_{\mathcal{P}}}$ set $\mathcal{V}$, e.g., $\{1, \ldots, e^{nR_{\mathcal{P}}}\}$, and denote by $\mathbf{P}(v)$ the $n$-tuple in $\boldsymbol{\mathcal{P}}$ that is indexed by $v \in \mathcal{V}$. For each receiving terminal $\Psi \in \{\mathcal{Y}, \mathcal{Z}\}$ associate with each ID message $m_\Psi \in \mathcal{M}_\Psi$ an index-set $\boldsymbol{\mathcal{V}}_{m_\Psi}$ and a bin $\boldsymbol{\mathcal{B}}_{m_\Psi}$ as follows. Select each element of $\mathcal{V}$ for inclusion in $\boldsymbol{\mathcal{V}}_{m_\Psi}$ independently with probability $e^{-n(R_{\mathcal{P}} - \tilde{R}_\Psi)}$, and let Bin $\boldsymbol{\mathcal{B}}_{m_\Psi}$ be the multiset that contains all the $n$-tuples in the pool that are indexed by $\boldsymbol{\mathcal{V}}_{m_\Psi}$,

$$\boldsymbol{\mathcal{B}}_{m_\Psi} = \big\{\mathbf{P}(v),\, v \in \boldsymbol{\mathcal{V}}_{m_\Psi}\big\}.$$

(Bin $\boldsymbol{\mathcal{B}}_{m_\Psi}$ is thus of expected size $e^{n\tilde{R}_\Psi}$.) Associate with each ID message-pair $(m_{\mathcal{Y}}, m_{\mathcal{Z}}) \in \mathcal{M}_{\mathcal{Y}} \times \mathcal{M}_{\mathcal{Z}}$ an index $V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}}$ as follows. If $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} \cap \boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}$ is not empty, then draw $V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}}$ uniformly over $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} \cap \boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}$. Otherwise draw $V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}}$ uniformly over $\mathcal{V}$. Reveal the pool $\boldsymbol{\mathcal{P}}$, the index-sets $\big\{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}}$ and $\big\{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}}$, the corresponding bins $\big\{\boldsymbol{\mathcal{B}}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}}$ and $\big\{\boldsymbol{\mathcal{B}}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}}$, and the indices $\big\{V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}}\big\}_{(m_{\mathcal{Y}}, m_{\mathcal{Z}}) \in \mathcal{M}_{\mathcal{Y}} \times \mathcal{M}_{\mathcal{Z}}}$ to all parties. The encoding and decoding are determined by

$$\boldsymbol{\mathcal{C}} = \Big(\boldsymbol{\mathcal{P}}, \big\{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}}, \big\{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}}, \big\{V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}}\big\}_{(m_{\mathcal{Y}}, m_{\mathcal{Z}}) \in \mathcal{M}_{\mathcal{Y}} \times \mathcal{M}_{\mathcal{Z}}}\Big). \tag{86}$$

*Encoding:* To send ID Message-Pair $(m_{\mathcal{Y}}, m_{\mathcal{Z}}) \in \mathcal{M}_{\mathcal{Y}} \times \mathcal{M}_{\mathcal{Z}}$, the encoder transmits the sequence $\mathbf{P}(V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}})$. ID Message-Pair $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$ is thus associated with the $\{0, 1\}$-valued PMF

$$\boldsymbol{\mathcal{Q}}_{m_{\mathcal{Y}}, m_{\mathcal{Z}}}(\mathbf{x}) = \mathbb{1}_{\mathbf{x} = \mathbf{P}(V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}})}, \quad \mathbf{x} \in \mathcal{X}^n. \tag{87}$$

Note that once the code (86) has been constructed, the encoder is deterministic: it maps ID Message-Pair $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$ to the $(m_{\mathcal{Y}}, m_{\mathcal{Z}})$-codeword $\mathbf{P}(V_{m_{\mathcal{Y}}, m_{\mathcal{Z}}})$.

*Decoding:* In this section the function $\delta(\cdot)$ maps every non-negative real number $u$ to $uH(P \times W)$. The decoders choose $\epsilon > 0$ sufficiently small so that $2\delta(\epsilon) < I(P, W_{\mathcal{Y}}) - \tilde{R}_{\mathcal{Y}}$ and $2\delta(\epsilon) < I(P, W_{\mathcal{Z}}) - \tilde{R}_{\mathcal{Z}}$. The $m'_{\mathcal{Y}}$-focused party at Terminal $\mathcal{Y}$ guesses that $m'_{\mathcal{Y}}$ was sent iff for some index $v \in \boldsymbol{\mathcal{V}}_{m'_{\mathcal{Y}}}$ the $n$-tuple $\mathbf{P}(v)$ in Bin $\boldsymbol{\mathcal{B}}_{m'_{\mathcal{Y}}}$ is jointly $\epsilon$-typical with the Terminal-$\mathcal{Y}$ output-sequence $Y^n$, i.e., iff $\big(\mathbf{P}(v), Y^n\big) \in \mathcal{T}_\epsilon^{(n)}(P \times W_{\mathcal{Y}})$ for some $v \in \boldsymbol{\mathcal{V}}_{m'_{\mathcal{Y}}}$. The set $\boldsymbol{\mathcal{D}}_{m'_{\mathcal{Y}}}$ of Terminal-$\mathcal{Y}$ output-sequences $\mathbf{y} \in \mathcal{Y}^n$ that result in the guess "$m'_{\mathcal{Y}}$ was sent" is thus

$$\boldsymbol{\mathcal{D}}_{m'_{\mathcal{Y}}} = \bigcup_{v \in \boldsymbol{\mathcal{V}}_{m'_{\mathcal{Y}}}} \mathcal{T}_\epsilon^{(n)}\big(P \times W_{\mathcal{Y}} \big| \mathbf{P}(v)\big). \tag{88}$$

Likewise, the $m'_{\mathcal{Z}}$-focused party at Terminal $\mathcal{Z}$ guesses that $m'_{\mathcal{Z}}$ was sent iff $\big(\mathbf{P}(v), Z^n\big) \in \mathcal{T}_\epsilon^{(n)}(P \times W_{\mathcal{Z}})$ for some $v \in \boldsymbol{\mathcal{V}}_{m'_{\mathcal{Z}}}$. The set $\boldsymbol{\mathcal{D}}_{m'_{\mathcal{Z}}}$ of Terminal-$\mathcal{Z}$ output-sequences $\mathbf{z} \in \mathcal{Z}^n$ that result in the guess "$m'_{\mathcal{Z}}$ was sent" is thus

$$\boldsymbol{\mathcal{D}}_{m'_{\mathcal{Z}}} = \bigcup_{v \in \boldsymbol{\mathcal{V}}_{m'_{\mathcal{Z}}}} \mathcal{T}_\epsilon^{(n)}\big(P \times W_{\mathcal{Z}} \big| \mathbf{P}(v)\big). \tag{89}$$

*Analysis of the Probabilities of Missed and Wrong Identification:* We first note that $\boldsymbol{\mathcal{C}}$ of (86) (together with the fixed blocklength $n$ and the chosen $\epsilon$) fully specifies the encoding and guessing rules. That is, the randomly constructed ID code

$$\big\{\boldsymbol{\mathcal{Q}}_{m_{\mathcal{Y}}, m_{\mathcal{Z}}}, \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}, \boldsymbol{\mathcal{D}}_{m_{\mathcal{Z}}}\big\}_{(m_{\mathcal{Y}}, m_{\mathcal{Z}}) \in \mathcal{M}_{\mathcal{Y}} \times \mathcal{M}_{\mathcal{Z}}} \tag{90}$$

is fully specified by $\boldsymbol{\mathcal{C}}$. Let $\mathbb{P}$ be the distribution of $\boldsymbol{\mathcal{C}}$, and let $\mathbb{E}$ denote expectation w.r.t. $\mathbb{P}$. Subscripts indicate conditioning on the event that some of the chance variables assume the values indicated by the subscripts, e.g., $\mathbb{P}_{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}}$ denotes the distribution conditional on $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} = \mathcal{V}_{m_{\mathcal{Y}}}$, and $\mathbb{E}_{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}}$ denotes the expectation w.r.t. $\mathbb{P}_{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}}$.

The maximum probabilities of missed and wrong identification of the randomly constructed ID code are the random variables

$$P^{\mathcal{Y}}_{\text{missed-ID}} = \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \frac{1}{|\mathcal{M}_{\mathcal{Z}}|} \times \sum_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \big(\boldsymbol{\mathcal{Q}}_{m_{\mathcal{Y}}, m_{\mathcal{Z}}} W^n\big)\big(Y^n \notin \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}\big), \tag{91a}$$

$$P^{\mathcal{Z}}_{\text{missed-ID}} = \max_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \frac{1}{|\mathcal{M}_{\mathcal{Y}}|} \times \sum_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \big(\boldsymbol{\mathcal{Q}}_{m_{\mathcal{Y}}, m_{\mathcal{Z}}} W^n\big)\big(Z^n \notin \boldsymbol{\mathcal{D}}_{m_{\mathcal{Z}}}\big), \tag{91b}$$

$$P^{\mathcal{Y}}_{\text{wrong-ID}} = \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \max_{m'_{\mathcal{Y}} \neq m_{\mathcal{Y}}} \frac{1}{|\mathcal{M}_{\mathcal{Z}}|} \times \sum_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \big(\boldsymbol{\mathcal{Q}}_{m_{\mathcal{Y}}, m_{\mathcal{Z}}} W^n\big)\big(Y^n \in \boldsymbol{\mathcal{D}}_{m'_{\mathcal{Y}}}\big), \tag{91c}$$

$$P^{\mathcal{Z}}_{\text{wrong-ID}} = \max_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \max_{m'_{\mathcal{Z}} \neq m_{\mathcal{Z}}} \frac{1}{|\mathcal{M}_{\mathcal{Y}}|} \times \sum_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \big(\boldsymbol{\mathcal{Q}}_{m_{\mathcal{Y}}, m_{\mathcal{Z}}} W^n\big)\big(Z^n \in \boldsymbol{\mathcal{D}}_{m'_{\mathcal{Z}}}\big). \tag{91d}$$

They are fully specified by $\boldsymbol{\mathcal{C}}$. To prove that for every choice of $\lambda_1^{\mathcal{Y}}, \lambda_2^{\mathcal{Y}}, \lambda_1^{\mathcal{Z}}, \lambda_2^{\mathcal{Z}} > 0$ and $n$ sufficiently large the collection of tuples (90) is with high probability 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 for the BC $W(y, z|x)$, we prove the following stronger result:

*Claim 14: The maximum probabilities of missed and wrong identification of the randomly constructed ID code* (90)*, $P^{\mathcal{Y}}_{\text{missed-ID}}$, $P^{\mathcal{Z}}_{\text{missed-ID}}$, $P^{\mathcal{Y}}_{\text{wrong-ID}}$, and $P^{\mathcal{Z}}_{\text{wrong-ID}}$, converge in probability to zero exponentially in the blocklength n, i.e.,*

$$\exists\, \tau > 0 \text{ s.t. } \lim_{n \to \infty} \mathbb{P}\Big[\max\big\{P^{\mathcal{Y}}_{\text{missed-ID}}, P^{\mathcal{Z}}_{\text{missed-ID}}, P^{\mathcal{Y}}_{\text{wrong-ID}}, P^{\mathcal{Z}}_{\text{wrong-ID}}\big\} \geq e^{-n\tau}\Big] = 0. \tag{92}$$

*Proof:* We will prove that

$$\exists\, \tau > 0 \text{ s.t. } \lim_{n \to \infty} \mathbb{P}\Big[\max\big\{P^{\mathcal{Y}}_{\text{missed-ID}}, P^{\mathcal{Y}}_{\text{wrong-ID}}\big\} \geq e^{-n\tau}\Big] = 0. \tag{93}$$

By swapping $\mathcal{Z}$ and $\mathcal{Y}$ throughout the proof it will then follow that (93) also holds when we replace $\mathcal{Y}$ with $\mathcal{Z}$, and (92) will then follow using the Union-of-Events bound.

To prove (93), we consider for each $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ two distributions on the set $\mathcal{V}$, which indexes the pool $\mathcal{P}$. We fix some $v^\star \in \mathcal{V}$ and define for every $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ the PMFs on $\mathcal{V}$

$$\boldsymbol{P}_V^{(m_{\mathcal{Y}})}(v) = \frac{1}{|\mathcal{M}_{\mathcal{Z}}|} \sum_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \mathbb{1}_{v=V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}}, \quad v \in \mathcal{V}, \tag{94a}$$

$$\tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}(v) = \begin{cases} \frac{1}{|\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}|} \sum_{v' \in \boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}} \mathbb{1}_{v=v'} & \text{if } \boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} \neq \emptyset, \\ \mathbb{1}_{v=v^\star} & \text{otherwise,} \end{cases} \quad v \in \mathcal{V}. \tag{94b}$$

The latter PMF is reminiscent of the distribution we encountered in (17) and (18) in the single-user case. The former is related to the BC setting when we view $M_{\mathcal{Z}}$ as uniform over $\mathcal{M}_{\mathcal{Z}}$. As we argue next, to establish (93) it suffices to show that the two are similar in the sense that

$$\exists\, \tau > 0 \text{ s.t. } \lim_{n\to\infty} \mathbb{P}\left[\max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} d\left(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\right) \geq e^{-n\tau}\right] = 0. \tag{95}$$

To see why, let us define for every $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ the PMFs on $\mathcal{X}^n$

$$\boldsymbol{Q}_{m_{\mathcal{Y}}}(\mathbf{x}) = \frac{1}{|\mathcal{M}_{\mathcal{Z}}|} \sum_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \boldsymbol{Q}_{m_{\mathcal{Y}},m_{\mathcal{Z}}}(\mathbf{x}), \quad \mathbf{x} \in \mathcal{X}^n, \tag{96a}$$

$$\tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}}(\mathbf{x}) = \begin{cases} \frac{1}{|\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}|} \sum_{v' \in \boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}} \mathbb{1}_{\mathbf{x}=\mathbf{P}(v')} & \text{if } \boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} \neq \emptyset, \\ \mathbb{1}_{\mathbf{x}=\mathbf{P}(v^\star)} & \text{otherwise,} \end{cases} \quad \mathbf{x} \in \mathcal{X}^n. \tag{96b}$$

The collection of tuples $\left\{\boldsymbol{Q}_{m_{\mathcal{Y}}}, \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}\right\}_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}}$ can be viewed as a randomly constructed ID code for the DMC $W_{\mathcal{Y}}(y|x)$ with maximum probability of missed identification

$$\begin{aligned} &\max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \left(\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n\right)\left(Y^n \notin \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}\right) \\ &\quad = \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \frac{1}{|\mathcal{M}_{\mathcal{Z}}|} \sum_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \left(\boldsymbol{Q}_{m_{\mathcal{Y}},m_{\mathcal{Z}}} W^n\right)\left(Y^n \notin \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}\right) \\ &\quad = P^{\mathcal{Y}}_{\text{missed-ID}} \end{aligned} \tag{97}$$

and maximum probability of wrong identification

$$\begin{aligned} &\max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \max_{m'_{\mathcal{Y}} \neq m_{\mathcal{Y}}} \left(\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n\right)\left(Y^n \in \boldsymbol{\mathcal{D}}_{m'_{\mathcal{Y}}}\right) \\ &\quad = \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \max_{m'_{\mathcal{Y}} \neq m_{\mathcal{Y}}} \frac{1}{|\mathcal{M}_{\mathcal{Z}}|} \\ &\qquad \times \sum_{m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}} \left(\boldsymbol{Q}_{m_{\mathcal{Y}},m_{\mathcal{Z}}} W^n\right)\left(Y^n \in \boldsymbol{\mathcal{D}}_{m'_{\mathcal{Y}}}\right) \\ &\quad = P^{\mathcal{Y}}_{\text{wrong-ID}}. \end{aligned} \tag{98, 99}$$

And $\left\{\tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}}, \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}\right\}_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}}$ has the same law as the randomly constructed ID code $\{\boldsymbol{Q}_m, \boldsymbol{\mathcal{D}}_m\}_{m \in \mathcal{M}}$ of Section II for the DMC $W = W_{\mathcal{Y}}$ with blocklength $n$, fixed element $v^\star$ of $\mathcal{V}$, decoding parameter $\epsilon$, size-$\exp(\exp(nR_{\mathcal{Y}}))$ set $\mathcal{M}_{\mathcal{Y}}$ of possible ID messages, expected bin rate $\tilde{R}_{\mathcal{Y}}$, and pool rate $R_{\mathcal{P}}$. (Note that $\epsilon$, $R_{\mathcal{Y}}$, $\tilde{R}_{\mathcal{Y}}$, and $R_{\mathcal{P}}$ are eligible for the random code construction in Section II, because $\epsilon$ is positive and sufficiently small so that $2\epsilon H(P \times W_{\mathcal{Y}}) < I(P, W_{\mathcal{Y}}) - \tilde{R}_{\mathcal{Y}}$, and because of (84) and (85).) Let $\tilde{P}^{\mathcal{Y}}_{\text{missed-ID}}$ and $\tilde{P}^{\mathcal{Y}}_{\text{wrong-ID}}$ denote the maximum probabilities of missed and wrong identification of the randomly constructed ID code $\left\{\tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}}, \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}\right\}_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}}$, i.e.,

$$\tilde{P}^{\mathcal{Y}}_{\text{missed-ID}} = \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \left(\tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n\right)\left(Y^n \notin \boldsymbol{\mathcal{D}}_{m_{\mathcal{Y}}}\right), \tag{100a}$$

$$\tilde{P}^{\mathcal{Y}}_{\text{wrong-ID}} = \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} \max_{m'_{\mathcal{Y}} \neq m_{\mathcal{Y}}} \left(\tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n\right)\left(Y^n \in \boldsymbol{\mathcal{D}}_{m'_{\mathcal{Y}}}\right). \tag{100b}$$

By Claim 6 on the single-user channel

$$\exists\, \tau > 0 \text{ s.t. } \lim_{n\to\infty} \mathbb{P}\left[\max\left\{\tilde{P}^{\mathcal{Y}}_{\text{missed-ID}}, \tilde{P}^{\mathcal{Y}}_{\text{wrong-ID}}\right\} \geq e^{-n\tau}\right] = 0. \tag{101}$$

And by definition of the Total-Variation distance

$$P^{\mathcal{Y}}_{\text{missed-ID}} \leq \tilde{P}^{\mathcal{Y}}_{\text{missed-ID}} + \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} d\left(\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}, \tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}\right), \tag{102a}$$

$$P^{\mathcal{Y}}_{\text{wrong-ID}} \leq \tilde{P}^{\mathcal{Y}}_{\text{wrong-ID}} + \max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} d\left(\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}, \tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}\right). \tag{102b}$$

For every $\tau_1$, $\tau_2$, and $\tau < \min\{\tau_1, \tau_2\}$ we have for all sufficiently-large $n$,

$$e^{-n\tau_1} + e^{-n\tau_2} \leq e^{-n\tau}. \tag{103}$$

This, combined with the Union-of-Events bound, (101), and (102), implies that to establish (93) it suffices to show that

$$\exists\, \tau > 0 \text{ s.t. } \lim_{n\to\infty} \mathbb{P}\left[\max_{m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}} d\left(\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}, \tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}\right) \geq e^{-n\tau}\right] = 0. \tag{104}$$

Consequently, to prove our claim that (95) implies (93), we only have to show that (95) implies (104). To that end define the conditional PMF

$$\boldsymbol{P}_{X^n|V}(\mathbf{x}|v) = \mathbb{1}_{\mathbf{x}=\mathbf{P}(v)}, \quad (\mathbf{x}, v) \in \mathcal{X}^n \times \mathcal{V}, \tag{105}$$

and note that for every $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$

$$\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}} = \boldsymbol{P}_V^{(m_{\mathcal{Y}})} \boldsymbol{P}_{X^n|V} W^n_{\mathcal{Y}}, \tag{106a}$$

$$\tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}} = \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})} \boldsymbol{P}_{X^n|V} W^n_{\mathcal{Y}}, \tag{106b}$$

where we used (94), (96), and (105), and in the first equality also (87). We can now upper-bound $d\left(\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}, \tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}\right)$ by

$$\begin{aligned} &d\left(\boldsymbol{Q}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}, \tilde{\boldsymbol{Q}}_{m_{\mathcal{Y}}} W^n_{\mathcal{Y}}\right) \\ &\quad = d\left(\boldsymbol{P}_V^{(m_{\mathcal{Y}})} \boldsymbol{P}_{X^n|V} W^n_{\mathcal{Y}}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})} \boldsymbol{P}_{X^n|V} W^n_{\mathcal{Y}}\right) \qquad (107) \\ &\quad \leq d\left(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\right), \qquad (108) \end{aligned}$$

where the last inequality follows from the Data-Processing inequality for the Total-Variation distance [13, Lemma 1]. From (108) we conclude that (95) implies (104) and hence also (93).

Having established that (95) implies (93), it remains to prove (95). Before we do that, we give an intuitive explanation

why (95) holds. Fix $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ and a realization $\mathcal{V}_{m_{\mathcal{Y}}}$ of the corresponding index-set $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}$, and assume that $\mathcal{V}_{m_{\mathcal{Y}}} \approx e^{n\tilde{R}_{\mathcal{Y}}}$. For every $m_{\mathcal{Z}} \in \mathcal{M}_{\mathcal{Z}}$, the probabilitiy that the intersection of $\mathcal{V}_{m_{\mathcal{Y}}}$ and $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}$ is empty is very small, and if the intersection is nonempty, then, by our random construction of $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}$ and $V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}$, the codeword-index $V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}$ is drawn uniformly at random from $\mathcal{V}_{m_{\mathcal{Y}}}$. Because $\mathcal{V}_{m_{\mathcal{Y}}}$ is exponential in $n$ and the cardinality of $\mathcal{M}_{\mathcal{Z}}$ is double-exponential in $n$, and because, by our random construction of $\{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ and $\{V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$, the codeword-indices $\{V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ are drawn independently of each other, (95) can be derived using concentration inequalities.

To prove (95) rigorously, fix some $\mu$ satisfying

$$0 < \mu < \tilde{R}_{\mathcal{Y}} - R_{\mathcal{Y}}, \tag{109}$$

and let

$$\delta_n = e^{-n\mu/2}. \tag{110}$$

Introduce the set $\mathcal{H}_\mu^{\mathcal{Y}}$ comprising the realizations $\{\mathcal{V}_\upsilon\}_{\upsilon\in\mathcal{M}_{\mathcal{Y}}}$ of the index-sets $\{\boldsymbol{\mathcal{V}}_\upsilon\}_{\upsilon\in\mathcal{M}_{\mathcal{Y}}}$ satisfying

$$|\mathcal{V}_\upsilon| > (1-\delta_n)e^{n\tilde{R}_{\mathcal{Y}}},\ \forall\, \upsilon \in \mathcal{M}_{\mathcal{Y}}. \tag{111}$$

We upper-bound $\max_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}} d\big(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\big)$ differently depending on whether or not $\{\boldsymbol{\mathcal{V}}_\upsilon\}$ is in $\mathcal{H}_\mu^{\mathcal{Y}}$, where $\{\boldsymbol{\mathcal{V}}_\upsilon\}$ is short for $\{\boldsymbol{\mathcal{V}}_\upsilon\}_{\upsilon\in\mathcal{M}_{\mathcal{Y}}}$. If $\{\boldsymbol{\mathcal{V}}_\upsilon\} \notin \mathcal{H}_\mu^{\mathcal{Y}}$, then we upper-bound it by one (which is an upper bound on the Total-Variation distance between any two probability measures) to obtain for every $\tau > 0$

$$\begin{aligned}
&\mathbb{P}\Bigg[\max_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}} d\Big(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\Big) \ge e^{-n\tau}\Bigg] \\
&\quad\le \mathbb{P}\Big[\{\boldsymbol{\mathcal{V}}_\upsilon\} \notin \mathcal{H}_\mu^{\mathcal{Y}}\Big] + \sum_{\{\mathcal{V}_\upsilon\}\in\mathcal{H}_\mu^{\mathcal{Y}}} \mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_\upsilon\} = \{\mathcal{V}_\upsilon\}\big] \\
&\qquad\times \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\Bigg[\max_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}} d\Big(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\Big) \ge e^{-n\tau}\Bigg].
\end{aligned} \tag{112}$$

We consider the two terms on the RHS of (112) separately, beginning with $\mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_\upsilon\} \notin \mathcal{H}_\mu^{\mathcal{Y}}\big]$. Following the proof of Lemma 5 in Section II, we will show that $\mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_\upsilon\} \notin \mathcal{H}_\mu^{\mathcal{Y}}\big]$ converges to zero as $n$ tends to infinity. This does not follow from Lemma 5, because here we require $\mu$ to satisfy (109) instead of the more restrictive condition (24) of Section II. For every fixed $\upsilon \in \mathcal{M}_{\mathcal{Y}}$ the $e^{nR_{\mathcal{P}}}$ binary random variables $\{\mathbb{1}_{\upsilon\in\boldsymbol{\mathcal{V}}_\upsilon}\}_{\upsilon\in\mathcal{V}}$ are IID, and

$$\mathbb{E}\Bigg[\sum_{\upsilon\in\mathcal{V}} \mathbb{1}_{\upsilon\in\boldsymbol{\mathcal{V}}_\upsilon}\Bigg] = \sum_{\upsilon\in\mathcal{V}} \mathbb{P}[\upsilon \in \boldsymbol{\mathcal{V}}_\upsilon] = e^{n\tilde{R}_{\mathcal{Y}}}. \tag{113}$$

Consequently, by the multiplicative Chernoff bound (6a) in Proposition 1,

$$\mathbb{P}\Big[|\boldsymbol{\mathcal{V}}_\upsilon| \le (1-\delta_n)\, e^{n\tilde{R}_{\mathcal{Y}}}\Big] = \mathbb{P}\Bigg[\sum_{\upsilon\in\mathcal{V}} \mathbb{1}_{\upsilon\in\boldsymbol{\mathcal{V}}_\upsilon} \le (1-\delta_n)\, e^{n\tilde{R}_{\mathcal{Y}}}\Bigg] \tag{114}$$

$$\le \exp\Big\{-\delta_n^2\, e^{n\tilde{R}_{\mathcal{Y}}-\log 2}\Big\} \tag{115}$$

$$= \exp\Big\{-e^{n(\tilde{R}_{\mathcal{Y}}-\mu)-\log 2}\Big\}. \tag{116}$$

The Union-of-Events bound thus implies that

$$\mathbb{P}\Big[\{\boldsymbol{\mathcal{V}}_\upsilon\} \notin \mathcal{H}_\mu^{\mathcal{Y}}\Big] \le |\mathcal{M}_{\mathcal{Y}}| \exp\Big\{-e^{n(\tilde{R}_{\mathcal{Y}}-\mu)-\log 2}\Big\} \tag{117}$$

$$\overset{(a)}{\to} 0\ (n\to\infty), \tag{118}$$

where $(a)$ holds because $|\mathcal{M}_{\mathcal{Y}}| = \exp(\exp(nR_{\mathcal{Y}}))$ and by (109).

Having established (118), we return to (112) and conclude the proof of (95) by showing that

$$\begin{aligned}
&\exists\, \tau > 0 \text{ s.t.} \\
&\lim_{n\to\infty} \max_{\{\mathcal{V}_\upsilon\}\in\mathcal{H}_\mu^{\mathcal{Y}}} \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\Bigg[\max_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}} d\Big(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\Big) \ge e^{-n\tau}\Bigg] \\
&\quad = 0.
\end{aligned} \tag{119}$$

(The proof of (119) ahead exploits the fact that the index-sets $\big\{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ are drawn at random. Likewise, when we prove (93) with $\mathcal{Y}$ replaced by $\mathcal{Z}$, we shall need the fact that the index-sets $\big\{\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}}$ are drawn at random. Hence Remark 7.) To prove (119), let us henceforth assume that $n$ is large enough so that the following two inequalities hold:

$$(1-\delta_n)e^{n\tilde{R}_{\mathcal{Y}}} \ge 1, \tag{120a}$$

$$\delta_n \le 1/2, \tag{120b}$$

where $\delta_n$ is defined in (110). (This is possible, because $\delta_n$ converges to zero as $n$ tends to infinity and $\tilde{R}_{\mathcal{Y}} > 0$.) Fix any realization $\{\mathcal{V}_\upsilon\}$ in $\mathcal{H}_\mu^{\mathcal{Y}}$. Rather than directly upper-bounding the maximum over $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ of $d\big(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\big)$ under $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$, we first consider $d\big(\boldsymbol{P}_V^{(m_{\mathcal{Y}})}, \tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\big)$ for a fixed $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$. By (111) (which holds because $\{\mathcal{V}_\upsilon\} \in \mathcal{H}_\mu^{\mathcal{Y}}$) and (120a), $\mathcal{V}_{m_{\mathcal{Y}}}$ is nonempty. For every fixed $\upsilon \in \mathcal{V}\setminus\mathcal{V}_{m_{\mathcal{Y}}}$ we therefore have that under $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$ the $\exp(\exp(nR_{\mathcal{Z}}))$ binary random variables $\big\{\mathbb{1}_{\upsilon=V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}}\big\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ are IID of mean

$$\begin{aligned}
&\mathbb{E}_{\{\mathcal{V}_\upsilon\}}\big[\mathbb{1}_{\upsilon=V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}}\big] \\
&\quad= \mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[V_{m_{\mathcal{Y}},m_{\mathcal{Z}}} = \upsilon\big] && (121)\\
&\quad\overset{(a)}{=} \frac{1}{|\mathcal{V}|}\mathbb{P}_{\{\mathcal{V}_\upsilon\}}\big[\mathcal{V}_{m_{\mathcal{Y}}} \cap \boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}} = \emptyset\big] && (122)\\
&\quad\overset{(b)}{=} \frac{1}{|\mathcal{V}|}\mathbb{P}\big[\mathcal{V}_{m_{\mathcal{Y}}} \cap \boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}} = \emptyset\big] && (123)\\
&\quad\overset{(c)}{=} \frac{1}{|\mathcal{V}|}\Big(1 - e^{-n(R_{\mathcal{P}}-\tilde{R}_{\mathcal{Z}})}\Big)^{|\mathcal{V}_{m_{\mathcal{Y}}}|} && (124)\\
&\quad\overset{(d)}{\le} \exp\Big\{-e^{-n(R_{\mathcal{P}}-\tilde{R}_{\mathcal{Z}})}\,|\mathcal{V}_{m_{\mathcal{Y}}}| - nR_{\mathcal{P}}\Big\} && (125)\\
&\quad\overset{(e)}{\le} (1-\delta_n)^{-1}\exp\Big\{-(1-\delta_n)e^{n(\tilde{R}_{\mathcal{Y}}+\tilde{R}_{\mathcal{Z}}-R_{\mathcal{P}})} - n\tilde{R}_{\mathcal{Y}}\Big\}, \\
&\qquad \upsilon \in \mathcal{V}\setminus\mathcal{V}_{m_{\mathcal{Y}}} && (126)
\end{aligned}$$

with the following justification. Equality $(a)$ holds because $\upsilon \notin \mathcal{V}_{m_{\mathcal{Y}}}$ and $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} = \mathcal{V}_{m_{\mathcal{Y}}}$ $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$-almost-surely, and therefore: if $\mathcal{V}_{m_{\mathcal{Y}}} \cap \boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}} \neq \emptyset$, then $V_{m_{\mathcal{Y}},m_{\mathcal{Z}}} \neq \upsilon$, and otherwise $V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}$ is uniform over $\mathcal{V}$. Equality $(b)$ holds because $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}$ is independent of $\{\boldsymbol{\mathcal{V}}_\upsilon\}_{\upsilon\in\mathcal{M}_{\mathcal{Y}}}$, and its distribution w.r.t. $\mathbb{P}_{\{\mathcal{V}_\upsilon\}}$ is thus the same as w.r.t. $\mathbb{P}$; $(c)$ holds because we have selected each element of $\mathcal{V}$ for inclusion in $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Z}}}$ independently with

probability $e^{-n(R_{\mathcal{P}}-\tilde{R}_{\mathcal{Z}})}$; ($d$) holds because $|\mathcal{V}| = e^{nR_{\mathcal{P}}}$ and because

$$1-\xi \le e^{-\xi}, \quad \xi \in \mathbb{R}; \tag{127}$$

and ($e$) holds because $0 \le \delta_n < 1$, by (111) (which holds because $\{\mathcal{V}_v\} \in \mathcal{H}_\mu^{\mathcal{Y}}$), and because $\tilde{R}_{\mathcal{Y}} < R_{\mathcal{P}}$. Similarly, for every fixed $v \in \mathcal{V}_{m_{\mathcal{Y}}}$ we have that under $\mathbb{P}_{\{\mathcal{V}_v\}}$ the $\exp(\exp(nR_{\mathcal{Z}}))$ binary random variables $\left\{\mathbb{1}_{v=V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}}\right\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ are IID of mean

$$\begin{aligned}
&\mathbb{E}_{\{\mathcal{V}_v\}}\left[\mathbb{1}_{v=V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}}\right] \\
&\quad= \mathbb{P}_{\{\mathcal{V}_v\}}\left[V_{m_{\mathcal{Y}},m_{\mathcal{Z}}} = v\right] && (128)\\
&\quad\overset{(a)}{=} \frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}\mathbb{P}_{\{\mathcal{V}_v\}}\left[V_{m_{\mathcal{Y}},m_{\mathcal{Z}}} \in \mathcal{V}_{m_{\mathcal{Y}}}\right] && (129)\\
&\quad= \frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}\left(1-\mathbb{P}_{\{\mathcal{V}_v\}}\left[V_{m_{\mathcal{Y}},m_{\mathcal{Z}}} \notin \mathcal{V}_{m_{\mathcal{Y}}}\right]\right) && (130)\\
&\quad\overset{(b)}{=} \frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}\left(1-\frac{|\mathcal{V}|-|\mathcal{V}_{m_{\mathcal{Y}}}|}{|\mathcal{V}|}\left(1-e^{-n(R_{\mathcal{P}}-\tilde{R}_{\mathcal{Z}})}\right)^{|\mathcal{V}_{m_{\mathcal{Y}}}|}\right) \\
&\quad= \frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}-\left(\frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}-\frac{1}{|\mathcal{V}|}\right)\left(1-e^{-n(R_{\mathcal{P}}-\tilde{R}_{\mathcal{Z}})}\right)^{|\mathcal{V}_{m_{\mathcal{Y}}}|} \\
&\quad\overset{(c)}{\in} \left[\frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}\left(1-\exp\left\{-e^{-n(R_{\mathcal{P}}-\tilde{R}_{\mathcal{Z}})}\,|\mathcal{V}_{m_{\mathcal{Y}}}|\right\}\right), \frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}\right] \\
&\quad\overset{(d)}{\subseteq} \left[\frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}-(1-\delta_n)^{-1}\exp\left\{-(1-\delta_n)e^{n(\tilde{R}_{\mathcal{Y}}+\tilde{R}_{\mathcal{Z}}-R_{\mathcal{P}})}\right.\right. \\
&\qquad\left.\left.-n\tilde{R}_{\mathcal{Y}}\right\}, \frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}\right], \quad v \in \mathcal{V}_{m_{\mathcal{Y}}}, && (131)
\end{aligned}$$

where ($a$) holds by symmetry; ($b$) holds by (124) and because $|\mathcal{V}\setminus\mathcal{V}_{m_{\mathcal{Y}}}| = |\mathcal{V}| - |\mathcal{V}_{m_{\mathcal{Y}}}|$; ($c$) holds by (127); and ($d$) holds by (111) (which holds because $\{\mathcal{V}_v\} \in \mathcal{H}_\mu^{\mathcal{Y}}$). Fix some $\kappa$ satisfying

$$0 < \kappa < \min\left\{R_{\mathcal{Z}}, \tilde{R}_{\mathcal{Y}} + \tilde{R}_{\mathcal{Z}} - R_{\mathcal{P}}\right\}, \tag{132}$$

and let

$$\xi_n = 4\exp\left\{-e^{n\kappa-\log 2}\right\}. \tag{133}$$

By (120b)

$$\xi_n/2 > (1-\delta_n)^{-1} \times \exp\left\{-(1-\delta_n)e^{n(\tilde{R}_{\mathcal{Y}}+\tilde{R}_{\mathcal{Z}}-R_{\mathcal{P}})}-n\tilde{R}_{\mathcal{Y}}\right\}. \tag{134}$$

Consequently, Hoeffding's inequality (Proposition 2) implies that for every fixed $v \in \mathcal{V}\setminus\mathcal{V}_{m_{\mathcal{Y}}}$

$$\begin{aligned}
&\mathbb{P}_{\{\mathcal{V}_v\}}\left[\left|\boldsymbol{P}_V^{(m_{\mathcal{Y}})}(v)-\tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}(v)\right| \ge \xi_n\right] \\
&\quad\overset{(a)}{=} \mathbb{P}_{\{\mathcal{V}_v\}}\left[\frac{1}{|\mathcal{M}_{\mathcal{Z}}|}\sum_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}\mathbb{1}_{v=V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}} \ge \xi_n\right] && (135)\\
&\quad\overset{(b)}{\le} \exp\left\{-2\,|\mathcal{M}_{\mathcal{Z}}|\left(\xi_n-(1-\delta_n)^{-1}\right.\right. \\
&\qquad\left.\left.\times\exp\left\{-(1-\delta_n)e^{n(\tilde{R}_{\mathcal{Y}}+\tilde{R}_{\mathcal{Z}}-R_{\mathcal{P}})}-n\tilde{R}_{\mathcal{Y}}\right\}\right)^2\right\} && (136)\\
&\quad\overset{(c)}{\le} \exp\left\{-|\mathcal{M}_{\mathcal{Z}}|\,\xi_n^2/2\right\}, && (137)
\end{aligned}$$

where ($a$) holds because $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} = \mathcal{V}_{m_{\mathcal{Y}}}$ $\mathbb{P}_{\{\mathcal{V}_v\}}$-almost-surely, because $\mathcal{V}_{m_{\mathcal{Y}}}$ is nonempty (which holds because $\{\mathcal{V}_v\} \in \mathcal{H}_\mu^{\mathcal{Y}}$ implies (111) and by (120a)), by (94), and because $v \notin \mathcal{V}_{m_{\mathcal{Y}}}$; ($b$) follows from Hoeffding's inequality (Proposition 2) and (126); and ($c$) holds by (134). Similarly, for every fixed $v \in \mathcal{V}_{m_{\mathcal{Y}}}$

$$\begin{aligned}
&\mathbb{P}_{\{\mathcal{V}_v\}}\left[\left|\boldsymbol{P}_V^{(m_{\mathcal{Y}})}(v)-\tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}(v)\right| \ge \xi_n\right] \\
&\quad\overset{(a)}{=} \mathbb{P}_{\{\mathcal{V}_v\}}\left[\left|\frac{1}{|\mathcal{M}_{\mathcal{Z}}|}\sum_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}\mathbb{1}_{v=V_{m_{\mathcal{Y}},m_{\mathcal{Z}}}}-\frac{1}{|\mathcal{V}_{m_{\mathcal{Y}}}|}\right| \ge \xi_n\right] \\
&\quad\overset{(b)}{\le} 2\exp\left\{-2\,|\mathcal{M}_{\mathcal{Z}}|\left(\xi_n-(1-\delta_n)^{-1}\right.\right. \\
&\qquad\left.\left.\times\exp\left\{-(1-\delta_n)e^{n(\tilde{R}_{\mathcal{Y}}+\tilde{R}_{\mathcal{Z}}-R_{\mathcal{P}})}-n\tilde{R}_{\mathcal{Y}}\right\}\right)^2\right\} && (138)\\
&\quad\overset{(c)}{\le} 2\exp\left\{-|\mathcal{M}_{\mathcal{Z}}|\,\xi_n^2/2\right\}, && (139)
\end{aligned}$$

where ($a$) holds because $\boldsymbol{\mathcal{V}}_{m_{\mathcal{Y}}} = \mathcal{V}_{m_{\mathcal{Y}}}$ $\mathbb{P}_{\{\mathcal{V}_v\}}$-almost-surely, because $\mathcal{V}_{m_{\mathcal{Y}}}$ is nonempty, by (94), and because $v \in \mathcal{V}_{m_{\mathcal{Y}}}$; ($b$) follows from Hoeffding's inequality (Proposition 2), (131), and the Union-of-Events bound; and ($c$) holds by (134). The Union-of-Events bound, (137), and (139) imply that

$$\begin{aligned}
&\mathbb{P}_{\{\mathcal{V}_v\}}\left[\exists\, v \in \mathcal{V}\colon \left|\boldsymbol{P}_V^{(m_{\mathcal{Y}})}(v)-\tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}(v)\right| \ge \xi_n\right] \\
&\qquad\le 2\,|\mathcal{V}|\exp\left\{-|\mathcal{M}_{\mathcal{Z}}|\,\xi_n^2/2\right\}. && (140)
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathbb{P}_{\{\mathcal{V}_v\}}\left[d\left(\boldsymbol{P}_V^{(m_{\mathcal{Y}})},\tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}\right) \ge |\mathcal{V}|\,\xi_n/2\right] \\
&\quad\overset{(a)}{=} \mathbb{P}_{\{\mathcal{V}_v\}}\left[\sum_{v\in\mathcal{V}}\left|\boldsymbol{P}_V^{(m_{\mathcal{Y}})}(v)-\tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}(v)\right| \ge |\mathcal{V}|\,\xi_n\right] \\
&\quad\le \mathbb{P}_{\{\mathcal{V}_v\}}\left[\exists\, v \in \mathcal{V}\colon \left|\boldsymbol{P}_V^{(m_{\mathcal{Y}})}(v)-\tilde{\boldsymbol{P}}_V^{(m_{\mathcal{Y}})}(v)\right| \ge \xi_n\right] \\
&\quad\overset{(b)}{\le} 2\,|\mathcal{V}|\exp\left\{-|\mathcal{M}_{\mathcal{Z}}|\,\xi_n^2/2\right\}, \quad \{\mathcal{V}_v\} \in \mathcal{H}_\mu^{\mathcal{Y}}, && (141)
\end{aligned}$$

where ($a$) holds by definition of the Total-Variation distance; and ($b$) holds by (140).

Having obtained (141) for every fixed $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$, we are now ready to tackle the maximum over $m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}$ and prove (119): By (84b), (85e), (132), and (133) there must exist a positive constant $\tau > 0$ and some $\eta_0 \in \mathbb{N}$ for which

$$|\mathcal{V}|\,\xi_n/2 \le e^{-n\tau}, \quad n \ge \eta_0. \tag{142}$$

For every $\tau > 0$ and $\eta_0 \in \mathbb{N}$ satisfying (142) and for all $n$ exceeding $\eta_0$

$$\begin{aligned}
&\max_{\{\mathcal{V}_v\}\in\mathcal{H}_\mu^{\mathcal{Y}}}\mathbb{P}_{\{\mathcal{V}_v\}}\left[\exists\, m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}\colon d\left(\boldsymbol{P}_V^{(m_{\mathcal{Y}})},\boldsymbol{U}_V^{(m_{\mathcal{Y}})}\right) \ge e^{-n\tau}\right] \\
&\quad\overset{(a)}{\le} \max_{\{\mathcal{V}_v\}\in\mathcal{H}_\mu^{\mathcal{Y}}}\mathbb{P}_{\{\mathcal{V}_v\}}\left[\exists\, m_{\mathcal{Y}} \in \mathcal{M}_{\mathcal{Y}}\colon d\left(\boldsymbol{P}_V^{(m_{\mathcal{Y}})},\boldsymbol{U}_V^{(m_{\mathcal{Y}})}\right)\right. \\
&\qquad\left.\ge |\mathcal{V}|\,\xi_n/2\right] && (143)\\
&\quad\overset{(b)}{\le} \max_{\{\mathcal{V}_v\}\in\mathcal{H}_\mu^{\mathcal{Y}}}\sum_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}}\mathbb{P}_{\{\mathcal{V}_v\}}\left[d\left(\boldsymbol{P}_V^{(m_{\mathcal{Y}})},\boldsymbol{U}_V^{(m_{\mathcal{Y}})}\right)\right. \\
&\qquad\left.\ge |\mathcal{V}|\,\xi_n/2\right] && (144)
\end{aligned}$$

$$\overset{(c)}{\leq} 2\,|\mathcal{V}|\,|\mathcal{M}_{\mathcal{Y}}|\exp\Big\{-|\mathcal{M}_{\mathcal{Z}}|\exp\big\{-e^{n\kappa}+3\log 2\big\}\Big\} \quad (145)$$

$$\overset{(d)}{\to} 0\ (n\to\infty), \quad (146)$$

where $(a)$ holds by (142), because $n$ exceeds $\eta_0$; $(b)$ follows from the Union-of-Events bound; $(c)$ holds by (141) and (133); and $(d)$ holds because $|\mathcal{V}| = e^{nR_{\mathcal{P}}}$, $|\mathcal{M}_{\mathcal{Y}}| = \exp(\exp(nR_{\mathcal{Y}}))$, $|\mathcal{M}_{\mathcal{Z}}| = \exp(\exp(nR_{\mathcal{Z}}))$, and by (132). □

### B. The Converse Part of Theorem 10

In this section we prove a strong converse to Theorem 10:

*Claim 15: For every rate-pair $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$, every positive constants $\lambda_1^{\mathcal{Y}}$, $\lambda_2^{\mathcal{Y}}$, $\lambda_1^{\mathcal{Z}}$, $\lambda_2^{\mathcal{Z}}$ satisfying*

$$\lambda_1^{\mathcal{Y}} + \lambda_2^{\mathcal{Y}} + \lambda_1^{\mathcal{Z}} + \lambda_2^{\mathcal{Z}} < 1, \quad (147)$$

*and every $\epsilon > 0$ there exists some $\eta_0 \in \mathbb{N}$ so that, for every blocklength $n \geq \eta_0$, every size-$\exp(\exp(nR_{\mathcal{Y}}))$ set $\mathcal{M}_{\mathcal{Y}}$ of possible ID messages for Receiver $\mathcal{Y}$, and every size-$\exp(\exp(nR_{\mathcal{Z}}))$ set $\mathcal{M}_{\mathcal{Z}}$ of possible ID messages for Receiver $\mathcal{Z}$, a necessary condition for 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 for the BC $W(y,z|x)$ to exist is that for some PMF $P$ on $\mathcal{X}$*

$$R_{\mathcal{Y}} < I(P, W_{\mathcal{Y}}) + \epsilon, \quad (148a)$$

$$R_{\mathcal{Z}} < I(P, W_{\mathcal{Z}}) + \epsilon. \quad (148b)$$

To prove Claim 15, we recall from Remark 9 that the following two conditions are necessary and sufficient for some collection of tuples

$$\big\{Q_{m_{\mathcal{Y}},m_{\mathcal{Z}}}, \mathcal{D}_{m_{\mathcal{Y}}}, \mathcal{D}_{m_{\mathcal{Z}}}\big\}_{(m_{\mathcal{Y}},m_{\mathcal{Z}})\in\mathcal{M}_{\mathcal{Y}}\times\mathcal{M}_{\mathcal{Z}}}$$

to be 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 for the BC $W(y,z|x)$: 1) $\big\{Q_{m_{\mathcal{Y}}}, \mathcal{D}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}}$ is an $\big(n, \mathcal{M}_{\mathcal{Y}}, \lambda_1^{\mathcal{Y}}, \lambda_2^{\mathcal{Y}}\big)$ ID code for the marginal channel $W_{\mathcal{Y}}(y|x)$; and 2) $\big\{Q_{m_{\mathcal{Z}}}, \mathcal{D}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ is an $\big(n, \mathcal{M}_{\mathcal{Z}}, \lambda_1^{\mathcal{Z}}, \lambda_2^{\mathcal{Z}}\big)$ ID code for $W_{\mathcal{Z}}(z|x)$, where $\big\{Q_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}}$ and $\big\{Q_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ are defined in (82). We shall use these conditions to establish Claim 15 following Han and Verdú's proof of the strong converse for identification via the DMC [3]. To that end we shall need some terminology and results from [3]. We begin with the following two definitions from [3]:

*Definition 16: An $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{M}}$ for the DMC $W(y|x)$ is* homogeneous *if for every $n$-type $P$ on $\mathcal{X}^n$*

$$Q_m\Big(\mathcal{T}_P^{(n)}\Big) = \frac{1}{|\mathcal{M}|}\sum_{\nu\in\mathcal{M}} Q_\nu\Big(\mathcal{T}_P^{(n)}\Big), \quad m\in\mathcal{M}. \quad (149)$$

*Definition 17: Given an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{M}}$ for the DMC $W(y|x)$, define for every $n$-type $P$ on $\mathcal{X}^n$ and $m\in\mathcal{M}$ the PMF*

$$Q_m^{(n,P)}(\mathbf{x}) = \begin{cases} \frac{Q_m(\mathbf{x})}{Q_m\left(\mathcal{T}_P^{(n)}\right)} & \text{if } \mathbf{x}\in\mathcal{T}_P^{(n)} \text{ and } Q_m\Big(\mathcal{T}_P^{(n)}\Big) > 0, \\ \frac{1}{\left|\mathcal{T}_P^{(n)}\right|} & \text{if } \mathbf{x}\in\mathcal{T}_P^{(n)} \text{ and } Q_m\Big(\mathcal{T}_P^{(n)}\Big) = 0, \\ 0 & \text{if } \mathbf{x}\notin\mathcal{T}_P^{(n)}. \end{cases}$$

*The ID code is $L$-*regular *if for every $n$-type $P$ on $\mathcal{X}^n$ and $m\in\mathcal{M}$ satisfying $Q_m\big(\mathcal{T}_P^{(n)}\big) > 0$ the PMF $Q_m^{(n,P)}$ on $\mathcal{T}_P^{(n)}$ is an $L$-type.*

Following the line of argument in [3], we shall construct from $\big\{Q_{m_{\mathcal{Y}}}, \mathcal{D}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}}$ and $\big\{Q_{m_{\mathcal{Z}}}, \mathcal{D}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ homogeneous $L$-regular ID codes. For the construction we shall need Proposition 18 and Lemma 19 ahead. Proposition 18 is a variation on [3, Proposition 3], and Lemma 19 is a generalization of [3, Lemma 1] similar to that in [14, Lemma 2].

*Proposition 18: For every $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{M}}$ for the DMC $W(y|x)$ and for every $\delta \geq \log 2/n$ there exists a subset $\mathcal{S}$ of $\mathcal{M}$ with*

$$|\mathcal{S}| \geq |\mathcal{M}|\exp\Big\{-e^{\log(1+n)(1+|\mathcal{X}|)+\log\delta}\Big\} \quad (150)$$

*for which we can construct from $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{S}}$ a homogeneous $(n, \mathcal{S}, \lambda_1', \lambda_2')$ ID code $\{Q_m', \mathcal{D}_m\}_{m\in\mathcal{S}}$ for $W(y|x)$ with*

$$\lambda_1' = \lambda_1 + e^{-n\delta+\log(1+n)\,|\mathcal{X}|}, \quad (151a)$$

$$\lambda_2' = \lambda_2 + e^{-n\delta+\log(1+n)\,|\mathcal{X}|}. \quad (151b)$$

*Moreover, if for some $\epsilon, \kappa > 0$*

$$Q_m\Big(X^n \in \big\{\mathbf{x}\in\mathcal{X}^n\colon I(P_{\mathbf{x}}, W) \leq R-\epsilon\big\}\Big) \geq \kappa, \quad m\in\mathcal{M}, \quad (152)$$

*then*

$$Q_m'\Big(X^n \in \big\{\mathbf{x}\in\mathcal{X}^n\colon I(P_{\mathbf{x}}, W) \leq R-\epsilon\big\}\Big) \geq \kappa, \quad m\in\mathcal{S}. \quad (153)$$

*Proof:* The proof is essentially that of [3, Proposition 3]. Additionally, we observe the following: if the PMFs $\{Q_m\}_{m\in\mathcal{M}}$ satisfy (152), then the PMFs $\{Q_m'\}_{m\in\mathcal{S}}$, which are constructed in the proof of [3, Proposition 3], satisfy (153). A proof can be found in [8, Appendix A.2] and in Section 1 of the supplementary material. □

*Lemma 19: For every DMC $W(y|x)$ there exists a positive constant $\delta_0 > 0$, which depends only on $|\mathcal{Y}|$, and a continuous, strictly-increasing function $\rho\colon [0,\delta_0] \to \mathbb{R}_0^+$ with $\rho(0) = 0$ so that, for every $\delta\in(0,\delta_0]$, every $\epsilon\in(0,1)$, and every blocklength $n \geq \eta_0$ (where $\eta_0\in\mathbb{N}$ depends only on $|\mathcal{X}|$, $|\mathcal{Y}|$, $\delta$, and $\epsilon$), it holds that for every $n$-type $P$ on $\mathcal{X}^n$, every PMF $Q$ on $\mathcal{T}_P^{(n)} \subseteq \mathcal{X}^n$, every $R \geq I(P,W) + \rho(\delta)$, and every $L = \lceil e^{nR}\rceil$ there exists an $L$-type $Q'$ on $\mathcal{T}_P^{(n)}$ that satisfies for every subset $\mathcal{D}$ of $\mathcal{Y}^n$*

$$(Q'W^n)(Y^n\in\mathcal{D}) \leq (1+\epsilon)(1-e^{-n\delta})^{-1}(QW^n)(Y^n\in\mathcal{D}) + e^{-n\delta}, \quad (154a)$$

$$(Q'W^n)(Y^n\in\mathcal{D}) \geq (1-\epsilon)(1-e^{-n\delta})(QW^n)(Y^n\in\mathcal{D}) - e^{-n\delta}. \quad (154b)$$

*Proof:* The proof is essentially that of [3, Lemma 1] with the differences being pointed out in the proof of [14, Lemma 2]. A proof can be found in [8, Appendix A.3] and in Section 2 of the supplementary material. □

Once we have constructed from $\big\{Q_{m_{\mathcal{Y}}}, \mathcal{D}_{m_{\mathcal{Y}}}\big\}_{m_{\mathcal{Y}}\in\mathcal{M}_{\mathcal{Y}}}$ and $\big\{Q_{m_{\mathcal{Z}}}, \mathcal{D}_{m_{\mathcal{Z}}}\big\}_{m_{\mathcal{Z}}\in\mathcal{M}_{\mathcal{Z}}}$ homogeneous $L$-regular ID codes, we shall use the following proposition to upper-bound the number of possible ID messages $|\mathcal{M}_{\mathcal{Y}}|$ and $|\mathcal{M}_{\mathcal{Z}}|$:

*Proposition 20: [3, Proposition 4] Let $\mathcal{M}$ be a finite set and $\lambda_1$, $\lambda_2$ positive constants satisfying $\lambda_1 + \lambda_2 < 1$. Every homogeneous $L$-regular $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code for the DMC $W(y|x)$ satisfies*

$$\log|\mathcal{M}| \leq n(1+n)^{|\mathcal{X}|} L \log|\mathcal{X}|. \tag{155}$$

Once we have upper-bounded $|\mathcal{M}_\mathcal{Y}|$ and $|\mathcal{M}_\mathcal{Z}|$, we shall infer from the upper bounds that for every $\epsilon > 0$ and $n$ sufficiently large the mixture PMF on $\mathcal{X}^n$

$$Q = \frac{1}{|\mathcal{M}_\mathcal{Y}|\,|\mathcal{M}_\mathcal{Z}|} \sum_{(m_\mathcal{Y}, m_\mathcal{Z}) \in \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}} Q_{m_\mathcal{Y}, m_\mathcal{Z}}$$

must assign notable probability mass to some sequence $\mathbf{x} \in \mathcal{X}^n$ that satisfies both $I(P_\mathbf{x}, W_\mathcal{Y}) > R_\mathcal{Y} - \epsilon$ and $I(P_\mathbf{x}, W_\mathcal{Z}) > R_\mathcal{Z} - \epsilon$. This implies Claim 15, because it implies that there must exist some PMF $P$ on $\mathcal{X}$ for which (148) holds.

We next establish Claim 15, proceeding as outlined above. In a first step we shall combine Proposition 18, Lemma 19, and Proposition 20 to obtain the following lemma:

*Lemma 21: For every DMC $W(y|x)$, every ID rate $R$, and every positive constants $\lambda_1$, $\lambda_2$, $\epsilon$, $\kappa$ satisfying $\lambda_1 + \lambda_2 < \kappa < 1$ there exists some $\eta_0 \in \mathbb{N}$ so that, for every blocklength $n \geq \eta_0$ and every size-$\exp(\exp(nR))$ set $\mathcal{M}$ of possible ID messages, a necessary condition for a collection of tuples $\{Q_m, \mathcal{D}_m\}_{m \in \mathcal{M}}$ to be an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code for the DMC $W(y|x)$ is that*

$$\frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} Q_m\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_\mathbf{x}, W) > R - \epsilon\big\}\Big) > 1 - \kappa - \exp\Big\{e^{n(R-\epsilon/2)}\Big\} / \exp\big\{e^{nR}\big\}. \tag{156}$$

*Proof:* See Appendix B. □

With Lemma 21 at hand, we are now ready to conclude the proof of Claim 15 by establishing that for every $\epsilon > 0$ and $n$ sufficiently large the mixture PMF on $\mathcal{X}^n$

$$Q = \frac{1}{|\mathcal{M}_\mathcal{Y}|\,|\mathcal{M}_\mathcal{Z}|} \sum_{(m_\mathcal{Y}, m_\mathcal{Z}) \in \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}} Q_{m_\mathcal{Y}, m_\mathcal{Z}}$$

must assign notable probability mass to some sequence $\mathbf{x} \in \mathcal{X}^n$ that satisfies both $I(P_\mathbf{x}, W_\mathcal{Y}) > R_\mathcal{Y} - \epsilon$ and $I(P_\mathbf{x}, W_\mathcal{Z}) > R_\mathcal{Z} - \epsilon$:

*Proof of Claim 15:* Fix $\kappa^\mathcal{Y}, \kappa^\mathcal{Z} > 0$ that satisfy the following three: 1) $\lambda_1^\mathcal{Y} + \lambda_2^\mathcal{Y} < \kappa^\mathcal{Y}$; 2) $\lambda_1^\mathcal{Z} + \lambda_2^\mathcal{Z} < \kappa^\mathcal{Z}$; and 3) $\kappa^\mathcal{Y} + \kappa^\mathcal{Z} < 1$. (This is possible because of (147).) By Remark 9 and Lemma 21 there must exist some $\eta_0' \in \mathbb{N}$ so that, for every blocklength $n \geq \eta_0'$, every size-$\exp(\exp(nR_\mathcal{Y}))$ set $\mathcal{M}_\mathcal{Y}$ of possible ID messages for Receiver $\mathcal{Y}$, and every size-$\exp(\exp(nR_\mathcal{Z}))$ set $\mathcal{M}_\mathcal{Z}$ of possible ID messages for Receiver $\mathcal{Z}$, the following conditions are necessary for a collection of tuples

$$\big\{Q_{m_\mathcal{Y}, m_\mathcal{Z}}, \mathcal{D}_{m_\mathcal{Y}}, \mathcal{D}_{m_\mathcal{Z}}\big\}_{(m_\mathcal{Y}, m_\mathcal{Z}) \in \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}}$$

to be 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 for the BC $W(y, z|x)$: the mixture PMFs on $\mathcal{X}^n$

$$Q_{m_\mathcal{Y}} = \frac{1}{|\mathcal{M}_\mathcal{Z}|} \sum_{m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}} Q_{m_\mathcal{Y}, m_\mathcal{Z}}, \quad m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y}, \tag{157a}$$

$$Q_{m_\mathcal{Z}} = \frac{1}{|\mathcal{M}_\mathcal{Y}|} \sum_{m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y}} Q_{m_\mathcal{Y}, m_\mathcal{Z}}, \quad m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}, \tag{157b}$$

$$Q = \frac{1}{|\mathcal{M}_\mathcal{Y}|\,|\mathcal{M}_\mathcal{Z}|} \sum_{(m_\mathcal{Y}, m_\mathcal{Z}) \in \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}} Q_{m_\mathcal{Y}, m_\mathcal{Z}} \tag{157c}$$

satisfy

$$Q\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_\mathbf{x}, W_\mathcal{Y}) > R_\mathcal{Y} - \epsilon\big\}\Big) = \frac{1}{|\mathcal{M}_\mathcal{Y}|} \sum_{m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y}} Q_{m_\mathcal{Y}}\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_\mathbf{x}, W_\mathcal{Y}) > R_\mathcal{Y} - \epsilon\big\}\Big) \tag{158}$$

$$> 1 - \kappa^\mathcal{Y} - \exp\big\{e^{n(R_\mathcal{Y} - \epsilon/2)}\big\} / \exp\big\{e^{nR_\mathcal{Y}}\big\} \tag{159}$$

and

$$Q\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_\mathbf{x}, W_\mathcal{Z}) > R_\mathcal{Z} - \epsilon\big\}\Big) = \frac{1}{|\mathcal{M}_\mathcal{Z}|} \sum_{m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}} Q_{m_\mathcal{Z}}\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_\mathbf{x}, W_\mathcal{Z}) > R_\mathcal{Z} - \epsilon\big\}\Big) \tag{160}$$

$$> 1 - \kappa^\mathcal{Z} - \exp\big\{e^{n(R_\mathcal{Z} - \epsilon/2)}\big\} / \exp\big\{e^{nR_\mathcal{Z}}\big\}. \tag{161}$$

The Union-of-Events bound, (159), and (161) imply that

$$Q\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_\mathbf{x}, W_\mathcal{Y}) > R_\mathcal{Y} - \epsilon, \; I(P_\mathbf{x}, W_\mathcal{Z}) > R_\mathcal{Z} - \epsilon\big\}\Big) > 1 - \kappa^\mathcal{Y} - \kappa^\mathcal{Z} - \exp\big\{e^{n(R_\mathcal{Y} - \epsilon/2)}\big\} / \exp\big\{e^{nR_\mathcal{Y}}\big\} - \exp\big\{e^{n(R_\mathcal{Z} - \epsilon/2)}\big\} / \exp\big\{e^{nR_\mathcal{Z}}\big\}. \tag{162}$$

Now let $\eta_0$ be the smallest integer $n \geq \eta_0'$ for which the RHS of (162) is positive (such an $n$ must exist, because $\epsilon > 0$ and $\kappa^\mathcal{Y} + \kappa^\mathcal{Z} < 1$). Then, for every blocklength $n \geq \eta_0$ a necessary condition for (162) to hold is that for some PMF $P$ on $\mathcal{X}$ (148) holds, and hence Claim 15 follows. □

## IV. Extensions

This section discusses two extensions: identification via the BC with more than two receivers (Section IV-A) and identification via the BC with a common message (Section IV-B). Identification via the BC with one-sided feedback is discussed in [8, Sec. 2.5.3].

### A. More Than Two Receivers

In this section we study identification via the BC with more than two receivers. As we shall see, it is easy to adapt the converse of Theorem 10 to this more general scenario, but in the direct part difficulties already arise when the number of receivers increases from two to three. To keep the exposition simple, we shall thus focus on the three-receiver BC. We inner-bound its ID capacity region and show that the bound is in some cases tight.

Consider a three-receiver BC of transition law $W(y_1, y_2, y_3|x)$, and for every $k \in \{1, 2, 3\}$ let $\mathcal{Y}_k$ denote the support of the channel output at Receiver $k$ and $W_k(y_k|x)$

the marginal channel to Receiver $k$. We begin with the basic definitions of an average-error ID code for the BC $W(y_1, y_2, y_3|x)$:

*Definition 22: Fix finite sets $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$, a blocklength $n \in \mathbb{N}$, and positive constants*

$$\lambda_1^{(k)},\ \lambda_2^{(k)},\quad k \in \{1, 2, 3\}.$$

*Associate with every ID message-triple $(m_1, m_2, m_3) \in \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3$ a PMF $Q_{m_1,m_2,m_3}$ on $\mathcal{X}^n$, and for each $k \in \{1, 2, 3\}$ associate with every $m_k \in \mathcal{M}_k$ an ID set $\mathcal{D}_{m_k} \subset \mathcal{Y}_k^n$. Define the mixture PMFs on $\mathcal{X}^n$*

$$Q_{m_1} = \frac{1}{|\mathcal{M}_2|\,|\mathcal{M}_3|} \sum_{m_2,m_3} Q_{m_1,m_2,m_3},\quad m_1 \in \mathcal{M}_1, \tag{163a}$$

$$Q_{m_2} = \frac{1}{|\mathcal{M}_1|\,|\mathcal{M}_3|} \sum_{m_1,m_3} Q_{m_1,m_2,m_3},\quad m_2 \in \mathcal{M}_2, \tag{163b}$$

$$Q_{m_3} = \frac{1}{|\mathcal{M}_1|\,|\mathcal{M}_2|} \sum_{m_1,m_2} Q_{m_1,m_2,m_3},\quad m_3 \in \mathcal{M}_3. \tag{163c}$$

*The collection of tuples*

$$\left\{Q_{m_1,m_2,m_3}, \mathcal{D}_{m_1}, \mathcal{D}_{m_2}, \mathcal{D}_{m_3}\right\}_{(m_1,m_2,m_3)\in\mathcal{M}_1\times\mathcal{M}_2\times\mathcal{M}_3}$$

*is an $\left(n, \{\mathcal{M}_k, \lambda_1^{(k)}, \lambda_2^{(k)}\}_{k\in\{1,2,3\}}\right)$ ID code for the BC $W(y_1, y_2, y_3|x)$ if for each $k \in \{1, 2, 3\}$ the collection of tuples $\left\{Q_{m_k}, \mathcal{D}_{m_k}\right\}_{m_k\in\mathcal{M}_k}$ is an $\left(n, \mathcal{M}_k, \lambda_1^{(k)}, \lambda_2^{(k)}\right)$ ID code for the marginal channel $W_k(y_k|x)$. A rate-triple $(R_1, R_2, R_3)$ is achievable if for every positive $\lambda_1^{(1)}$, $\lambda_2^{(1)}$, $\lambda_1^{(2)}$, $\lambda_2^{(2)}$, $\lambda_1^{(3)}$, and $\lambda_2^{(3)}$ and for every sufficiently-large blocklength $n$ there exists an $\left(n, \{\mathcal{M}_k, \lambda_1^{(k)}, \lambda_2^{(k)}\}_{k\in\{1,2,3\}}\right)$ ID code for the BC with*

$$\begin{cases} \frac{1}{n}\log\log|\mathcal{M}_k| \geq R_k & \text{if } R_k > 0, \\ & \quad k \in \{1, 2, 3\}. \\ |\mathcal{M}_k| = 1 & \text{if } R_k = 0, \end{cases}$$

*The ID capacity region $\mathcal{C}_3$ of the three-receiver BC is the closure of the set of all achievable rate-triples.*

Our next result is an outer bound on the ID capacity region of the three-receiver BC:

*Theorem 23: The ID capacity region $\mathcal{C}_3$ of the BC $W(y_1, y_2, y_3|x)$ is contained in the set $\mathcal{R}_{3\text{-ob}}$ of all rate-triples $(R_1, R_2, R_3) \in (\mathbb{R}_0^+)^3$ that for some PMF $P$ on $\mathcal{X}$ satisfy*

$$R_k \leq I(P, W_k),\ \forall\, k \in \{1, 2, 3\}. \tag{164}$$

*Proof:* The proof follows along the line of argument in Section III-B. It can be found in [8, Appendix A.4] and in Section 3 of the supplementary material. □

We can adapt the two-receiver broadcast ID code of Section III-A to obtain the following inner bound on the ID capacity region of the three-receiver BC:

*Theorem 24: The ID capacity region $\mathcal{C}_3$ of the BC $W(y_1, y_2, y_3|x)$ contains the set $\mathcal{R}_{3\text{-ib}}$ of all rate-triples $(R_1, R_2, R_3) \in (\mathbb{R}_0^+)^3$ that for some PMF $P$ on $\mathcal{X}$ satisfy*

$$R_k \leq \min\left\{I(P, W_k), \sum_{l\in\{1,2,3\}\setminus\{k\}} I(P, W_l)\right\},\quad \forall\, k \in \{1, 2, 3\}. \tag{165}$$

*The interior of $\mathcal{R}_{3\text{-ib}}$ is achieved by codes with deterministic encoders.*

*Proof:* A proof can be found in [8, Appendix A.5] and in Section 4 of the supplementary material. □

By comparing Theorems 10 and 24, we see that to adapt the broadcast ID code of Section III-A to the three-receiver BC we additionally need the constraints

$$R_k < \sum_{l\in\{1,2,3\}\setminus\{k\}} I(P, W_l),\ \forall\, k \in \{1, 2, 3\}, \tag{166}$$

which have no counterpart in the two-receiver case. We next explain where we use (166). To this end we briefly describe how to extend the random code construction of Section III-A to the three-receiver BC. Fix a PMF $P$ on $\mathcal{X}$, a blocklength $n$, ID rates $R_k$, $k \in \{1, 2, 3\}$, expected bin rates $\tilde{R}_k$, $k \in \{1, 2, 3\}$, and a pool rate $R_{\mathcal{P}}$ satisfying

$$R_k < \tilde{R}_k < \min\left\{I(P, W_k), R_{\mathcal{P}}\right\},\ \forall\, k \in \{1, 2, 3\}. \tag{167}$$

Draw $e^{nR_{\mathcal{P}}}$ $n$-tuples $\sim P^n$ independently, index them, and place them in a pool $\boldsymbol{\mathcal{P}}$. For each receiving terminal $k \in \{1, 2, 3\}$ associate with each ID message $m_k \in \mathcal{M}_k$ a Bin $\boldsymbol{\mathcal{B}}_{m_k}$ by randomly selecting each indexed element of the pool for inclusion in $\boldsymbol{\mathcal{B}}_{m_k}$ independently with probability $e^{-n(R_{\mathcal{P}}-\tilde{R}_k)}$. Associate with every ID message-triple $(m_1, m_2, m_3)$ an $n$-tuple we call the $(m_1, m_2, m_3)$-codeword as follows. If at least one indexed pool-element is contained in all three bins $\boldsymbol{\mathcal{B}}_{m_1}$, $\boldsymbol{\mathcal{B}}_{m_2}$, and $\boldsymbol{\mathcal{B}}_{m_3}$, then draw the $(m_1, m_2, m_3)$-codeword uniformly over the indexed pool-elements that are contained in all three bins. Otherwise draw the $(m_1, m_2, m_3)$-codeword uniformly over the pool. To send ID message-triple $(m_1, m_2, m_3)$, the encoder transmits the $(m_1, m_2, m_3)$-codeword. For each $k \in \mathcal{M}_k$ the $m_k'$-focused party at Terminal $k$ guesses that $m_k'$ was sent if at least one element of the $m_k'$-th bin is jointly typical with the channel outputs that it observes. Therefore, if the $(m_1, m_2, m_3)$-codeword is not an element of Bin $\boldsymbol{\mathcal{B}}_{m_k}$, then the probability that the $m_k$-focused party at Terminal $k$ erroneously guesses that $m_k$ was not sent is high.

Note that for every ID message-triple $(m_1, m_2, m_3)$ the expected number of indexed pool-elements that are contained in all three bins $\boldsymbol{\mathcal{B}}_{m_1}$, $\boldsymbol{\mathcal{B}}_{m_2}$, and $\boldsymbol{\mathcal{B}}_{m_3}$ is $e^{n(\sum_{k=1}^3 \tilde{R}_k - 2R_{\mathcal{P}})}$ $(= e^{nR_{\mathcal{P}}} \prod_{k=1}^3 e^{-n(R_{\mathcal{P}}-\tilde{R}_k)})$, which is smaller than one unless

$$2R_{\mathcal{P}} \leq \sum_{k=1}^3 \tilde{R}_k. \tag{168}$$

Therefore, if (168) does not hold, then with high probability the $(m_1, m_2, m_3)$-codeword is not contained in all three bins $\boldsymbol{\mathcal{B}}_{m_1}$, $\boldsymbol{\mathcal{B}}_{m_2}$, and $\boldsymbol{\mathcal{B}}_{m_3}$, and our scheme will thus fail. This, combined with (167), implies that the code can be reliable only if (166) holds. Note that in the two-receiver scenario the counterpart to (168) is

$$R_{\mathcal{P}} \leq \tilde{R}_{\mathcal{Y}} + \tilde{R}_{\mathcal{Z}}. \tag{169}$$

Unlike (168) in the three-receiver scenario, (169) in the two-receiver scenario can be satisfied by choosing $R_{\mathcal{P}}$ sufficiently small and hence without constraining the rate-pair $(R_{\mathcal{Y}}, R_{\mathcal{Z}})$.

As the following example shows, the inner bound of Theorem 24 need not be tight:

*Example 25: Consider a deterministic BC $W(y_1, y_2, y_3|x)$ with input $X = (X_1, X_2, X_3)$, where for each $k \in \{1, 2, 3\}$ $X_k$ is binary, and with output $Y = (Y_1, Y_2, Y_3)$, where*

$$Y_k = X_k, \quad k \in \{1, 2\}, \tag{170a}$$
$$Y_3 = X. \tag{170b}$$

*For this channel the inner bound $\mathcal{R}_{3\text{-ib}}$ of Theorem 24 evaluates to the set of all rate-triples $(R_1, R_2, R_3) \in (\mathbb{R}_0^+)^3$ that satisfy*

$$R_k \leq \log 2, \ \forall k \in \{1, 2\}, \tag{171a}$$
$$R_3 \leq 2 \log 2. \tag{171b}$$

*Since the BC is deterministic, the encoder can compute all outputs from the inputs that it produces, and the ID capacity region $\mathcal{C}_3$ does thus not increase if the encoder if furnished with perfect feedback. Therefore, Theorem 23 and [18, Corollary 3], which holds under the maximum-error criterion, imply that $\mathcal{C}_3$ is the set of all rate-triples $(R_1, R_2, R_3) \in (\mathbb{R}_0^+)^3$ that satisfy*

$$R_k \leq \log 2, \ \forall k \in \{1, 2\}, \tag{172a}$$
$$R_3 \leq 3 \log 2. \tag{172b}$$

*Consequently, $\mathcal{R}_{3\text{-ib}} \subsetneq \mathcal{C}_3$.*

The inner bound of Theorem 24 is in some cases tight, e.g., if no receiver is "much more capable" than the other two:

*Corollary 26: If the BC $W(y_1, y_2, y_3|x)$ satisfies for every PMF $P$ on $\mathcal{X}$*

$$2 \max_{k \in \{1,2,3\}} I(P, W_k) \leq \sum_{l \in \{1,2,3\}} I(P, W_l), \tag{173}$$

*then its ID capacity region $\mathcal{C}_3$ is the set of all rate-triples $(R_1, R_2, R_3) \in (\mathbb{R}_0^+)^3$ that for some PMF $P$ on $\mathcal{X}$ satisfy* (164)*, and every rate-pair in the interior of $\mathcal{C}_3$ can be achieved using ID codes with deterministic encoders.*

*Proof:* This follows from Theorems 23 and 24, because for such a BC $\mathcal{R}_{3\text{-ob}} = \mathcal{R}_{3\text{-ib}}$. □

### B. A Common Message

In this section we consider the two-receiver BC $W(y, z|x)$ and adapt the coding scheme in Section III-A to solve for the capacity region of a more general scenario where the receivers' ID messages need not be independent but can have a common part. We thus assume that the ID message intended for Terminal $\mathcal{Y}$ is a tuple comprising a private message and a common message, and likewise for Terminal $\mathcal{Z}$. We begin with the basic definitions of an average-error ID code for the BC $W(y, z|x)$ with a common message:

*Definition 27: Fix finite sets $\mathcal{M}$, $\mathcal{M}_\mathcal{Y}$, and $\mathcal{M}_\mathcal{Z}$, a blocklength $n \in \mathbb{N}$, and positive constants $\lambda_1^\mathcal{Y}$, $\lambda_2^\mathcal{Y}$, $\lambda_1^\mathcal{Z}$, $\lambda_2^\mathcal{Z}$. Associate with every ID message-triple $(m, m_\mathcal{Y}, m_\mathcal{Z}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}$ a PMF $Q_{m,m_\mathcal{Y},m_\mathcal{Z}}$ on $\mathcal{X}^n$, with every $(m, m_\mathcal{Y}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Y}$ an ID set $\mathcal{D}_{m,m_\mathcal{Y}} \subset \mathcal{Y}^n$, and with every $(m, m_\mathcal{Z}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Z}$ an ID set $\mathcal{D}_{m,m_\mathcal{Z}} \subset \mathcal{Z}^n$. Define the mixture PMFs on $\mathcal{X}^n$*

$$Q_{m,m_\mathcal{Y}} = \frac{1}{|\mathcal{M}_\mathcal{Z}|} \sum_{m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}} Q_{m,m_\mathcal{Y},m_\mathcal{Z}}, \quad (m, m_\mathcal{Y}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Y}, \tag{174a}$$
$$Q_{m,m_\mathcal{Z}} = \frac{1}{|\mathcal{M}_\mathcal{Y}|} \sum_{m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y}} Q_{m,m_\mathcal{Y},m_\mathcal{Z}}, \quad (m, m_\mathcal{Z}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Z}. \tag{174b}$$

*The collection of tuples*

$$\left\{Q_{m,m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}_{m,m_\mathcal{Y}}, \mathcal{D}_{m,m_\mathcal{Z}}\right\}_{(m,m_\mathcal{Y},m_\mathcal{Z}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}}$$

*is an $\left(n, \mathcal{M}, \mathcal{M}_\mathcal{Y}, \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z}\right)$ ID code for the BC $W(y, z|x)$ with a common message if the following two requirements are met: 1) $\left\{Q_{m,m_\mathcal{Y}}, \mathcal{D}_{m,m_\mathcal{Y}}\right\}_{(m,m_\mathcal{Y}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Y}}$ is an $\left(n, \mathcal{M} \times \mathcal{M}_\mathcal{Y}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y}\right)$ ID code for the marginal channel $W_\mathcal{Y}(y|x)$; and 2) $\left\{Q_{m,m_\mathcal{Z}}, \mathcal{D}_{m,m_\mathcal{Z}}\right\}_{(m,m_\mathcal{Z}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Z}}$ is an $\left(n, \mathcal{M} \times \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z}\right)$ ID code for $W_\mathcal{Z}(z|x)$. A rate-triple $(R, R_\mathcal{Y}, R_\mathcal{Z})$ is achievable if for every positive $\lambda_1^\mathcal{Y}$, $\lambda_2^\mathcal{Y}$, $\lambda_1^\mathcal{Z}$, and $\lambda_2^\mathcal{Z}$ and for every sufficiently-large blocklength $n$ there exists an $\left(n, \mathcal{M}, \mathcal{M}_\mathcal{Y}, \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z}\right)$ ID code for the BC with*

$$\begin{cases} \frac{1}{n} \log \log |\mathcal{M}| \geq R & \text{if } R > 0, \\ |\mathcal{M}| = 1 & \text{if } R = 0, \end{cases}$$
$$\begin{cases} \frac{1}{n} \log \log |\mathcal{M}_\mathcal{Y}| \geq R_\mathcal{Y} & \text{if } R_\mathcal{Y} > 0, \\ |\mathcal{M}_\mathcal{Y}| = 1 & \text{if } R_\mathcal{Y} = 0, \end{cases}$$
$$\begin{cases} \frac{1}{n} \log \log |\mathcal{M}_\mathcal{Z}| \geq R_\mathcal{Z} & \text{if } R_\mathcal{Z} > 0, \\ |\mathcal{M}_\mathcal{Z}| = 1 & \text{if } R_\mathcal{Z} = 0. \end{cases}$$

*The ID capacity region $\mathcal{C}_{\text{cm}}$ of the BC with a common message is the closure of the set of all achievable rate-triples.*

We restrict our analysis to positive ID rates $R_\mathcal{Y}$, $R_\mathcal{Z}$, because if to some receiver we send only the common message, then for the other receiver the imposed average-error criterion will turn into a maximum-error criterion. Theorem 10 allows for the following generalization:

*Theorem 28: The ID capacity region $\mathcal{C}_{\text{cm}}$ of the BC $W(y, z|x)$ with a common message and positive private rates $R_\mathcal{Y}$, $R_\mathcal{Z}$ is the set of all rate-triples $(R, R_\mathcal{Y}, R_\mathcal{Z}) \in (\mathbb{R}_0^+)^3$ that for some PMF $P$ on $\mathcal{X}$ satisfy*

$$R, \ R_\mathcal{Y} \leq I(P, W_\mathcal{Y}), \tag{175a}$$
$$R, \ R_\mathcal{Z} \leq I(P, W_\mathcal{Z}), \tag{175b}$$
$$R_\mathcal{Y}, \ R_\mathcal{Z} > 0. \tag{175c}$$

*It is achieved by codes with deterministic encoders.*

*Proof:* The proof is similar to that of Theorem 10. It can be found in [8, Appendix A.6] and in Section 5 of the supplementary material. □

Comparing Theorems 28 and 10 we see that the common message appears to come for free at all rates up to $\min\left\{I(P, W_\mathcal{Y}), I(P, W_\mathcal{Z})\right\}$. This can be explained as follows. The ID rate is the iterated logarithm of the number of ID messages normalized by the blocklength $n$, and for $n$ sufficiently

large and for all nonnegative real numbers $R_1$ and $R_2$

$$\exp(\exp(nR_1))\exp(\exp(nR_2)) \approx \exp\big(\exp\big(n\max\{R_1, R_2\}\big)\big).$$

So far, we assumed that each receiver identifies the common message and its private message jointly. Next, we assume that each receiver identifies the common message and its private message separately. We begin with the basic definitions of an average-error ID code for the BC $W(y,z|x)$ with a common message and where each receiver identifies the common message and its private message separately:

*Definition 29: Fix finite sets $\mathcal{M}$, $\mathcal{M}_\mathcal{Y}$, and $\mathcal{M}_\mathcal{Z}$, a blocklength $n \in \mathbb{N}$, and positive constants $\lambda_1^\mathcal{Y}$, $\lambda_2^\mathcal{Y}$, $\lambda_1^\mathcal{Z}$, $\lambda_2^\mathcal{Z}$. Associate with every ID message-triple $(m, m_\mathcal{Y}, m_\mathcal{Z}) \in \mathcal{M} \times \mathcal{M}_\mathcal{Y} \times \mathcal{M}_\mathcal{Z}$ a PMF $Q_{m,m_\mathcal{Y},m_\mathcal{Z}}$ on $\mathcal{X}^n$, with every $m \in \mathcal{M}$ ID sets $\mathcal{D}_m^\mathcal{Y} \subset \mathcal{Y}^n$ and $\mathcal{D}_m^\mathcal{Z} \subset \mathcal{Z}^n$, with every $m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y}$ an ID set $\mathcal{D}_{m_\mathcal{Y}} \subset \mathcal{Y}^n$, and with every $m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}$ an ID set $\mathcal{D}_{m_\mathcal{Z}} \subset \mathcal{Z}^n$. Define the mixture PMFs on $\mathcal{X}^n$*

$$Q_m = \frac{1}{|\mathcal{M}_\mathcal{Y}|\,|\mathcal{M}_\mathcal{Z}|}\sum_{m_\mathcal{Y},m_\mathcal{Z}} Q_{m,m_\mathcal{Y},m_\mathcal{Z}}, \quad m \in \mathcal{M},$$

$$Q_{m_\mathcal{Y}} = \frac{1}{|\mathcal{M}|\,|\mathcal{M}_\mathcal{Z}|}\sum_{m,m_\mathcal{Z}} Q_{m,m_\mathcal{Y},m_\mathcal{Z}}, \quad m_\mathcal{Y} \in \mathcal{M}_\mathcal{Y},$$

$$Q_{m_\mathcal{Z}} = \frac{1}{|\mathcal{M}|\,|\mathcal{M}_\mathcal{Y}|}\sum_{m,m_\mathcal{Y}} Q_{m,m_\mathcal{Y},m_\mathcal{Z}}, \quad m_\mathcal{Z} \in \mathcal{M}_\mathcal{Z}.$$

*The collection of tuples*

$$\big\{Q_{m,m_\mathcal{Y},m_\mathcal{Z}}, \mathcal{D}_m^\mathcal{Y}, \mathcal{D}_{m_\mathcal{Y}}, \mathcal{D}_m^\mathcal{Z}, \mathcal{D}_{m_\mathcal{Z}}\big\}_{(m,m_\mathcal{Y},m_\mathcal{Z})\in\mathcal{M}\times\mathcal{M}_\mathcal{Y}\times\mathcal{M}_\mathcal{Z}}$$

*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 for the BC $W(y,z|x)$ with a common message and where each receiver identifies the common message and its private message separately if the following four requirements are met: 1) $\big\{Q_m, \mathcal{D}_m^\mathcal{Y}\big\}_{m\in\mathcal{M}}$ is an $\big(n, \mathcal{M}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y}\big)$ ID code for the marginal channel $W_\mathcal{Y}(y|x)$; 2) $\big\{Q_{m_\mathcal{Y}}, \mathcal{D}_{m_\mathcal{Y}}\big\}_{m_\mathcal{Y}\in\mathcal{M}_\mathcal{Y}}$ is an $\big(n, \mathcal{M}_\mathcal{Y}, \lambda_1^\mathcal{Y}, \lambda_2^\mathcal{Y}\big)$ ID code for $W_\mathcal{Y}(y|x)$; 3) $\big\{Q_m, \mathcal{D}_m^\mathcal{Z}\big\}_{m\in\mathcal{M}}$ is an $\big(n, \mathcal{M}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z}\big)$ ID code for $W_\mathcal{Z}(z|x)$; and 4) $\big\{Q_{m_\mathcal{Z}}, \mathcal{D}_{m_\mathcal{Z}}\big\}_{m_\mathcal{Z}\in\mathcal{M}_\mathcal{Z}}$ is an $\big(n, \mathcal{M}_\mathcal{Z}, \lambda_1^\mathcal{Z}, \lambda_2^\mathcal{Z}\big)$ ID code for $W_\mathcal{Z}(z|x)$. A rate-triple $(R, R_\mathcal{Y}, R_\mathcal{Z})$ is achievable if for every positive $\lambda_1^\mathcal{Y}$, $\lambda_2^\mathcal{Y}$, $\lambda_1^\mathcal{Z}$, and $\lambda_2^\mathcal{Z}$ and for every sufficiently-large blocklength $n$ there exists 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 for the BC with*

$$\begin{cases} \frac{1}{n}\log\log|\mathcal{M}| \geq R & \text{if } R > 0, \\ |\mathcal{M}| = 1 & \text{if } R = 0, \end{cases}$$

$$\begin{cases} \frac{1}{n}\log\log|\mathcal{M}_\mathcal{Y}| \geq R_\mathcal{Y} & \text{if } R_\mathcal{Y} > 0, \\ |\mathcal{M}_\mathcal{Y}| = 1 & \text{if } R_\mathcal{Y} = 0, \end{cases}$$

$$\begin{cases} \frac{1}{n}\log\log|\mathcal{M}_\mathcal{Z}| \geq R_\mathcal{Z} & \text{if } R_\mathcal{Z} > 0, \\ |\mathcal{M}_\mathcal{Z}| = 1 & \text{if } R_\mathcal{Z} = 0. \end{cases}$$

*The ID capacity region $\mathcal{C}_{\text{cm-s}}$ of the BC with a common message and where each receiver identifies the common message and its private message separately is the closure of the set of all achievable rate-triples.*

When each receiver identifies the common message and its private message separately, we can argue similarly as for the three-receiver BC to obtain the following result:

*Theorem 30: The ID capacity region $\mathcal{C}_{\text{cm-s}}$ of the BC $W(y,z|x)$ with a common message and where each receiver identifies the common message and its private message separately is contained in the set of all rate-triples $(R, R_\mathcal{Y}, R_\mathcal{Z}) \in (\mathbb{R}_0^+)^3$ that for some PMF $P$ on $\mathcal{X}$ satisfy*

$$R,\ R_\mathcal{Y} \leq I(P, W_\mathcal{Y}), \tag{177a}$$

$$R,\ R_\mathcal{Z} \leq I(P, W_\mathcal{Z}), \tag{177b}$$

*and it contains the set of all rate-tiples $(R, R_\mathcal{Y}, R_\mathcal{Z}) \in (\mathbb{R}_0^+)^3$ that for some PMF $P$ on $\mathcal{X}$ satisfy* (177) *and*

$$R_\mathcal{Y} \leq 2I(P, W_\mathcal{Z}), \tag{178a}$$

$$R_\mathcal{Z} \leq 2I(P, W_\mathcal{Y}). \tag{178b}$$

*Proof:* Pretend that the common ID message were intended for a third receiver whose marginal channel is time-invariant but can be either $W_\mathcal{Y}(y|x)$ or $W_\mathcal{Z}(z|x)$. To establish the outer and inner bound, respectively, we can now argue as in the proofs of Theorems 23 and 24, which can be found in [8, Appendices A.4 and A.5] and in Sections 3 and 4 of the supplementary material. □

## V. Summary

We have shown that the ID capacity region of the two-receiver BC is the set of rate-pairs for which, for some distribution on the channel input, each receiver's ID rate does not exceed the mutual information between the channel input and the output that it observes (Theorem 10). The capacity region's interior is achieved by codes with deterministic encoders (Remark 12). The results hold under the average-error criterion, which requires that each receiver identify the message intended for it reliably in expectation over the uniform ID message intended for the other receiving terminal. Previously, identification via the BC was studied under the maximum-error criterion, which requires that each receiver identify the message intended for it reliably irrespective of the realization of the ID message intended for the other receiving terminal. Both criteria—average- and maximum-error—consistently extend Ahlswede and Dueck's identification-via-channels problem to the broadcast setting.

The average-error criterion is suitable whenever the receivers' ID messages are independent and uniform over their supports. As we have seen, our coding scheme can be adapted to solve for the capacity region of a more general scenario where the receivers' ID messages are not independent but have a common part (Theorem 28). We also discussed an extension to the BC with more than two receivers (Theorems 23). In particular, we obtained the ID capacity region of the three-receiver BC whenever no receiver is "much more capable" than the other two (Corollary 26).

The question whether for some BCs the average-error ID capacity region can be strictly larger than the maximum-error ID capacity region remains open. We do know that the ID capacity regions differ when only deterministic encoders are allowed: under the average-error criterion deterministic

encoders can achieve every rate-pair in the interior of the ID capacity region (Remark 12), but under the maximum-error criterion they cannot achieve any positive ID rates [1].

## APPENDIX A
## A PROOF OF LEMMA 5

We use the Union-of-Events bound to show that $\mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_m\}_{m\in\mathcal{M}} \notin \mathcal{G}_\mu\big]$ converges to zero. We begin with the events $|\boldsymbol{\mathcal{V}}_m| \le (1-\delta_n)e^{n\tilde{R}}$ and $|\boldsymbol{\mathcal{V}}_{m'}| \ge (1+\delta_n)e^{n\tilde{R}}$. For every $v \in \mathcal{M}$ the binary random variables $\big\{\mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_v}\big\}_{v\in\mathcal{V}}$ are IID, and

$$\mathbb{E}\Big[\sum_{v\in\mathcal{V}} \mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_v}\Big] = \sum_{v\in\mathcal{V}} \mathbb{P}[v \in \boldsymbol{\mathcal{V}}_v] = e^{n\tilde{R}}. \tag{179}$$

Consequently, by the multiplicative Chernoff bounds in Proposition 1,

$$\begin{aligned}
&\mathbb{P}\Big[|\boldsymbol{\mathcal{V}}_m| \le (1-\delta_n)\, e^{n\tilde{R}}\Big] \\
&\quad= \mathbb{P}\Big[\sum_{v\in\mathcal{V}} \mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_m} \le (1-\delta_n)\, e^{n\tilde{R}}\Big] && (180)\\
&\quad\le \exp\big\{-\delta_n^2\, e^{n\tilde{R}-\log 2}\big\} && (181)\\
&\quad= \exp\big\{-e^{n(\tilde{R}-\mu)-\log 2}\big\}, && (182)
\end{aligned}$$

and

$$\mathbb{P}\Big[|\boldsymbol{\mathcal{V}}_{m'}| \ge (1+\delta_n)\, e^{n\tilde{R}}\Big] \le \exp\big\{-e^{n(\tilde{R}-\mu)-\log 3}\big\}. \tag{183}$$

As to $|\boldsymbol{\mathcal{V}}_{m,m'}| \ge e^{n(\tilde{R}-\mu/2)+\log 2}$, note that for every $v \in \mathcal{V}$

$$\mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_{m,m'}} = \mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_m}\, \mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_{m'}},$$

where $\mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_m}$ and $\mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_{m'}}$ are independent because $m \ne m'$. Hence, the binary random variables $\big\{\mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_{m,m'}}\big\}_{v\in\mathcal{V}}$ are IID of mean

$$\begin{aligned}
\mathbb{E}\Big[\sum_{v\in\mathcal{V}} \mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_{m,m'}}\Big] &= \sum_{v\in\mathcal{V}} \mathbb{P}[v \in \boldsymbol{\mathcal{V}}_m]\, \mathbb{P}[v \in \boldsymbol{\mathcal{V}}_{m'}] && (184)\\
&= e^{n(2\tilde{R}-R_{\mathcal{P}})}. && (185)
\end{aligned}$$

Fix some $\xi$ satisfying

$$R_{\mathcal{P}} - \tilde{R} - \mu \le \xi \le R_{\mathcal{P}} - \tilde{R} - \mu/2, \tag{186}$$

and let

$$\kappa_n = e^{n\xi}. \tag{187}$$

Observe that

$$\begin{aligned}
&\mathbb{P}\Big[|\boldsymbol{\mathcal{V}}_{m,m'}| \ge e^{n(\tilde{R}-\mu/2)+\log 2}\Big]\\
&\quad\overset{(a)}{\le} \mathbb{P}\Big[|\boldsymbol{\mathcal{V}}_{m,m'}| \ge e^{n(2\tilde{R}-R_{\mathcal{P}}+\xi)+\log 2}\Big] && (188)\\
&\quad\overset{(b)}{\le} \mathbb{P}\Big[|\boldsymbol{\mathcal{V}}_{m,m'}| \ge (1+\kappa_n)\, e^{n(2\tilde{R}-R_{\mathcal{P}})}\Big] && (189)\\
&\quad= \mathbb{P}\Big[\sum_{v\in\mathcal{V}} \mathbb{1}_{v\in\boldsymbol{\mathcal{V}}_{m,m'}} \ge (1+\kappa_n)\, e^{n(2\tilde{R}-R_{\mathcal{P}})}\Big] && (190)\\
&\quad\overset{(c)}{\le} \exp\big\{-\kappa_n\, e^{n(2\tilde{R}-R_{\mathcal{P}})-\log 3}\big\} && (191)\\
&\quad\overset{(d)}{\le} \exp\big\{-e^{n(\tilde{R}-\mu)-\log 3}\big\}, && (192)
\end{aligned}$$

where $(a)$ holds because (186) implies that $\tilde{R} - R_{\mathcal{P}} + \xi \le -\mu/2$; $(b)$ holds by (187) and because (24) implies that $\mu < R_{\mathcal{P}} - \tilde{R}$, and hence it follows from (186) that $\xi > 0$; $(c)$ follows from the multiplicative Chernoff bound (7) in Proposition 1; and $(d)$ holds by (187) and because (186) implies that $-\mu \le \tilde{R} - R_{\mathcal{P}} + \xi$. The Union-of-Events bound, (182), (183), and (192) imply that

$$\begin{aligned}
&\mathbb{P}\big[\{\boldsymbol{\mathcal{V}}_m\}_{m\in\mathcal{M}} \notin \mathcal{G}_\mu\big]\\
&\quad\le |\mathcal{M}|\Big(\exp\big\{-e^{n(\tilde{R}-\mu)-\log 2}\big\}\\
&\qquad + |\mathcal{M}|\exp\big\{-e^{n(\tilde{R}-\mu)-\log 3}\big\}\Big) && (193)\\
&\quad\overset{(a)}{\to} 0\,(n\to\infty), && (194)
\end{aligned}$$

where $(a)$ holds because $|\mathcal{M}| = \exp(\exp(nR))$ and by (24).

## APPENDIX B
## A PROOF OF LEMMA 21

Choose

$$\gamma = \Big(1 - \frac{\lambda_1+\lambda_2}{\kappa}\Big)/2, \tag{195}$$

and note that $\gamma > 0$. Pick $\delta > 0$ sufficiently small so that it satisfies the requirement in Lemma 19 and so that $\rho(\delta) < \epsilon/2$, where $\rho(\cdot)$ denotes the same function as in Lemma 19, and let $\epsilon' = \rho(\delta)$. We henceforth assume that $n$ is sufficiently large so that the following four inequalities hold:

$$\begin{aligned}
\log 2/n &\le \delta, && (196a)\\
(1+\gamma/4)\big(1-e^{-n\delta}\big)^{-1} + e^{-n\delta} &\le 1+\gamma/2, && (196b)\\
\Big(\lambda_1+\lambda_2+2e^{-n\delta+\log(1+n)\,|\mathcal{X}|}\Big)/\kappa + \gamma &< 1, && (196c)
\end{aligned}$$

and

$$\begin{aligned}
&\exp\Big\{e^{n(R-\epsilon+\epsilon')+\log(1+n)(1+|\mathcal{X}|)+\log\log|\mathcal{X}|}\\
&\qquad + e^{\log(1+n)(1+|\mathcal{X}|)+\log\delta}\Big\} < \exp\Big\{e^{n(R-\epsilon/2)}\Big\}. && (196d)
\end{aligned}$$

Let $\mathcal{M}$ be some size-$\exp(\exp(nR))$ set, and assume that the collection of tuples $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{M}}$ is an $(n, \mathcal{M}, \lambda_1, \lambda_2)$ ID code for the DMC $W(y|x)$. Pick

$$\begin{aligned}
\mathcal{K} = \Big\{m \in \mathcal{M}\colon\, &Q_m\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n\colon I(P_{\mathbf{x}}, W)\\
&\le R - \epsilon\big\}\Big) \ge \kappa\Big\}, && (197)
\end{aligned}$$

and note that $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{K}}$ is an $(n, \mathcal{K}, \lambda_1, \lambda_2)$ ID code for the DMC $W(y|x)$. By (196a), (197), and Proposition 18 there exists a subset $\mathcal{S}$ of $\mathcal{K}$ with

$$|\mathcal{S}| \ge |\mathcal{K}|\exp\Big\{-e^{\log(1+n)(1+|\mathcal{X}|)+\log\delta}\Big\} \tag{198}$$

for which we can construct from $\{Q_m, \mathcal{D}_m\}_{m\in\mathcal{S}}$ a homogeneous $(n, \mathcal{S}, \lambda_1', \lambda_2')$ ID code $\{Q_m', \mathcal{D}_m\}_{m\in\mathcal{S}}$ for $W(y|x)$ with

$$\begin{aligned}
\lambda_1' &= \lambda_1 + e^{-n\delta+\log(1+n)\,|\mathcal{X}|}, && (199a)\\
\lambda_2' &= \lambda_2 + e^{-n\delta+\log(1+n)\,|\mathcal{X}|}, && (199b)
\end{aligned}$$

and

$$Q'_m\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_{\mathbf{x}}, W) \le R - \epsilon\big\}\Big) \ge \kappa, \quad m \in \mathcal{S}.$$

For every $m \in \mathcal{S}$ define the PMF on $\mathcal{X}^n$

$$Q''_m(\mathbf{x}) = \begin{cases} \frac{Q'_m(\mathbf{x})}{Q'_m(X^n \in \{\mathbf{x}' \in \mathcal{X}^n \colon I(P_{\mathbf{x}'}, W) \le R - \epsilon\})} \\ \quad \text{if } I(P_{\mathbf{x}}, W) \le R - \epsilon, \qquad \mathbf{x} \in \mathcal{X}^n. \\ 0 \quad \text{otherwise,} \end{cases}$$

Let

$$\lambda''_1 = \frac{\lambda'_1}{\kappa} \quad \text{and} \quad \lambda''_2 = \frac{\lambda'_2}{\kappa}, \tag{200}$$

and note that the collection of tuples $\{Q''_m, \mathcal{D}_m\}_{m \in \mathcal{S}}$ is a homogeneous $(n, \mathcal{S}, \lambda''_1, \lambda''_2)$ ID code for $W(y|x)$, because for every distinct pair $m, m' \in \mathcal{S}$

$$(Q''_m W^n)(Y^n \notin \mathcal{D}_m) \le \frac{\lambda'_1}{\kappa} = \lambda''_1, \tag{201a}$$

$$(Q''_m W^n)(Y^n \in \mathcal{D}_{m'}) \le \frac{\lambda'_2}{\kappa} = \lambda''_2. \tag{201b}$$

By Lemma 19 there exists some $\eta'_0 \in \mathbb{N}$, which depends only on $|\mathcal{X}|$, $|\mathcal{Y}|$, $\delta$, and $\gamma$, so that for every $n \ge \eta'_0$ we can, for every $n$-type $P$ on $\mathcal{X}^n$ for which

$$I(P, W) \le R - \epsilon$$

and for every $m \in \mathcal{M}$, approximate the PMF $(Q''_m)^{(n,P)}$ on $\mathcal{T}_P^{(n)}$ by an $e^{n(R-\epsilon+\epsilon')}$-type $(Q'''_m)^{(n,P)}$ on $\mathcal{T}_P^{(n)}$ that satisfies for every subset $\mathcal{D}$ of $\mathcal{Y}^n$

$$\begin{aligned} &\big((Q'''_m)^{(n,P)} W^n\big)(Y^n \in \mathcal{D}) \\ &\quad \le (1 + \gamma/4)(1 - e^{-n\delta})^{-1}\big((Q''_m)^{(n,P)} W^n\big)(Y^n \in \mathcal{D}) \\ &\qquad + e^{-n\delta} \qquad (202) \\ &\quad \le \big((Q''_m)^{(n,P)} W^n\big)(Y^n \in \mathcal{D}) + \gamma/2, \qquad (203) \end{aligned}$$

where in the second inequality we used (196b). For every $m \in \mathcal{S}$ define the PMF

$$Q'''_m(\mathbf{x}) = Q''_m\Big(\mathcal{T}_P^{(n)}\Big)(Q'''_m)^{(n,P)}(\mathbf{x}), \quad P \in \Gamma^{(n)}, \ \mathbf{x} \in \mathcal{T}_P^{(n)}. \tag{204}$$

By (203) it holds for every subset $\mathcal{D}$ of $\mathcal{Y}^n$ that

$$\begin{aligned} &(Q'''_m W^n)(Y^n \in \mathcal{D}) \\ &\quad = \sum_{P \in \Gamma^{(n)}} Q''_m\Big(\mathcal{T}_P^{(n)}\Big)\big((Q'''_m)^{(n,P)} W^n\big)(Y^n \in \mathcal{D}) \qquad (205) \\ &\quad \le \sum_{P \in \Gamma^{(n)}} Q''_m\Big(\mathcal{T}_P^{(n)}\Big)\Big(\big((Q''_m)^{(n,P)} W^n\big)(Y^n \in \mathcal{D}) + \gamma/2\Big) \qquad (206) \\ &\quad = (Q''_m W^n)(Y^n \in \mathcal{D}) + \gamma/2. \qquad (207) \end{aligned}$$

Let

$$\lambda'''_1 = \lambda''_1 + \frac{\gamma}{2} \quad \text{and} \quad \lambda'''_2 = \lambda''_2 + \frac{\gamma}{2}. \tag{208}$$

By (207) and because $\{Q''_m, \mathcal{D}_m\}_{m \in \mathcal{S}}$ is a homogeneous $(n, \mathcal{S}, \lambda''_1, \lambda''_2)$ ID code for $W(y|x)$, the collection of tuples $\{Q'''_m, \mathcal{D}_m\}_{m \in \mathcal{S}}$ is a homogeneous $e^{n(R-\epsilon+\epsilon')}$-regular $(n, \mathcal{S}, \lambda'''_1, \lambda'''_2)$ ID code for $W(y|x)$, and by (196c), (199), and (200)

$$\lambda'''_1 + \lambda'''_2 < 1. \tag{209}$$

Consequently, Proposition 20 implies that

$$\log|\mathcal{S}| \le n(1+n)^{|\mathcal{X}|} e^{n(R-\epsilon+\epsilon')} \log|\mathcal{X}|, \tag{210}$$

and by (198)

$$\begin{aligned} |\mathcal{K}| &\le \exp\Big\{e^{n(R-\epsilon+\epsilon')+\log(1+n)(1+|\mathcal{X}|)+\log\log|\mathcal{X}|} \\ &\qquad + e^{\log(1+n)(1+|\mathcal{X}|)+\log\delta}\Big\} \qquad (211) \\ &< \exp\Big\{e^{n(R-\epsilon/2)}\Big\}, \qquad (212) \end{aligned}$$

where in the second inequality we used (196d). We are now ready to conclude the proof:

$$\begin{aligned} &\frac{1}{|\mathcal{M}|}\sum_{m \in \mathcal{M}} Q_m\Big(X^n \in \big\{\mathbf{x} \in \mathcal{X}^n \colon I(P_{\mathbf{x}}, W) > R - \epsilon\big\}\Big) \\ &\quad \overset{(a)}{>} (1 - \kappa)\frac{|\mathcal{M}| - |\mathcal{K}|}{|\mathcal{M}|} \qquad (213) \\ &\quad \overset{(b)}{>} 1 - \kappa - \exp\Big\{e^{n(R-\epsilon/2)}\Big\}/\exp\Big\{e^{nR}\Big\}, \quad n \ge \eta_0, \qquad (214) \end{aligned}$$

where $(a)$ holds by (197); and $(b)$ holds by (212) when we choose $\eta_0$ to be the smallest integer no smaller than $\eta'_0$ that satisfies (196).

## Acknowledgments

The authors would like to thank the anonymous reviewer and the Associate Editor whose valuable comments helped substantially improve the quality of the paper.

## References

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

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.

[3] T. Han and S. Verdú, "New results in the theory of identification via channels," *IEEE Trans. Inf. Theory*, vol. 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*, vol. 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," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2001, p. 146.

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

[7] R. Ahlswede, "General theory of information transfer: Updated," *Discrete Appl. Math.*, vol. 156, no. 9, pp. 1348–1388, May 2008.

[8] A. Bracher, "Identification and zero-error codes," Ph.D. dissertation, ETH Zurich, Zürich, Switzerland, 2016.

[9] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[10] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Statist. Assoc.*, vol. 58, pp. 13–30, Mar. 1963.

[11] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York, NY, USA: Springer-Verlag, 1998.

[12] S. Ross and E. Peköz, *A Second Course In Probability*. Boston, MA, USA: www.ProbabilityBookstore.com, 2007.

[13] C. Canonne, D. Ron, and R. A. Servedio. (Jan. 2015). "Testing probability distributions using conditional samples." [Online]. Available: https://arxiv.org/abs/1211.2664

[14] R. Ahlswede and Z. Zhang, "New directions in the theory of identification via channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1040–1050, Jul. 1995.

[15] F. M. J. Willems, "The maximal-error and average-error capacity region of the broadcast channel are identical: A direct proof," *Problem Control Inf. Theory*, vol. 19, no. 4, pp. 339–347, 1990.

[16] R. Ahlswede and G. Dueck, "Identification in the presence of feedback—A discovery of new capacity formulas," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 30–36, Jan. 1989.

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

[18] R. Ahlswede and B. Verboven, "On identification via multiway channels with feedback," *IEEE Trans. Inf. Theory*, vol. 37, no. 6, pp. 1519–1526, Nov. 1991.

**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.

**Amos Lapidoth** (S'89–M'95–SM'00–F'04) received the B.A. degree in mathematics (*summa cum* laude, 1986), the B.Sc. degree in electrical engineering (*summa cum laude*, 1986), and the M.Sc. degree in electrical engineering (1990) all from the Technion-Israel Institute of Technology. He received the Ph.D. degree in electrical engineering from Stanford University in 1995.

In the years 1995-1999 he was an Assistant and Associate Professor at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, and was the KDD Career Development Associate Professor in Communications and Technology. He is now Professor of Information Theory at the Swiss Federal Institute of Technology in Zurich. His research interests are in digital communications and information theory.

Dr. Lapidoth served in the years 2003-2004 and 2009 as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.